Fachbereich 10
Mathematik und
Informatik

Master Thesis

Analysis of a Dynamic Cell Imaging Model in Positron Emission Tomography

Marco Jonas Mauritz

Matr. 449734

Supervisor: Prof. Dr. Benedikt Wirth
Assisting Supervisor: Prof. Dr. Caterina Ida Zeppieri

Institut für Analysis und Numerik, University of Münster, Germany

January 17, 2020

Contents

CHAPTER 1

Introduction and Overview

This thesis deals with the mathematical analysis of a dynamic cell imaging model. Based on physical considerations, we derive a reconstruction method that determines the temporal evolution of the distribution of radioactively marked cells from measured PET (positron-emission tomography) data. The reconstruction is obtained by minimizing a certain objective function. We show existence of minimizers of this function and consider Γ-convergence for the case of an intensity of radiation tending to infinity.

The reconstruction method was proposed by Schmitzer, Schäfers and Wirth [1]. Its intended use is to track single or small number of cells. Cell tracking is of special interest in the research field of immunotherapy in order to understand the underlying biological processes better. This kind of therapy uses modified immune cells to specifically destroy harmful cells such as bacteria or cancer cells in the human body [2].

Determining the underlying material distribution form PET data is an ill-posed problem. The proposed method differs from conventional PET reconstruction and is similar to the method of [3] and [4]. It uses the information of all detected events to find the temporal evolution of the radioactive material distribution. Additionally, this information is used to establish temporal consistency between different time steps. This is achieved using optimal transport regularization. Within this regularization approach, among all temporally evolving material distribution being consistent with the PET measurement, the one with least kinetic energy is chosen as the reconstruction. Moreover, the temporal evolution of the material distributions ρ needs to satisfy the continuity equation $\partial_t \rho + \nabla_x \omega = 0$ where ω describes the material flux inducing the temporal variation of ρ. A great advantage of the method considered in this thesis is that the reconstruction complexity is independent of the number of cells to be tracked. Additionally, the number of tracked cells is determined [1].

We model the distribution of the radioactive material, i.e. the labeled cells, using a non-negative Radon measure $dt \otimes \rho_t$ in space-time. The decay of the material is then modeled with a Poisson point process with intensity measure $dt \otimes \frac{1}{T_{1/2}} \rho_t$ where $T_{1/2}$ is the radionuclid's half-life. From physical considerations we derive a forward operator $A = A^s + A^d$ describing the detection process of the emitted photons. Within the modeling we distinguish between scattered photons (described by A^s) and unscattered, hence normally detected ones (described by A^d). The forward operator transforms ρ_t to a new Radon measure κ_t that describes the intensity of the photons at the detectors. The actual detection process is again described by a Poisson point process with intensity measure $dt \otimes \frac{1}{T_{1/2}} \kappa_t$. We use a maximum a posteriori (MAP) estimate to reconstruct the material distribution from a given measurement which results in a minimization problem. Due to the stochastic character of our model, the objective function has a stochastic component, i.e. the measurement is produced in a stochastic way for a given material distribution. The objective function reads

$$J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}(\rho,\omega) = \sum_{ikl} \left[A_{ikl}\rho - E_{ikl} \log\left(pA_{ikl}^{s}\rho + A_{ikl}^{d}\rho \right) \right] + \beta\mathcal{S}(\rho,\omega)$$

where E_{ikl} is the (stochastic) number of photons during the time interval τ_i at detector pair Γ_{kl},

$$A_{ikl}^{/s/d} = \frac{1}{T_{1/2}} \int_{\tau_i} \kappa_t^{/s/d}(\Gamma_{kl})dt$$

and $\mathcal{S}$ is the regularization term that penalizes mass movement.

We show that this functional almost surely has a minimizer in a space of Radon measures $\mathbb{M}$. In the next step we compute two Γ-limits for half-lifes $T_{1/2} \to 0$ which means that the intensity of the radioactive material tends to infinity. For the first Γ-limit we shrink our time intervals. We find for a true underlying material density $dt \otimes \rho_t^\dagger$ (which leads to a measurement process with intensity measure $dt \otimes \frac{1}{T^n}\kappa_t^\dagger$) and for $\mathcal{E}^n = T^n J_{\beta^n,T^n}^{E^n,\Delta t^n,(\Gamma_{kl})_{kl}} + C^n$ we receive the Γ-convergence

$$\Gamma - \lim_{n\to\infty} \mathcal{E}^n = \mathcal{E}^\infty$$

almost surely with

$$\mathcal{E}^\infty(\rho,\omega) = \sum_{kl} \int_0^T \left[\kappa_t(\Gamma_{kl}) - \log\left(p\kappa_t^s(\Gamma_{kl}) + \kappa_t^d(\Gamma_{kl}) \right) \kappa_t^\dagger(\Gamma_{kl}) \right] dt + \beta\mathcal{S}(\rho,\omega).$$

For the second Γ-limit we additionally shrink the detector sizes and end up with

$$\Gamma - \lim_{n\to\infty} \mathcal{E}^n = \mathcal{E}^\infty$$

almost surely with $\mathcal{E}^n = T^n J_{\beta^n,T^n}^{E^n,\Delta t^n,(\Gamma_{kl}^n)_{kl}} + C^n$ and with the limit functional

$$\mathcal{E}^\infty(\rho,\omega) = \int_{\partial\Omega_\delta\times\partial\Omega_\delta} \int_0^T \left(\kappa_t(x) - \log\left(p\kappa_t^s(x) + \kappa_t^d(x) \right) \kappa_t^\dagger(x) \right) dt d\phi(x) + \beta\mathcal{S}(\rho,\omega).$$

The structure of this thesis is as follows. In Chapter 2 we introduce important concepts that will be needed in the following sections. Starting with Section 2.1 we give a short introduction to optimal transport and the Wasserstein metric. Then we consider curves in probability or Radon spaces and their connection to the continuity equation. This results in the Benamou-Brenier formula which will be used for our regularization approach. In the following section we introduce Poisson point processes which are needed for our modeling approach. We define integrals with respect to such processes and give formulas for expectation, variance and higher moments for those integrals. Section 2.3 deals with the physical modeling of the radioactive decay and the detection process of the emitted photons which is

described by the forward operator. Using this forward operator, we derive the reconstruction formula for identifying the distribution of radioactive material, i.e. the labeled cells. The final section of the preliminary chapter shortly introduces Γ-convergence which will be the notion of convergence for our limit consideration $T_{1/2} \to 0$.
Chapter 3 provides the mathematical analysis of the reconstruction formula. First, we examine important properties of the forward operator. Then we prove existence of minimizers of the stochastic objective function $J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}$. We continue by analyzing the limit behavior of $J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}$ for a half-life $T_{1/2}$ tending to zero. This is done by means of Γ-convergence. Within this limit process we first only shrink the time intervals of the detection process. In the next step we shrink the time intervals as well as the detector sizes.

CHAPTER 2

Preliminaries and Model Derivation

2.1 Optimal Transport and the Continuity Equation

Our ill-posed inverse problem will be regularized by means of optimal transport. Within this approach we use the continuity equation $\partial_t \rho_t + \nabla_x (v_t \rho_t) = 0$ to enforce temporal consistency of the reconstruction between different time steps. Therefore, the relevant facts about optimal transport and its connection to the continuity equation are considered in this section. The depictions of this section are mostly oriented at [5] and [6].

Definition 2.1.1 (Radon measure, [7])**.** *Let X be a locally compact and separable metric space, $\mathcal{B}(X)$ its Borel σ-algebra, and consider the measure space $(X, \mathcal{B}(X))$.*

(a) A positive measure on $(X, \mathcal{B}(X))$ is called a Borel measure. If a Borel measure is finite on compact sets, it is called a positive Radon measure. We denote the set of positive Radon measures on $(X, \mathcal{B}(X))$ by $\mathcal{M}_+(X)$.

(b) A measure $\mu \colon \mathcal{B}(X) \to \mathbb{R}^m$ for $m \geq 1$ is said to be a finite Radon measure and the set of finite $\mathbb{R}^m$-valued Radon measures on $(X, \mathcal{B}(X))$ is denoted by $\mathcal{M}(X)^m$.

Definition 2.1.2 (Narrow and weak-* convergence of measures, [5] and [7])**.** *Let X be a locally compact and separable metric space.*

(a) Let μ, $(\mu_n)_n \in \mathcal{M}(X)$ be finite Radon measures. We say that $(\mu_n)_n$ converges narrowly to μ if

$$\lim_{n\to\infty} \int_X \varphi d\mu_n = \int_X \varphi d\mu$$

for all $\varphi \in C_b^0(X)$, the space of all continuous and bounded functions on X.

(b) Let μ, $(\mu_n)_n \in \mathcal{M}(X)$ be finite measures. We say that $(\mu_n)_n$ weak- converges to μ if*

$$\lim_{n\to\infty} \int_X \varphi d\mu_n = \int_X \varphi d\mu$$

for all $\varphi \in C_0(X)$, the space of continuous functions vanishing at infinity. We write $\mu_n \overset{*}{\rightharpoonup} \mu$.

We consider the Kantorovich formulation of optimal transport. In this formulation of optimal transport one tries to find a way to move mass from one distribution to another while trying to minimize the cost of this mass movement.

Problem 2.1.3 (Kantorovich problem, [6]). *Given two metric spaces X, Y, two probability measures $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ and a cost function $c\colon X \times Y \to [0, \infty]$, we consider the problem*

$$\inf \left\{ K(\gamma) := \int_{X\times Y} c d\gamma \mid \gamma \in \Pi(\mu,\nu) \right\}$$

where $\Pi(\mu,\nu)$ is the set of transport plans defined as

$$\Pi(\mu,\nu) := \{\gamma \in \mathcal{P}(X \times Y) \mid (\pi_1)_\#\gamma = \mu,\ (\pi_2)_\#\gamma = \nu\}.$$

Here, $(\pi_1)_\#$ and $(\pi_2)_\#$ are the push-forward measure with respect to the projections of $X\times Y$ onto X and Y.

Remark 2.1.4 ([8]). If we deal with finite positive measures that are not normalized to one, a natural generalization of the transport problem for finite μ, ν with equal mass, i.e. $\mu(X) = \nu(Y)$, would be to consider

$$\hat{\Pi}(\mu,\nu) = \{\gamma \in \mathcal{P}(X \times Y) \mid |\mu|(\pi_x)_\#\gamma = \mu,\ |\nu|(\pi_y)_\#\gamma = \nu\}$$

and

$$|\mu| \inf \left\{ K(\gamma) \mid \gamma \in \hat{\Pi}(\mu,\nu) \right\}$$

which is just a scaling of the situation above. Setting

$$\Pi(\mu,\nu) := \{\gamma \in \mathcal{M}_+(X \times Y) \mid (\pi_x)_\#\gamma = \mu,\ (\pi_y)_\#\gamma = \nu\}$$

and noting that $|\mu|\hat{\Pi}(\mu,\nu) = \Pi(\mu,\nu)$, we can stick to the notation of Problem 2.1.3 by changing the definition of $\Pi(\mu,\nu)$ and are able to work with any finite measures.

Definition 2.1.5 (Wasserstein distance, [6]). *Let $\Omega \subset \mathbb{R}^d$ and let d be a metric on Ω. We set*

$$\mathcal{M}_p := \left\{ \mu \in \mathcal{M}_+(\Omega) \mid \int |x|^p d\mu < +\infty \right\}.$$

Then

$$\begin{aligned} W_p\colon \mathcal{M}_p \times \mathcal{M}_p &\to [0,+\infty) \\ (\mu,\nu) &\mapsto \min \left\{ \int_{\Omega\times\Omega} |x-y|^p d\gamma \mid \gamma \in \Pi(\nu,\mu) \right\}^{\frac{1}{p}} \end{aligned}$$

is called the p-Wasserstein distance on (Ω, d). One can show that this function indeed defines a distance on $\mathcal{M}_p$.
Further, we define the Wasserstein space of order p as $\mathcal{M}_p$ endowed with the distance W_p and denote it by $\mathbb{W}_p^{\mathcal{M}_p}(\Omega)$. In the special case where all measures are normalized to one, we denote by $\mathbb{W}_p^{\mathcal{P}_p}(\Omega)$ the space

$$\mathcal{P}_p(\Omega) := \left\{ \mu \in \mathcal{M}_+(\Omega) \mid \mu(\Omega) = 1 \text{ and } \int |x|^p d\mu < +\infty \right\}$$

endowed with the distance W_p.

Definition 2.1.6 (Continuity equation, [6], [5]). *Let $\Omega \subset \mathbb{R}^d$ be either a bounded subset of $\mathbb{R}^d$ or $\mathbb{R}^d$ itself and let $I = (0,T) \subset \mathbb{R}$ be a time interval, let $(\mu_t)_{t\in I} \subset \mathcal{P}(\Omega)$ be a family of probability measures and let $v\colon \mathbb{R}^d \times I \to \mathbb{R}^d$, $(x,t) \mapsto v_t(x)$ be a Borel velocity field such that*

$$\int_0^T \int_\Omega |v_t(x)|\, d\mu_t(x)dt < +\infty. \tag{2.1}$$

We say that the pair (μ_t, v_t) satisfies the continuity equation

$$\partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0 \tag{2.2}$$

on $\Omega \times (0,T)$ in the sense of distributions if for all $\varphi \in C_c^1(\overline{\Omega} \times (0,T))$ it holds

$$\int_0^T \int_\Omega (\partial_t \varphi(x,t) + \langle \nabla_x \varphi(x,t), v_t(x)\rangle)\, d\mu_t(x)dt = 0. \tag{2.3}$$

Remark 2.1.7. (a) The formulation includes homogeneous Neumann boundary conditions on $\partial\Omega$ for v_t if Ω is not $\mathbb{R}^d$ itself.

(b) In the case $\Omega = \mathbb{R}^d$ it suffices that (2.3) holds for test functions $\varphi \in C_c^\infty(\mathbb{R}^d)$. By a regularization argument we then find that $\varphi \in C_c^1(\mathbb{R}^d \times (0,T))$ is also possible. Therefore, we need to consider $\varphi_\varepsilon = \varphi * \rho_\varepsilon \in C_c^\infty(\mathbb{R}^d \times (0,T))$ for a standard mollifier $(\rho_\varepsilon)_\varepsilon$. In addition, the integrability condition on the velocity field v allows us to consider bounded $C^1(\mathbb{R}^d \times (0,T))$ functions with bounded gradient whose support has a compact projection in $(0,T)$. This time we approximate such a test function φ by $\varphi_R = \varphi\chi_R$ where $\chi_R \in C_c^\infty(\mathbb{R}^d)$ with $0 \le \chi_R \le 1$, $|\nabla \chi_R| \le 2$ and $\chi_R = 1$ on $B_R(0)$ [5, Remark 8.1.1].

The penalty term that will be used to regularize our ill-posed problem is given by the Benamou-Brenier functional $\mathcal{B}$. We are going to prove the relation of this functional to optimal transport, i.e.

$$\begin{aligned} W_p^p(\mu,\nu) &= \min\left\{\mathcal{B}_p(\rho,\omega) \mid \partial_t\rho + \nabla_x \cdot \omega = \delta_0 \otimes \mu - \delta_T \otimes \nu\right\} \\ &= \min\left\{\int_0^1 \int_\Omega |v_t(x)|^p\, d\rho_t(x)dt \mid \partial_t \rho + \nabla\cdot(v_t\rho_t) = 0,\ \rho_0 = \mu, \rho_1 = \nu\right\} \end{aligned} \tag{2.4}$$

(under some suitable assumptions that will be stated later, see Theorem 2.1.16). The important theorem for proving the above formula (2.4) is the following.

Theorem 2.1.8 (Absolutely continuous curves and the continuity equation [5]). *Let I be an open interval in $\mathbb{R}$, let $\mu_t\colon I \to \mathbb{W}_p^{\mathcal{P}_p}(\mathbb{R}^d)$ be an absolutely continuous curve. Then there exists a Borel vector field $v\colon \mathbb{R} \times I \to \mathbb{R}^d$ such that*

$$v_t \in L^p(\mu_t;\mathbb{R}^d) \quad \text{and} \quad \|v_t\|_{L^p(\mu_t;\mathbb{R}^d)} \le |\mu'|(t) \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in I$$

and the continuity equation 2.2 holds in the sense of distributions. Here, $|\mu'|$ denotes the metric derivative of the absolutely continuous curve μ.
Conversely, if a narrowly continuous curve $\mu_t\colon I \to \mathbb{W}_p^{\mathcal{P}_p}(\mathbb{R}^d)$ satisfies the continuity equation (2.2) for some Borel field v with $\|v_t\|_{L^p(\mu_t;\mathbb{R}^d)} \in L^1(I)$, then the curve $(\mu_t)_t$ is absolutely continuous and $|\mu'|(t) \le \|v_t\|_{L^p(\mu_t;\mathbb{R}^d)}$ for $\mathcal{L}^1$-a.e. $t \in I$.

Remark 2.1.9. We stated Theorem 2.1.8 for the continuity equation on $[0,T]\times\mathbb{R}^d$ because smoothing of measures, which is needed in the proof, is easier in this context compared the case of considering the continuity equation on $[0,T]\times\Omega$ for $\Omega\subset\mathbb{R}^d$. Similar results hold true for $\Omega\subset\mathbb{R}^d$ compact, see for example [6, Theorem 5.14].

We start with proving the first statement of Theorem 2.1.8.

Proof: "AC ⇒ existence of vector field". Using a Lipschitz reparametrization of μ_t ([5, Lemma 1.1.4]) and time rescaling of the continuity equation ([5, Lemma 8.1.3]) in the end allows us to assume that $|\mu'|(t)\in L^\infty(I)$.
We start with investigating the behavior of the functions $t\mapsto\mu_t(\varphi)=\int_{\mathbb{R}^d}\varphi d\mu_t$ for $\varphi\in C_c^\infty(\mathbb{R}^d)$. For $s,t\in I$ and $\gamma_{s,t}\in\Pi_0(\mu_s,\mu_t)$ (the set of optimal plans with given marginals μ_s and μ_t) we find, using the Hölder inequality,

$$\begin{aligned}|\mu_t(\varphi)-\mu_s(\varphi)| &= \left|\int_{\mathbb{R}^d}\varphi(y)d((\pi_2)_\#\gamma_{s,t})(y)-\int_{\mathbb{R}^d}\varphi(x)d((\pi_1)_\#\gamma_{s,t})(x)\right|\\ &= \left|\int_{\mathbb{R}^d\times\mathbb{R}^d}(\varphi(y)-\varphi(x))d\gamma_{s,t}(x,y)\right|\\ &\le \operatorname{Lip}(\varphi)W_p(\mu_s,\mu_t)\end{aligned}$$

which shows absolute continuity of $t\mapsto\mu_t(\varphi)$. Now that we know that this function is absolute continuous we can investigate its derivative. Therefore, we define

$$H(x,y):=\begin{cases}|\nabla\varphi(x)| & \text{if } x=y,\\ \frac{|\varphi(x)-\varphi(y)|}{|x-y|} & \text{if } x\neq y\end{cases}$$

which is a bounded and upper semi-continuous function. Using this function, $\gamma_{s,s+h}\in\Pi_0(\mu_s,\mu_{s+h})$ and again Hölder's inequality, we find

$$\begin{aligned}\frac{|\mu_{s+h}(\varphi)-\mu_s(\varphi)|}{|h|} &\le \frac{1}{|h|}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|\,H(x,y)d\gamma_{s,s+h}(x,y)\\ &\le \frac{W_p(\mu_s,\mu_{s+h})}{|h|}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d}H^q(x,y)d\gamma_{s,s+h}(x,y)\right)^{\frac{1}{q}}.\end{aligned}$$

Next, for any point $t\in I$ for which the metric derivative of $s\mapsto\mu_s(\varphi)$ exists, it holds

$$\limsup_{h\to0}\frac{|\mu_{t+h}(\varphi)-\mu_t(\varphi)|}{|h|}\le|\mu'|(t)\left(\int_{\mathbb{R}^d}|H|^q(x,x)d\mu_t(x)\right)^{\frac{1}{q}}=|\mu'|(t)\,\|\nabla\varphi\|_{L^q(\mu_t;\mathbb{R}^d)}\tag{2.5}$$

where we used the narrow convergence $\gamma_{s,s+h}\to(id,id)_\#\mu_t$ as $h\to0$. This convergence is due to the fact that the marginals are narrowly converging and thus any limit point is an element of $\Pi_0(\mu_t,\mu_t)$ and is concentrated on the diagonal of $\mathbb{R}^d\times\mathbb{R}^d$.
Let $Q=\mathbb{R}^d\times I$ and let $\mu=\int\mu_t dt\in\mathcal{P}(Q)$ be the measure having $(\mu_t)_t$ as disintegration. From the mean value theorem we get for any $\varphi\in C_c^\infty(Q)$ and h small enough

$$\frac{|\varphi(x,s)-\varphi(x,s-h)|}{|h|}\le|\partial_t\varphi(x,\xi_h)|\le\sup_{(x,s)\in\operatorname{supp}(\varphi)}|\partial_t\varphi(x,s)|\lesssim1.$$

This shows that $\left(\frac{\varphi(x,s)-\varphi(x,s-h)}{h}\right)_{h>0}$ is uniformly bounded on a compact subset of Q and thus dominated convergence yields

$$\begin{aligned}\int_Q \partial_s\varphi(x,s)d\mu(x,s) &= \lim_{h\downarrow 0}\int_Q \frac{\varphi(x,s)-\varphi(x,s-h)}{h}d\mu(x,s)\\ &= \lim_{h\downarrow 0}\int_I \frac{1}{h}\left(\int_{\mathbb{R}^d}\varphi(x,s)d\mu_s(x) - \int_{\mathbb{R}^d}\varphi(x,s)d\mu_{s+h}(x)\right)ds.\end{aligned}$$

With (2.5) the last expression can be further estimated and we get with Fatou's lemma and Hölder's inequality

$$\begin{aligned}\left|\int_Q \partial_s\varphi(x,s)d\mu(x,s)\right| &\le \limsup_{h\downarrow 0}\int_I \frac{1}{h}\left|\mu_{s+h}(\varphi)-\mu_s(\varphi)\right| ds\\ &\le \int_J |\mu'|(s)\left(\int_{\mathbb{R}^d}|\nabla_x\varphi(x,s)|^q\, d\mu_s(x)\right)^{\frac{1}{q}} ds\\ &\le \left(\int_J |\mu'|^p(s)ds\right)^{\frac{1}{p}}\left(\int_Q |\nabla_x\varphi(x,s)|^q\, d\mu(x,s)\right)^{\frac{1}{q}} \qquad (2.6)\end{aligned}$$

where $J\subset I$ is an interval such that $\operatorname{supp}(\varphi)\subset J\times\mathbb{R}^d$. Next, let us denote by $\mathcal{V}$ the closure of $V=\{\nabla_x\varphi \mid \varphi\in C_c^\infty(Q)\}$ in the space $L^q(\mu;\mathbb{R}^d)$. We define a linear functional

$$L\colon V\to\mathbb{R},\quad L(\nabla\varphi) := -\int_Q \partial_s\varphi(x,s)d\mu(x,s).$$

Due to (2.6) we can uniquely extend L to a bounded linear functional on $\mathcal{V}$.
We consider now the minimization problem

$$\min_{\omega\in\mathcal{V}}\left\{\frac{1}{q}\int_Q |\omega(x,s)|^q\, d\mu(x,s) - L(\omega) =: F(\omega)\right\}$$

and show that it has a unique solution in $\mathcal{V}$. Therefore, let $(\omega_n)_n$ be a minimizing sequence. Then for any $\varphi\in C_c^\infty(Q)$ and n large enough we have

$$\begin{aligned}+\infty > F(\nabla_x\varphi) &\ge \frac{1}{q}\int_Q |\omega_n(x,s)|^q\, d\mu(x,s) - L(\omega_n)\\ &\ge \frac{1}{q}\|\omega_n\|^q_{L^q(\mu;\mathbb{R}^d)} - C\|\omega_n\|_{L^q(\mu;\mathbb{R}^d)}\\ &= \|\omega_n\|_{L^q(\mu;\mathbb{R}^d)}\left(\frac{1}{q}\|\omega_n\|^{q-1}_{L^q(\mu;\mathbb{R}^d)} - C\right),\end{aligned}$$

showing that the minimizing sequence is uniformly bounded (for $\|\omega_n\|_{L^q(\mu;\mathbb{R}^d)} > ((C+\varepsilon)q)^{\frac{1}{q-1}}$, the uniform bound is established by the above estimate and for $\|\omega_n\|_{L^q(\mu;\mathbb{R}^d)} \le ((c+\varepsilon)q)^{\frac{1}{q-1}}$ the norm is bounded anyway). This uniform bound of the norm shows the existence of an $\omega\in\mathcal{V}$ such that $\omega_n\rightharpoonup\omega$ weakly in $L^q(\mu;\mathbb{R}^d)$ along a subsequence. Due to the weak lower semi-continuity of the objective function F we find that ω is a minimizer. Finally, ω is the unique minimizer, since the $\int_Q |\omega(x,s)|^q\, d\mu(x,s)$-part of the objective function is strictly convex. Now that we know that F admits a unique minimizer ω we have that

$$\frac{d}{dt}F(\omega+tv)|_{t=0} = 0$$

for every $v \in \mathcal{V}$. For the Gateaux differential we find

$$\begin{aligned}\frac{d}{dt}F(\omega+tv)|_{t=0} &= \frac{d}{dt}\left(\frac{1}{q}\int_Q |\omega(x,s)+tv(x,s)|^q \, d\mu(x,s) - L(\omega+tv)\right)|_{t=0} \\ &= \int_Q |\omega(x,s)|^{q-2} \langle \omega(x,s), v(x,s)\rangle d\mu(x,s) - L(v)\end{aligned}$$

where the change of integration and differentiation is justified by

$$\begin{aligned}|\omega+tv|^q \leq (|\omega|+|v|)^q \in L^1(\mu) \quad \text{and} \\ \left||\omega|^{q-2}\langle \omega, v\rangle\right| \leq |\omega|^{q-1}|v| \in L^1(\mu)\end{aligned}$$

for $|t| \leq 1$. Using the function

$$j_q\colon L^q(\mu;\mathbb{R}^d) \to L^p(\mu;\mathbb{R}^d), \quad f \mapsto \begin{cases} |f|^{q-2} f & \text{if } f \neq 0, \\ 0 & \text{if } f = 0 \end{cases}$$

for $\frac{1}{p}+\frac{1}{q}=1$ and choosing the vector field $v_t(x) = j_q(\omega)(x,t)$ we get

$$\begin{aligned}0 &= \int_Q \langle \nabla_x \varphi(x,t), v_t(x)\rangle d\mu(x,t) - L(\nabla_x\varphi) \\ &= \int_Q \langle \nabla_x \varphi(x,t), v_t(x)\rangle d\mu(x,t) + \int_Q \partial_t\varphi(x,t) d\mu(x,t)\end{aligned}$$

for every $\varphi \in C_c^\infty(Q)$, showing that (μ_t, v_t) satisfies the continuity equation in the sense of distributions.
We are left to show the norm estimate

$$\|v_t\|_{L^p(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t).$$

Therefore, we take a sequence $(\nabla_x\varphi_n)_n \subset V$ that converges to ω in $L^q(\mu;\mathbb{R}^d)$. Let $\eta \in C_c^\infty(J)$ for an interval $J \subset I$ with $0 \leq \eta \leq 1$. Then the convergence $\nabla_x\varphi_n \to \omega$ yields, using Hölder's inequality and the reversed triangle inequality,

$$\begin{aligned}\left|\int_Q \eta\langle v,\omega\rangle d\mu - \int_Q \eta\langle v, \nabla_x\varphi_n\rangle d\mu\right| &\leq \int_Q |v|\,|\nabla_x\varphi_n - \omega|\, d\mu \\ &\leq \|v\|_{L^p(\mu;\mathbb{R}^d)} \|\nabla_x\varphi_n - \omega\|_{L^q(\mu;\mathbb{R}^d)} \xrightarrow{n\to\infty} 0\end{aligned}$$

and

$$\left|\|\omega\|_{L^q(\mu;\mathbb{R}^d)} - \|\nabla_x\varphi_n\|_{L^p(\mu;\mathbb{R}^d)}\right| \leq \|\omega - \nabla_x\varphi_n\|_{L^p(\mu;\mathbb{R}^d)} \xrightarrow{n\to\infty} 0.$$

For the function j_q we have the identities [5, section 8.3]

$$\begin{aligned}\omega = j_p(v) \quad &\Longleftrightarrow \quad v = j_q(\omega) \quad \text{and} \\ \|j_p(v)\|^q_{L^q(\mu;\mathbb{R}^d)} &= \|v\|^p_{L^p(\mu;\mathbb{R}^d)} = \int_Q \langle j_p(v), v\rangle d\mu.\end{aligned}$$

Taking everything together we get

$$\begin{aligned}
&\int_Q \eta(s)\left|v(x,s)\right|^p d\mu(x,s) = \int_Q \eta\left|v\right|^{p-2}\langle v,v\rangle d\mu\\
&= \int_Q \eta\langle v,\omega\rangle d\mu = \lim_{n\to\infty}\int_Q \eta\langle v,\nabla_x\varphi_n\rangle d\mu\\
&= \lim_{n\to\infty} L(\nabla_x(\eta\varphi_n)) \le \left(\int_J \left|\mu'\right|^p(s)ds\right)^{\frac{1}{p}} \lim_{n\to\infty}\left(\int_{\mathbb{R}^d\times J}\left|\nabla_x(\eta\varphi_n)\right|^q d\mu\right)^{\frac{1}{q}}\\
&\le \left(\int_J \left|\mu'\right|^p(s)ds\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}^d\times J}\left|\omega\right|^q d\mu\right)^{\frac{1}{q}}\\
&\le \left(\int_J \left|\mu'\right|^p(s)ds\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}^d\times J}\left|v\right|^p d\mu\right)^{\frac{1}{q}}.
\end{aligned}$$

Finally, taking a smooth approximation $(\eta_n)_n \subset C_c^\infty(J)$ of the characteristic function of J yields

$$\int_J\int_{\mathbb{R}^d}\left|v_s(x)\right|^p d\mu_s(x)ds \le \int_J \left|\mu'\right|^p(s)ds$$

and therefore

$$\|v_t\|_{L^p(\mu_t;\mathbb{R}^d)} \le \left|\mu'\right|(t)$$

for $\mathcal{L}^1$-a.e. $t\in I$. □

The converse implication of Theorem 2.1.8 will be proven using regularization of measures and a relation between solutions of an ordinary differential equation and distributional solutions of the continuity equation. Therefore, we start by collecting some auxiliary results from [5, Section 8.1] first.

Proposition 2.1.10. *Let μ_t, $t\in[0,T]$, be a narrowly continuous family of Borel probability measures solving the continuity equation w.r.t a Borel vector field v_t satisfying*

$$\int_0^T\left(\sup_B |v_t| + \mathrm{Lip}(v_t,B)\right)dt < \infty \quad \text{for every compact set } B\subset\mathbb{R}^d$$

and satisfying (2.1). *Then for μ_0-a.e. $x\in\mathbb{R}^d$ the characteristic system*

$$X_s(x,s)=x,\ \ \frac{d}{dt}X_t(x,s)=v_t(X_t(x,s))$$

admits a globally defined solution $X_t(x)$ in $[0,T]$ and

$$\mu_t = (X_t)_\#\mu_0 \quad \forall t\in[0,T].$$

Lemma 2.1.11 (Approximation by regular curves, [5])**.** *Let $p\ge 1$ and let μ_t be a time continuous solution of the continuity equation w.r.t. a velocity field satisfying the p-integrability condition*

$$\int_0^T\int_{\mathbb{R}^d}\left|v_t(x)\right|^p d\mu_t(x)dt < +\infty.$$

Let $(\eta_\varepsilon)_\varepsilon \subset C^\infty(\mathbb{R}^d)$ *be a family of strictly positive mollifiers (e.g.* $\rho_\varepsilon(x) = (2\pi\varepsilon)^{-d/2}$ $\exp(-|x|^2/2\varepsilon)$*), and set*

$$\mu_t^\varepsilon := \mu_t * \rho_\varepsilon,\ E_t^\varepsilon := (v_t\mu_t) * \rho_\varepsilon,\ v_t^\varepsilon := \frac{dE_t^\varepsilon}{d\mu_t^\varepsilon}.$$

Then μ_t^ε *is a continuous solution of the continuity equation w.r.t.* v_t^ε*, which satisfy the local regularity assumption*

$$\int_0^T \left(\sup_B |v_t^\varepsilon| + \mathrm{Lip}(v_t^\varepsilon, B)\right) dt < \infty \quad \text{for every compact set } B \subset \mathbb{R}^d$$

and the uniform integrability bounds

$$\int_{\mathbb{R}^d} |v_t^\varepsilon(x)|^p \, d\mu_t^\varepsilon(x) \leq \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) \ \forall t \in (0,T).$$

Moreover, $E_t^\varepsilon \to v_t\mu_t$ *narrowly and*

$$\lim_{\varepsilon\downarrow 0} \|v_t^\varepsilon\|_{L^p(\mu_t^\varepsilon;\mathbb{R}^d)} = \|v_t\|_{L^p(\mu_t;\mathbb{R}^d)} \ \ \forall t \in (0,T).$$

Lemma 2.1.12. *Let* $p \geq 1$*,* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *and let* E *be an* $\mathbb{R}^m$*-valued measure in* $\mathbb{R}^d$ *with finite total variation and absolutely continuous w.r.t.* μ*. Then*

$$\int_{\mathbb{R}^d} \left|\frac{d(E*\rho)}{d(\mu*\rho)}\right|^p \mu * \rho dx \leq \int_{\mathbb{R}^d} \left|\frac{dE}{d\mu}\right|^p d\mu$$

for any convolution kernel ρ*.*

Now we are ready to prove the second part of Theorem 2.1.8.

Proof: "Existence of vector field $\Rightarrow$ μ *is absolutely continuous".* We start with applying Lemma 2.1.11 in order to get regular curves μ_t^ε, v_t^ε satisfying the continuity equation. From Proposition 2.1.10 we get the representation

$$\mu_t^\varepsilon = (T_t^\varepsilon)_\# \mu_0^\varepsilon$$

where T_t^ε is the maximal solution of the ODE $\frac{d}{dt}T_t^\varepsilon = v_t^\varepsilon(T_t^\varepsilon)$ with initial condition $T_0^\varepsilon = x$ (we refer to Section 8.1 in [5] for further details). We find for $t_1 \leq t_2 \in [0,T]$, using Hölder's inequality, Fubini and Lemma 2.1.12,

$$\begin{aligned}
\int_{\mathbb{R}^d} |T_{t_2}^\varepsilon(x) - T_{t_1}^\varepsilon(x)|^p \, d\mu_0^\varepsilon(x) &\leq \int_{\mathbb{R}}^d \left|\int_{t_1}^{t_2} \frac{d}{dt}T_t^\varepsilon(x) dt\right|^p d\mu_0^\varepsilon(x) \\
&\leq (t_2 - t_1)^{p-1} \int_{\mathbb{R}^d} \int_{t_1}^{t_2} \left|\frac{d}{dt}T_t^\varepsilon(x)\right|^p dt d\mu_0^\varepsilon(x) \\
&= (t_2 - t_1)^{p-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |v_t^\varepsilon(T_t^\varepsilon(x))|^p \, d\mu_0^\varepsilon(x) dt \\
&= (t_2 - t_1)^{p-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |v_t^\varepsilon(x)|^p \, d\mu_t^\varepsilon(x) dt \\
&\leq (t_2 - t_1)^{p-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) dt
\end{aligned}$$

where the use of Fubini is justified by the uniform integrability bounds on the approximation v_t^ε and $\|v_t\|_{L^p(\mu;\mathbb{R}^d)} \in L^1(I)$, i.e. we have

$$\int_I \int_{\mathbb{R}^d} |v_t^\varepsilon(x)|^p \, d\mu_t^\varepsilon(x) dt \leq \int_I \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) dt < +\infty.$$

Consider now the transport plan $\gamma^\varepsilon := (T_{t_1}^\varepsilon, T_{t_2}^\varepsilon)_{\#}\mu_0^\varepsilon$. The above calculation shows

$$\begin{aligned} W_p^p(\mu_{t_1}^\varepsilon, \mu_{t_2}^\varepsilon) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y| \, d\gamma^\varepsilon(x,y) = \int_{\mathbb{R}^d} |T_{t_2}^\varepsilon(x) - T_{t_1}^\varepsilon(x)|^p \, d\mu_0^\varepsilon(x) \\ &\leq (t_2 - t_1)^{p-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) dt. \end{aligned}$$

We have that μ_t^ε converges narrowly to μ_t for all $t \in I$ as $\varepsilon \to 0$. Since the Wasserstein distance is narrowly lower semi-continuous [5, Proposition 7.1.3] we find

$$W_p^p(\mu_{t_1}, \mu_{t_2}) \leq (t_2 - t_1)^{p-1} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) dt.$$

As t_1 and t_2 were arbitrary, this implies the absolute continuity of $(\mu_t)_t$ and for the metric derivative we find for $\mathcal{L}^1$-a.e. $t \in I$

$$\begin{aligned} |\mu'|(t) &= \lim_{h \to 0} \frac{W_p(\mu_{t+h}, \mu_t)}{|h|} \leq \left(\lim_{h \to 0} \frac{1}{|h|} \int_t^{t+h} \int_{\mathbb{R}^d} |v_t(x)|^p \, d\mu_t(x) dt \right)^{\frac{1}{p}} \\ &= \|v_t\|_{L^p(\mu_t, \mathbb{R}^d)}. \end{aligned}$$

□

The last part of this section describes the connection between the Wasserstein distance and the continuity equation. First, we introduce the Benamou-Brenier functional and state some important properties.

Definition 2.1.13 (Benamou-Brenier functional)**.** *Let $p > 1$, $\Omega \subset \mathbb{R}^d$ and*

$$\Phi_p \colon \mathbb{R} \times \mathbb{R}^d \to [0, +\infty], \quad (t,x) \mapsto \begin{cases} \frac{|x|^p}{t^{p-1}} & \text{if } t > 0, \\ 0 & \text{if } (t,x) = (0,0), \\ +\infty & \text{if } t < 0 \text{ or } t = 0, x \neq 0. \end{cases}$$

We define the Benamou-Brenier functional

$$\mathcal{B}_p \colon \mathcal{M}([0,T] \times \Omega) \times \mathcal{M}([0,T] \times \Omega)^d \to [0, +\infty], \quad (\rho, \omega) \mapsto \int_{[0,T] \times \Omega} \Phi_p\left(\frac{d\rho}{d\lambda}, \frac{d\omega}{d\lambda}\right) d\lambda$$

where λ is any nonnegative Borel measure such that $|(\rho, \omega)| \ll \lambda$, i.e. the total variation of the measure (ρ, ω) is absolutely continuous w.r.t. λ [9].

Lemma 2.1.14 (Properties of $\mathcal{B}_p$)**.** *Let $p > 1$ and $\Omega \subset \mathbb{R}^d$.*

(a) The definition of $\mathcal{B}_p$ does not depend on the choice of λ;

(b) *It is $\mathcal{B}_p(\rho,\omega) < +\infty$ only if $\rho \geq 0$ and $\omega \ll \rho$. In this case we can write*

$$\mathcal{B}_p(\rho,\omega) = \int_{[0,T]\times\Omega} \left|\frac{d\omega}{d\rho}\right|^p d\rho;$$

(c) *For $\Omega \subset \mathbb{R}^d$ compact or $\mathbb{R}^d$ itself, $\mathcal{B}_p$ is convex and lower semi-continuous w.r.t. weak-* convergence.*

Proof. (a) Since the function Φ_p from the definition of $\mathcal{B}_p$ is 1-homogeneous, the definition of $\mathcal{B}_p$ does not depend on the choice of λ [9].

(b) To see that $\mathcal{B}_p(\rho,\omega)$ is only finite if $\rho \geq 0$ [6, Proposition 5.18], suppose there would exist a measurable set A, $\lambda(A) > 0$, such that $\frac{d\rho}{d\lambda} < 0$ λ-a.s. on A, i.e. $\rho(A) < 0$. Then

$$\mathcal{B}_p(\rho,\omega) \geq \int_A \Phi_p\left(\frac{d\rho}{d\lambda}, \frac{d\omega}{d\lambda}\right) d\lambda = +\infty$$

yields a contradiction. To see that $\mathcal{B}_p(\rho,\omega)$ is finite only if $\omega \ll \rho$ [6, Proposition 5.18], suppose we had a measurable set $A \subset [0,T]\times\mathbb{R}^d$ with $\rho(A) = 0$, $\lambda(A) \neq 0$ and $\omega(A) \neq 0$. If such a set does not exist then we either have $\omega \ll \rho$ or $\lambda \ll \rho$, and the letter case implies $\omega \ll \lambda \ll \rho$ and hence $\omega \ll \rho$ as well. For such a set A we find $\frac{d\rho}{d\lambda}|_A = 0$ λ-a.s. and $\left|\frac{d\omega}{d\lambda}|_B\right| > 0$ λ-a.s. for some measurable $B \subset A$. With this we find

$$\mathcal{B}_p(\rho,\omega) \geq \int_B \Phi_p\left(\frac{d\rho}{d\lambda}, \frac{d\omega}{d\lambda}\right) d\lambda = +\infty$$

which contradicts our assumption $\mathcal{B}_p(\rho,\omega) < +\infty$.
If it holds $\omega \ll \rho$, then we can write

$$\mathcal{B}_p(\rho,\omega) = \int_{[0,T]\times\Omega} \Phi_p\left(1, \frac{d\omega}{d\rho}\right) d\rho = \int_{[0,T]\times\Omega} \left|\frac{d\omega}{d\rho}\right|^p d\rho.$$

(c) For $\frac{1}{p} + \frac{1}{q} = 1$ we consider the closed convex set

$$K_q = \left\{(a,b) \in \mathbb{R}\times\mathbb{R}^d \mid a + \frac{1}{q}|b|^q \leq 0\right\}$$

and the indicator function

$$\iota_{K_q}(t,x) = \begin{cases} 0 & \text{if } (t,x) \in K_q, \\ +\infty & \text{else.} \end{cases}$$

Then we have for the function Φ_p from the definition of $\mathcal{B}_p$ [6, Lemma 5.17]

$$p(\iota_{K_q})^* = \Phi_p$$

where $(\cdot)^*$ denotes the Fenchel-Legendre conjugate. Since the Fenchel-Legendre conjugate is always convex and lower semi-continuous, we have that Φ_p has these properties. This directly guarantees the convexity of $\mathcal{B}_p$. Its lower semi-continuity w.r.t.

the weak-* convergence follows from [7, Theorem 2.34] which states the lower semi-continuity of functionals of the form $\int f\left(\frac{d\omega}{d\rho}\right) d\rho$. To see this, consider a sequence $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$. If $\liminf_{n\to\infty} \mathcal{B}_p(\rho^n, \omega^n) = +\infty$, then there is nothing to show. We can thus assume w.l.o.g. that $\mathcal{B}(\rho^n, \omega^n) < +\infty$ for all n and $(\mathcal{B}(\rho^n, \omega^n))_n$ is bounded. From part (b) we get that $\omega^n \ll \rho^n$. Then [7, Theorem 2.34 and Example 2.36] shows

$$\int_{[0,T]\times\Omega} \left|\frac{d\omega}{d\rho}\right|^p d\rho \leq \liminf_{n\to\infty} \int_{[0,T]\times\Omega} \left|\frac{d\omega^n}{d\rho^n}\right|^p d\rho^n.$$

Note, that we need to extend the measures by zero to an open set $A \supset [0,T]\times\Omega$ in order to be able to apply the mentioned theorem. Since $[0,T]\times\Omega$ is compact, local weak-* convergence of $(\tilde{\rho}^n, \tilde{\omega}^n)$ on A follows from weak-* convergence of (ρ^n, ω^n) on $[0,T]\times\Omega$. Such an extension for a measure μ is given by

$$\tilde{\mu}\colon \mathcal{B}(A) \to [0, +\infty],\ S \mapsto \mu(S \cap [0,T]\times\Omega).$$

□

Within our reconstruction method we use $\mathcal{B}_p$ to regularize the problem (see Section 2.3.3 and 2.3.4). We are interested in a convex regularization term in order to have a convex objective function which makes the optimization easier. Therefore, we will consider a slightly more general formulation of the continuity equation [10, Definition 1.1.1 without source term] that will basically result in the previous formulation 2.3 within our modeling approach. We consider either a compact subset $\Omega \subset \mathbb{R}^d$ or $\Omega = \mathbb{R}^d$.
Let $\mu, \nu \in \mathcal{M}_+([0,T]\times\Omega)$ with $|\mu| = |\nu|$. We say that the pair $(\rho, \omega) \in \mathbb{M} := \mathcal{M}_+([0,T]\times\Omega) \times \mathcal{M}([0,T]\times\Omega)^d$ satisfies the continuity equation

$$\partial_t \rho + \nabla_x \omega = \delta_0 \otimes \mu - \delta_T \otimes \nu$$

between μ and ν in the distributional sense if for all $\varphi \in C_c^1([0,T]\times\Omega)$ it holds

$$\int_{[0,T]\times\Omega} \partial_t \varphi(t,x) d\rho(t,x) + \int_{[0,T]\times\Omega} \langle \nabla_x \varphi(t,x), d\omega(t,x)\rangle = \int_\Omega \varphi(T,\cdot) d\nu - \int_\Omega \varphi(0,\cdot) d\mu. \tag{2.7}$$

One can remove the time boundary constraints by testing against $\varphi \in C_c^1((0,T)\times\Omega)$.

Lemma 2.1.15 (Conservation of mass and continuous representative, [10]). *Let $\Omega \subset \mathbb{R}^d$ be either compact or $\mathbb{R}^d$ itself. If (ρ, ω) satisfies the continuity equation 2.7 from ρ_0 to ρ_T and $\mathcal{B}_p(\rho, \omega) < +\infty$, then ρ and ω admit a disintegration with respect to the Lebesgue measure in time, i.e. we have $\rho = \rho_t \otimes dt$ and $\omega = \omega_t \otimes dt$ and the weak derivative of the function $(0,T) \to \mathbb{R},\ t \mapsto \rho_t(\Omega)$ is zero. Moreover, we can write*

$$\mathcal{B}_p(\rho, \omega) = \int_0^T \int_\Omega \left|\frac{d\omega_t}{d\rho_t}\right|^p d\rho_t dt.$$

Additionally, there exists a narrowly continuous curve $\tilde{\rho} \in \mathcal{M}([0,T]\times\Omega)$ that dt-a.e. equals ρ_t such that $t \mapsto \tilde{\rho}_t(\Omega)$ is continuous and constant.

Proof. We use functions $\varphi \in C^1([0,T]\times\Omega)$ which are constant in the space variable. This gives us

$$\begin{aligned}(\varphi(T)-\varphi(0))\rho_0(\Omega) &= \varphi(T)\rho_T(\Omega)-\varphi(0)\rho_0(\Omega)\\ &= \int_0^T\int_\Omega \varphi'(t)d\rho(t,x) = \int_0^T \varphi'(t)d(\pi^t_{\#}\rho)(t)\end{aligned}$$

where $\pi^t : [0,T]\times\Omega \to \mathbb{R},\ (t,x)\mapsto t$. The above equation implies that $\pi^t_{\#}\rho = \rho_0(\Omega)dt$. By the disintegration theorem [5, theorem 5.3.1] one always has $\rho = \tilde{\rho}_t \otimes \pi^t_{\#}\rho$ for $\tilde{\rho}_t \in \mathcal{M}(\Omega)$ for all t. Thus, we end up with the decomposition $\rho = \rho_t \otimes dt$ when we set $\rho_t = \rho_0(\Omega)\tilde{\rho}_t$. The disintegration in time can be used to derive that the function $t \mapsto \rho_t(\Omega)$ is constant in time. Let $\varphi \in C_c^\infty((0,T))$. Then, using that ρ satisfies the continuity equation,

$$\int_0^T \varphi'(t)\rho_t(\Omega)dt = \int_0^T\int_\Omega \varphi'(t)d\rho_t dt = 0,$$

implying that the weak derivative of $t\mapsto\rho_t(\Omega)$ is zero.
Next, using the assumption $\mathcal{B}_p(\rho,\omega) < +\infty$, we get $\omega \ll \rho$ by Lemma 2.1.14 which gives us a disintegration w.r.t. the Lebesgue measure in time of ω because of $\rho = \rho_t\otimes dt$. Thus, it holds

$$\mathcal{B}_p(\rho,\omega) = \int_0^T\int_\Omega \left|\frac{d\omega_t}{d\rho_t}\right|^p d\rho_t dt.$$

Finally, [5, Lemma 8.1.2] (if $\Omega=\mathbb{R}^d$) and [10, Proposition 1.1.3] (if Ω is compact) guarantee the existence of a narrowly continuous representative $\tilde{\rho}_t$ for which $t\mapsto\rho_t(\Omega)$ is continuous and constant. □

Theorem 2.1.16 ([5],[6])**.** *Let $\Omega\subset\mathbb{R}^d$ be either compact and convex or $\mathbb{R}^d$ itself. Then for $p\geq 1$ and $\mu,\nu\in\mathcal{M}_+(\Omega)$ with equal mass it holds*

$$\begin{aligned}W_p^p(\mu,\nu) &= T^{p-1}\min\left\{\int_0^T\int_\Omega |v_t(x)|^p\, d\rho_t(x)dt \mid \partial_t\rho_t+\nabla_x(v_t\rho_t)=0,\ \rho_0=\mu,\ \rho_T=\nu\right\}\\ &= T^{p-1}\min\left\{\mathcal{B}_p(\rho,\omega)\mid \partial_t\rho+\nabla_x\omega = \delta_0\otimes\mu-\delta_T\otimes\nu\right\}.\end{aligned}$$

Proof. The statement follows from Theorem 2.1.8 and the fact that we have constant speed geodesics in $\mathcal{P}(\Omega)$ [5, Theorem 7.2.2], [6, Theorem 5.27]. We first assume $\mu,\nu\in\mathcal{P}(\Omega)$. For any absolutely continuous curve $(\rho_t)_t \subset \mathbb{W}_p^{\mathcal{P}_p}(\Omega)$ with $\rho_0=\mu$ and $\rho_T=\nu$ (note, that $\rho_t = \frac{t}{T}\nu+\frac{T-t}{T}\mu$ is such a curve) we have, using Theorem 2.1.8 and Hölder's inequality,

$$W_p^p(\mu,\nu) \leq \left(\int_0^T |\rho'|\,dt\right)^p \leq T^{p-1}\int_0^T |\rho'|^p(t)dt \leq T^{p-1}\int_0^T \|v_t\|^p_{L^p(\rho_t;\Omega)}dt.$$

Next, take a constant speed geodesic

$$\gamma\colon [0,1]\to\mathcal{P}(\Omega)\quad\text{with } \gamma(0)=\mu,\ \gamma(1)=\nu.$$

It holds $W_p(\gamma_t,\gamma_s) = W_p(\gamma_0,\gamma_1)\,|t-s|$ and thus we have $|\gamma'| = W_p(\gamma_0,\gamma_1) = \text{const}$. Applying again Theorem 2.1.8 we find

$$W_p^p(\mu,\nu) = \left(\int_0^1 |\gamma'|\,dt\right)^p = W_p^p(\gamma_0,\gamma_1) = \int_0^1 |\gamma'|^p(t)dt \geq \int_0^1 \|v_t\|^p_{L^p(\gamma_t;\Omega)}dt.$$

Finally, we use rescaling in time [5, Lemma 8.1.3] to achieve a curve $\tilde{\gamma}\colon [0,T] \to \mathcal{P}(\Omega)$. Therefore, consider $f : [0,T] \to [0,1],\ t \mapsto \frac{t}{T}$. Then it holds

$$\begin{aligned}\int_0^1 \int_\Omega |v_t(x)|^p \, d\gamma_t(x)dt &= \frac{1}{T}\int_0^T \int_\Omega \left|v_{f(t)}(x)\right|^p d\gamma_{f(t)}(x)dt \\ &= T^{p-1}\int_0^T \int_\Omega \left|(f^{-1})'(t)v_{f(t)}(x)\right|^p d\gamma_{f(t)}(x)dt\end{aligned}$$

and $(\gamma_{f(t)}, (f^{-1})'(t)v_{f(t)})$ solves the continuity equation on $\Omega \times [0,T]$. In total, this gives us the first desired equation

$$W_p^p(\mu,\nu) = T^{p-1}\min\left\{\int_0^T \int_\Omega |v_t(x)|^p \, d\rho_t(x)dt \mid \partial_t\rho_t + \nabla_x(v_t\rho_t) = 0,\ \rho_0 = \mu,\ \rho_T = \nu\right\}.$$

Switching to the variable $\omega_t = v_t\rho_t$ and taking into account Lemma 2.1.15 yields the second equation

$$W_p^p(\mu,\nu) = T^{p-1}\min\left\{\mathcal{B}_p(\rho,\omega) \mid \partial_t\rho + \nabla_x\omega = \delta_0 \otimes \mu - \delta_T \otimes \nu\right\}.$$

For $\mu, \nu \in \mathcal{M}_+(\Omega)$ we have with the above considerations

$$\begin{aligned}W_p^p(\mu,\nu) &= |\mu|\, W_p^p\left(\frac{\mu}{|\mu|}, \frac{\nu}{|\nu|}\right) \\ &= |\mu|\, T^{p-1}\min\left\{\int_0^T \int_\Omega |v_t(x)|^p \, d\rho_t(x)dt \mid \partial_t\rho_t + \nabla_x(v_t\rho_t) = 0,\ \rho_0 = \frac{\mu}{|\mu|},\ \rho_T = \frac{\nu}{|\nu|}\right\} \\ &= T^{p-1}\min\left\{\int_0^T \int_\Omega |v_t(x)|^p \, d\rho_t(x)dt \mid \partial_t\rho_t + \nabla_x(v_t\rho_t) = 0,\ \rho_0 = \mu,\ \rho_T = \nu\right\} \\ &= T^{p-1}\min\left\{\mathcal{B}_p(\rho,\omega) \mid \partial_t\rho + \nabla_x\omega = \delta_0 \otimes \mu - \delta_T \otimes \nu\right\}.\end{aligned}$$

□

The advantage of considering the minimization problem

$$\min\left\{\mathcal{B}_p(\rho,\omega) \mid \partial_t\rho + \nabla_x\omega = \delta_0 \otimes \mu - \delta_T \otimes \nu\right\}$$

instead of

$$\min\left\{\int_0^T \int_\Omega |v_t(x)|^p \, d\rho_t(x)dt \mid \partial_t\rho_t + \nabla_x(v_t\rho_t) = 0,\ \rho_0 = \mu,\ \rho_T = \nu\right\}$$

is that the first problem is convex with linear constraints and the second one is non-convex ($(x,t) \mapsto t\,|x|^p$ is not convex) with non-linear constraints (due to the term $v_t\rho_t$) [6, Section 6.1].

We want to use optimal transport to regularize the ill-posed problem considered in this thesis and described in Section 2.3.4. Therefore, we introduce

$$\mathcal{S}(\rho,\omega) := \begin{cases} \int_0^T \int_\Omega \left(\frac{d\omega_t}{d\rho_t}\right)^2 d\rho_t dt & \text{if } \rho \geq 0, \omega \ll \rho \text{ and (2.7) holds,} \\ +\infty & \text{else} \end{cases}$$

The reconstruction of the material distribution will be found by minimizing an objective function of the form $(\rho, \omega) \mapsto f(\rho) + \lambda \mathcal{S}(\rho, \omega)$. Due to the constraint that (ρ, ω) should satisfy the continuity equation and since $\mathcal{S}(\rho, \omega)$ needs to be finite at the minimum, Lemma 2.1.15 implies that we can write

$$\mathcal{B}_2(\rho, \omega) = \mathcal{S}(\rho, \omega)$$

in this case.

2.2 Poisson Point Processes

In Section 2.3.4 we derive a reconstruction method that uses PET data to determine a temporally evolving distribution of radioactive material. Due to the stochastic character of radioactive decay we use a stochastic model to describe the important physical processes that lead to the PET data. This is done using Poisson point processes (PPP). Therefore, we introduce Poisson point processes in this section and state the most important properties needed in this thesis. Mainly, this section is taken from [11] and [12].
We start with the definition of a point process. Point processes can be seen as a random collection of sets in some space $\mathbb{X}$. To define this rigorously, let $(\mathbb{X}, \mathcal{X})$ be a measurable space and let $\mathbf{N}_{<\infty}(\mathbb{X}) =: \mathbf{N}_{<\infty}$ denote the space of all measures μ on $\mathbb{X}$ satisfying $\mu(B) \in \mathbb{N}_0$ for all $B \in \mathcal{X}$. Let $\mathbf{N} := \mathbf{N}(\mathbb{X})$ be the space of all measures that can be written as a countable sum of measures from $\mathbf{N}_{<\infty}$. Let further $\mathcal{N} := \mathcal{N}(\mathbf{X})$ denote the σ-algebra generated by the collection of all subsets of $\mathbf{N}$ having the form

$$\{\mu \in \mathbf{N} \mid \mu(B) = k\}, \quad B \in \mathcal{X}, k \in \mathbb{N}_0.$$

This means that $\mathcal{N}$ is the smallest σ-algebra on $\mathbf{N}$ such that $\mu \mapsto \mu(B)$ is measurable for all $B \in \mathcal{X}$.

Definition 2.2.1 (Point process)**.** *A point process on $\mathbb{X}$ is a measurable map*

$$\eta\colon (\Omega, \mathcal{F}) \to (\mathbf{N}, \mathcal{N}),$$

i.e. an $\mathbf{N}$-valued random variable.

For a point process η and $B \in \mathcal{X}$ we call $\eta(B)$ the (random) number of points of η in B. Denote by $\eta(B)$ the mapping $\omega \mapsto \eta(\omega, B) := \eta(\omega)(B)$. Then $\eta(B)$ is a random variable taking values in $\overline{\mathbb{N}}_0 := \mathbb{N}_0 \cup \{+\infty\}$ due to the definition of η and $\mathcal{N}$, i.e. we have

$$\{\eta(B) = k\} = \{\omega \in \Omega \mid \eta(\omega, B) = k\} \in \mathcal{F}, \quad B \in \mathcal{X}, k \in \overline{\mathbb{N}}_0. \tag{2.8}$$

On the other hand, a mapping $\eta\colon \Omega \to \mathbf{N}$ is a point process if (2.8) holds.
An important characteristic of a point process is its mean number of points lying in an measurable set which is characterized by the process' intensity measure:

Definition 2.2.2 (Intensity measure)**.** *The intensity measure of a point process η on $\mathbb{X}$ is the measure λ defined by*

$$\lambda(B) := \mathbb{E}\left[\eta(B)\right], \quad B \in \mathcal{X}.$$

Basic properties of the expectation show that the intensity measure indeed is a measure.

A Poisson point process is a point process with the number of points in a given set being distributed according to a Poisson distribution.

Definition 2.2.3 (Poisson point process)**.** *Let λ be an s-finite measure, i.e. a countable sum of finite measures on $\mathbb{X}$. A Poisson point process with intensity measure λ is a point process η on $\mathbb{X}$ with the properties:*

1. *For every* $B \in \mathcal{X}$ *the distribution of* $\eta(B)$ *is Poisson with parameter* $\lambda(B)$*, i.e.* $\mathbb{P}(\eta(B) = k) = \frac{\lambda(B)^k}{k!} \exp(-\lambda(B))$ *for all* $k \in \mathbb{N}_0$.
2. *For every* $m \in \mathbb{N}$ *and all pairwise disjoint sets* $B_1, \ldots, B_m \in \mathcal{X}$ *the random variables* $\eta(B_1), \ldots, \eta(B_m)$ *are independent.*

Poisson processes with σ-finite intensity measure (which is satisfied in our model as we are dealing with finite measures) on complete separable metric spaces $\mathbb{X}$ are proper, i.e. there exist random elements $X_1, X_2, \ldots$ in $\mathbb{X}$ and an $\overline{\mathbb{N}}_0$-valued random variable κ such that almost surely $\eta = \sum_{n=1}^{\kappa} \delta_{X_n}$ [11, Corollary 6.5]. This way the point process can be interpreted as a countable random set of points in $\mathbb{X}$ (with possible repetitions). Moreover, a Poisson process η on a complete separable metric space $\mathbb{X}$ with s-finite intensity measure λ is simple, i.e. we have $\eta(\omega)(\{x\}) \leq 1$ with probability one, if and only if for its intensity measure it holds $\lambda(\{x\}) = 0$ for all $x \in \mathbb{X}$ [11, Proposition 6.9]. As these characterizations hold for the intensity measures in our model, we are considering simple and proper Poisson processes and thus our considered processes fit the intuition of a point process being a random set of points.

An interesting result on (Poisson) point processes is Campbell's formula which relates the expectation of an integral with respect to a (Poisson) point process to an integral with respect to its intensity measure. We only state the version for Poisson processes here, since this allows us to give a formula for the variance as well. Parts of the following theorem are also valid for more general point processes (see [11], Proposition 2.7).

Theorem 2.2.4 (Campbell's formula [11],[12]). *Let* η *be a Poisson point process on* $(\mathbb{X}, \mathcal{X})$ *with* σ*-finite intensity measure* λ*. Let* $u \colon \mathbb{X} \to \mathbb{R}$ *be measurable and set* $D := \{z \in \mathbb{C} \mid Re(z) < 0\}$*. Then* $S_u := \int u(x)\eta(dx) = \sum_{x \in \eta} u(x)$ *is a random variable and the sum converges absolutely if and only if*

$$\int_{\mathbb{X}} (|u(x)| \wedge 1)\lambda(dx) < +\infty. \tag{2.9}$$

If this condition holds, then for $u \geq 0$ *and* $\theta \in \overline{D}$

$$\mathbb{E}\left[e^{\theta S_u}\right] = \exp\left\{\int_{\mathbb{X}} (e^{\theta u(x)} - 1)\lambda(dx)\right\}. \tag{2.10}$$

For measurable u *the above formula holds for* $\theta \in i\mathbb{R}$*. Moreover, it is*

$$\mathbb{E}\left[\int u(x)\eta(dx)\right] = \int_{\mathbb{X}} u(x)\lambda(dx) \tag{2.11}$$

in the sense that the expectation exists if and only if the integral $\int_{\mathbb{X}} u(x)\lambda(dx)$ *converges. If* (2.11) *converges, then*

$$\mathbb{V}\left[\int u(x)\eta(dx)\right] = \int_{\mathbb{X}} u^2(x)\lambda(dx), \tag{2.12}$$

finite or infinite.

Proof. We start with proving that S_u is a random variable. First, let $u(x) = \mathbb{1}_B(x)$ for some $B \in \mathcal{X}$. We get

$$\int u(x)\eta(dx) = \eta(B)$$

which is measurable (i.e. S_u is a random variable) due to the definition of the σ-algebra $\mathcal{N}$. Extending this result first to measurable simple functions, then to non-negative measurable functions and finally to arbitrary measurable functions by using standard techniques of measure theory shows that S_u is a random variable.
We now prove (2.10). We start with considering simple functions, i.e. functions taking only finitely many different values and vanishing outside a set of finite measure. For such a function $u(x) = \sum_{i=1}^n u_i \mathbb{1}_{A_i}(x)$ with pairwise disjoint measurable sets A_i the random variables $\eta_i = \eta(A_i) \sim \mathrm{Poi}(\lambda_i)$, $\lambda_i = \lambda(A_i)$, are independent and

$$S_u = \sum_{x\in\eta} u(x) = \sum_{i=1}^n u_i\eta_i.$$

This gives us

$$\begin{aligned}\mathbb{E}\left[e^{\theta S_u}\right] &= \mathbb{E}\left[e^{\theta\sum_{i=1}^n u_i\eta_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{\theta u_i\eta_i}\right] = \prod_{i=1}^n \sum_{k=1}^\infty e^{\theta u_i k}\frac{\lambda_i^k}{k!}e^{-\lambda_i} = \prod_{i=1}^n e^{\lambda_i(e^{\theta u_i}-1)} \\ &= e^{\sum_{i=1}^n \lambda_i(e^{\theta u_i}-1)} = e^{\sum_{i=1}^n \int_{A_i}(e^{\theta u(x)}-1)\lambda(dx)} = \exp\left\{\int_{\mathbb{X}}(e^{\theta u(x)}-1)\lambda(dx)\right\}.\end{aligned}$$

For simple functions this equation holds for all $\theta \in \mathbb{C}$. Next, we take $u \geq 0$ and $\theta = -t$ for some $t \in (0,\infty)$. There exists an increasing sequence of simple functions u_j converging to u. Then $S_{u_j}(\omega) \xrightarrow{j\to\infty} S_u(\omega)$ for every realization $\omega \in \Omega$ by monotone convergence. Using dominated and monotone convergence we find

$$\begin{aligned}\mathbb{E}\left[e^{-tS_u}\right] &= \lim_{j\to\infty}\mathbb{E}\left[e^{-tS_{u_j}}\right] = \lim_{j\to\infty}\exp\left\{\int_{\mathbb{X}}(e^{-tu_j(x)}-1)\lambda(dx)\right\} \\ &= \lim_{j\to\infty}\exp\left\{-\int_{\mathbb{X}}(1-e^{-tu_j(x)})\lambda(dx)\right\} \\ &= \exp\left\{-\int_{\mathbb{X}}(1-e^{-tu(x)})\lambda(dx)\right\}.\end{aligned}$$

If (2.9) holds, we can compute the limit $t \to 0$ for the last integral using dominated convergence. From the mean value theorem we get

$$\left|e^{-tu(x)} - 1\right| \leq |tu(x)|$$

and for $u \geq 0$ it always holds

$$\left|1 - e^{-tu(x)}\right| \leq 1.$$

Thus, $\left|1 - e^{-tu(x)}\right| \leq |tu(x)| \wedge 1 \leq (t \vee 1)(|u(x)| \wedge 1)$. The last function is assumed to be integrable. Hence, we get by dominated convergence

$$\lim_{t\to 0}\mathbb{E}\left[e^{-tS_u}\right] = \lim_{t\to 0}\exp\left\{-\int_{\mathbb{X}}(1-e^{-tu(x)})\lambda(dx)\right\} = 1.$$

This shows that S_u is a finite random variable because if we had $S_u = \infty$ on some measurable set A with $\mathbb{P}(A) > 0$ we would get

$$1 = \lim_{t \to 0} \mathbb{E}\left[e^{-tS_u}\right] = \lim_{t \to 0} \mathbb{E}\left[e^{-tS_u} \mathbb{1}_{A^c}\right] \leq \mathbb{E}\left[\mathbb{1}_{A^c}\right] = 1 - \mathbb{P}(A) < 1.$$

This contradiction shows $S_u < +\infty$ almost surely.
The next step is to show that both sides of (2.10) are analytic functions on D and continuous on $\overline{D}$ if we take $u \geq 0$. From this we can conclude by the identity theorem for analytic functions that (2.10) holds for $u \geq 0$ because we have already shown equality on the negative real line. To show analyticity on D of the two functions

$$\theta \mapsto \mathbb{E}\left[e^{\theta S_u}\right] \quad \text{and} \quad \theta \mapsto \exp\left\{\int_{\mathbb{X}} (e^{\theta u(x)} - 1)\lambda(dx)\right\} \tag{2.13}$$

in (2.10), we use Morera's theorem. To apply this theorem, we need to show continuity of the two functions on D as well as that the integral over each closed and piecewise continuously differentiable curve in D vanishes. Note, that the domain of each function in (2.13) can be extended to $\overline{D}$ because the expectation and the integral exist on $\overline{D}$. This is established by the following considerations that show the continuity of the functions.
We start with the left side. Let $\theta_n \xrightarrow{n\to\infty} \theta$ in $\overline{D}$. Since $\mathrm{Re}(\theta_n) \leq 0$, we have $\left|e^{\theta_n S_u}\right| \leq 1$. Dominated converges then yields

$$\lim_{n\to\infty} \mathbb{E}\left[e^{\theta_n S_u}\right] = \mathbb{E}\left[e^{\theta S_u}\right]$$

which establishes continuity. Next, let γ be a closed and piecewise C^1 curve in D. Using Fubini's theorem and analyticity of the function $z \mapsto e^{zS_u}$ almost everywhere (Cauchy's integral theorem states that $\oint_\gamma f(z)dt = 0$ for γ closed and piecewise C^1 and f holomorphic and from $S_u < +\infty$ we get that $z \mapsto e^{zS_u}$ is analytic a.e.) we have

$$\oint_\gamma \mathbb{E}\left[e^{zS_u}\right] dz = \int_0^1 \mathbb{E}\left[e^{\gamma(t)S_u}\right] \gamma'(t)dt = \mathbb{E}\left[\int_0^1 e^{\gamma(t)S_u}\gamma'(t)dt\right] = \mathbb{E}\left[\oint_\gamma e^{zS_u}dz\right] = 0.$$

The use of Fubini is justified by

$$\mathbb{E}\left[\int_0^1 \left|e^{\gamma(t)S_u}\gamma'(t)\right| dt\right] \leq \|\gamma'\|_\infty \mathbb{E}\left[\int_0^1 e^{\gamma(t)S_u}dt\right] \lesssim 1.$$

A similar approach shows that the right side of (2.10) is holomorphic on D as well. Let again $\theta_n^1 + i\theta_n^2 = \theta_n \to \theta$ in $\overline{D}$. We can estimate

$$\begin{aligned}\left|e^{\theta_n u(x)} - 1\right| &= \left|e^{\theta_n^1 u(x)} e^{i\theta_n^2 u(x)} - 1\right| \leq \left|e^{\theta_n^1 u(x)}(e^{i\theta_n^2 u(x)} - 1)\right| + \left|e^{\theta_n^1 u(x)} - 1\right| \\ &\leq 2 \wedge |\theta_n^2| u(x) + 1 \wedge \theta_n^1 u(x) \leq (2 \vee |\theta_n^2| + 1 \vee \theta_n^1)(1 \wedge u(x)) \lesssim 1 \wedge u(x),\end{aligned} \tag{2.14}$$

where we used $\left|e^{it} - 1\right| \leq 2 \wedge |t|$ ([13, Lemma 4.14]) and $\left|e^{-tu(x)} - 1\right| \leq 1 \wedge |tu(x)|$ (which we have derived above). Then, dominated convergence yields continuity

$$\exp\left\{\int_{\mathbb{X}} (e^{\theta_n u(x)} - 1)\lambda(dx)\right\} \xrightarrow{n\to\infty} \exp\left\{\int_{\mathbb{X}} (e^{\theta u(x)} - 1)\lambda(dx)\right\}.$$

To show that the function is holomorphic, it suffices to show that the function $z \mapsto \int_{\mathbb{X}}(e^{zu(x)} - 1)\lambda(dx)$ is holomorphic on D because the composition of holomorphic functions again is holomorphic. Therefore, let γ be a closed piecewise C^1 curve in D. Using Cauchy's integral theorem applied to the holomorphic functions $z \mapsto (e^{zu(x)} - 1)$ we have

$$\begin{aligned}\oint_\gamma \int_{\mathbb{X}} (e^{zu(x)} - 1)\lambda(dx)dz &= \int_0^1 \int_{\mathbb{X}} (e^{\gamma(t)u(x)} - 1)\lambda(dx)\gamma'(t)dt \\ &= \int_{\mathbb{X}} \int_0^1 (e^{\gamma(t)u(x)} - 1)\gamma'(t)dt\lambda(dx) \\ &= \int_{\mathbb{X}} \oint_\gamma \left(e^{zu(x)} - 1\right) dz\lambda(dx) = 0,\end{aligned}$$

where the change of the order of integration is justified by Fubini's theorem since we have with (2.14)

$$\begin{aligned}\int_{\mathbb{X}} \int_0^1 \left|(e^{\gamma(t)u(x)} - 1)\gamma'(t)\right| dt\lambda(dx) &\le \int_{\mathbb{X}} \int_0^1 (2 \vee |\gamma^2|(t) + 1 \vee \gamma^1(t))(1 \wedge u(x)) \left|\gamma'(t)\right| dt\lambda(dx) \\ &\le \left\|\gamma' \left(2 \vee |\gamma^2| + 1 \vee \gamma^1\right)\right\|_\infty \int_{\mathbb{X}} 1 \wedge u(x)\lambda(dx) \lesssim 1.\end{aligned}$$

Next, we show that (2.9) is necessary for S_u to converge. If (2.9) does not hold, we have $S_u = +\infty$ with probability one. To see this, we use the mean value theorem to deduce for $0 \le u(x) < 1$

$$\frac{e^{-tu(x)} - 1}{t} = -u(x)e^{-\tilde{t}u(x)} \quad \Rightarrow \quad e^{-tu(x)} - 1 \le -tu(x)e^{-t} \quad \Leftrightarrow \quad 1 - e^{-tu(x)} \ge tu(x)e^{-t}$$

where $0 < \tilde{t} < 1$. Then we find

$$\begin{aligned}\int_{\mathbb{X}} (1 - e^{-tu(x)})\lambda(dx) &= \int_{\mathbb{X}} \mathbb{1}_{1 \le u}(x)(1 - e^{-tu(x)})\lambda(dx) + \int_{\mathbb{X}} \mathbb{1}_{1 > u}(x)(1 - e^{-tu(x)})\lambda(dx) \\ &\ge \int_{\mathbb{X}} \mathbb{1}_{1 \le u}(x)(1 - e^{-t})\lambda(dx) + te^{-t} \int_{\mathbb{X}} \mathbb{1}_{1 > u}(x)u(x)\lambda(dx) \\ &\ge (1 - e^{-t}) \int_{\mathbb{X}} \mathbb{1}_{1 \le u}(x)\lambda(dx) + te^{-t} \int_{\mathbb{X}} \mathbb{1}_{1 > u}(x)u(x)\lambda(dx) \\ &\ge (1 - e^{-t} \wedge te^{-t}) \int_{\mathbb{X}} (1 \wedge u(x))\lambda(dx) = +\infty,\end{aligned}$$

showing that $\mathbb{E}\left[e^{-tS_u}\right] = 0$, hence $S_u = +\infty$ almost surely. Up to now the theorem is proved for $u \ge 0$.

For an arbitrary measurable function u we consider

$$u^+ = u \vee 0 \quad \text{and} \quad u^- = -u \vee 0$$

and the restrictions η^+ and η^- of η to the sets $\{u > 0\}$ and $\{u < 0\}$ respectively. Moreover, let

$$S_u^+ = \sum_{x \in \eta} u^+(x) = \sum_{x \in \eta^+} u(x) \quad \text{and} \quad S_u^- = \sum_{x \in \eta} u^-(x) = \sum_{x \in \eta^-} u(x).$$

The Poisson processes η^+ and η^- are independent since $\{u > 0\}$ and $\{u < 0\}$ are disjoint and from the definition of the σ-algebra $\mathcal{N}$ we get that the function

$$f\colon (\mathbf{N}, \mathcal{N}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R})), \quad \xi \mapsto \int u d\xi$$

is measurable for $u \geq 0$. Thus, $S_u^+ = f \circ \eta^+$ and $S_u^- = f \circ \eta^-$ are independent. We now apply the results above. This yields convergence of the sums S_u^+ and S_u^- if and only if the integrals

$$\int_{\mathbb{X}} u^+(x) \wedge 1 \lambda(dx) \quad \text{and} \quad \int_{\mathbb{X}} u^-(x) \wedge 1 \lambda(dx)$$

and hence

$$\int_{\mathbb{X}} u^+(x) \wedge 1 \lambda(dx) + \int_{\mathbb{X}} u^-(x) \wedge 1 = \int_{\mathbb{X}} (|u(x)| \wedge 1) \lambda(dx)$$

are finite. Since the sums S_u^+ and S_u^- converge if and only if

$$S_u = \sum_{x \in \eta} u(x)$$

converges absolutely, (2.9) is a necessary and sufficient condition for S_u to converge absolutely. Finally, if (2.9) holds and $\theta \in i\mathbb{R}$, using the independence of S_u^+ and S_u^-, we get

$$\begin{aligned}
\mathbb{E}\left[e^{\theta S_u}\right] &= \mathbb{E}\left[e^{\theta S_u^+ - \theta S_u^-}\right] = \mathbb{E}\left[e^{\theta S_u^+}\right] \mathbb{E}\left[e^{-\theta S_u^-}\right] \\
&= \exp\left\{\int_{\mathbb{X}} (e^{\theta u^+(x)} - 1) \lambda(dx)\right\} \exp\left\{\int_{\mathbb{X}} (e^{-\theta u^-(x)} - 1) \lambda(dx)\right\} \\
&= \exp\left\{\int_{\mathbb{X}} \left(e^{\theta u^+(x)} + e^{-\theta u^-(x)} - 2\right) \lambda(dx)\right\} \\
&= \exp\left\{\int_{\mathbb{X}} (e^{\theta u(x)} - 1) \lambda(dx)\right\}.
\end{aligned}$$

Next, we prove the formulas for the expectation and the variance. Therefore, we take a simple function u and purely imaginary θ in (2.10) and consider the formula

$$\varphi_{S_u}(t) = \mathbb{E}\left[e^{itS_u}\right] = \exp\left\{\int_{\mathbb{X}} (e^{itu(x)} - 1) \lambda(dx)\right\}.$$

It is

$$\mathbb{E}[S_u] = \frac{1}{i} \frac{d}{dt} \varphi_{S_u}(0) \quad \text{and} \quad \mathbb{E}\left[S_u^2\right] = \frac{1}{i^2} \frac{d^2}{dt^2} \varphi_{S_u}(0)$$

if the second derivative of the characteristic function $\varphi_{S_u}(t)$ exists [14, Theorem 15.34 and Theorem 15.31]. We get

$$\begin{aligned}
\frac{d}{dt} \mathbb{E}\left[e^{itS_u}\right]\Bigg|_{t=0} &= \exp\left\{\int_{\mathbb{X}} (e^{itu(x)} - 1) \lambda(dx)\right\} \frac{d}{dt} \int_{\mathbb{X}} (e^{itu(x)} - 1) \lambda(dx)\Bigg|_{t=0} \\
&= i \int_{\mathbb{X}} u(x) \lambda(dx),
\end{aligned}$$

where the change of integration and differentiation is justified by $x \mapsto (e^{itu(x)} - 1)$ being dominated in a neighborhood around $t = 0$ by an integrable function due to $u(x)$ taking nonzero, finite values only in a set of finite measure. The same is true for the derivative $\frac{d}{dt}(e^{itu(x)} - 1) = u(x)e^{itu(x)}$. Continuing in a similar way we compute

$$\begin{aligned}\frac{d^2}{dt^2}\mathbb{E}\left[e^{itS_u}\right]\bigg|_{t=0} &= \frac{d}{dt}\left[\exp\left\{\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\right\}\frac{d}{dt}\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\right]\bigg|_{t=0} \\ &= \exp\left\{\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\right\}\left(\frac{d}{dt}\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\right)^2\bigg|_{t=0} \\ &\quad + \exp\left\{\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\right\}\frac{d^2}{dt^2}\int_{\mathbb{X}}(e^{itu(x)} - 1)\lambda(dx)\bigg|_{t=0} \\ &= \left(i\int_{\mathbb{X}}u(x)\lambda(dx)\right)^2 + i^2\int_{\mathbb{X}}u^2(x)\lambda(dx),\end{aligned}$$

using the same arguments as above for the change of integration and differentiation. This gives us

$$\mathbb{V}\left[\int u(x)\eta(dx)\right] = \mathbb{E}\left[\left(\int u(x)\eta(dx)\right)^2\right] - \mathbb{E}\left[\int u(x)\eta(dx)\right]^2 = \int_{\mathbb{X}} u^2(x)\lambda(dx).$$

Thus, we justified the formulas (2.11) and (2.12) in the case of u being a simple function. Next, for $u \geq 0$ measurable there exists a sequence of simple functions u_n converging monotonically to u. Monotone convergence yields the formulas for expectation and variance for nonnegative functions as long as $\int_{\mathbb{X}} u(x)\lambda(dx) < +\infty$. For arbitrary measurable functions u we split u into positive and negative part $u = u_+ - u_-$ both for which the two formulas hold as u_+ and u_- are non-negative. Thus, $\mathbb{E}[S_u]$ exists if and only if the integral $\int_{\mathbb{X}} u(x)\lambda(dx)$ converges, i.e. both integrals $\int_{\mathbb{X}} u_+(x)\lambda(dx)$ and $\int_{\mathbb{X}} u_-(x)\lambda(dx)$ are finite. This establishes (2.11). In this case the variance is well-defined and formula (2.12) follows with similar arguments.

□

In the theorem above we only have expressions for the expectation and the variance of the random variable S_u. Later we can improve our results by using higher order central moments of S_u instead of the variance. The expressions for those moments are more complex and we will give a formula in the following lemma.

Lemma 2.2.5. *Let η be a Poisson point process on $(\mathbb{X}, \mathcal{X})$ with σ-finite intensity measure λ, and let $n \in \mathbb{N}$ and $u\colon \mathbb{X} \to \mathbb{R}$ be measurable such that*

$$\int_{\mathbb{X}} |u|^j \, d\lambda < +\infty$$

for $j = 1, 2, \dots, N$ where $N = n$ if n is even and $N = n + 1$ if n is odd. We set $S_u = \int u(x)\eta(dx) = \sum_{x\in\eta} u(x)$. Then we have

$$\mathbb{E}\left[(S_u - \mathbb{E}[S_u])^n\right] = \sum_{k\cdot s = n, k_1 = 0} \frac{n!}{k!(s!)^k} a^k$$

where the sum runs over all $k \in \mathbb{N}^\infty$ satisfying $\sum_{l=1}^\infty l k_l = n$ and $k_1 = 0$. We used the abbreviations $a_l = \int_{\mathbb{X}} u^l(x)\lambda(dx)$, $a^k = \prod_{l=1}^\infty a_l^{k_l}$, $s = (1, 2, 3, \dots) \in \mathbb{N}^\infty$, $k! = \prod_{l=1}^\infty k_l!$ and $(s!)^k = \prod_{l=1}^\infty (l!)^{k_l}$.
Especially, the above formula implies that $\mathbb{E}\left[(S_u - \mathbb{E}[S_u])^n\right]$ is a polynomial of degree at most $\lfloor \frac{n}{2} \rfloor$ in the variables $a_1, a_2, \dots, a_n$.

Proof. To find an expression for $\mathbb{E}\left[(S_u - \mathbb{E}[S_u])^n\right]$, we need a formula for the moments of S_u. We are going to derive these moment formulas from the characteristic function of S_u that is given by

$$\varphi_{S_u}(t) = \mathbb{E}\left[e^{itS_u}\right] = \exp\left\{\int_{\mathbb{X}} (e^{itu(x)} - 1)\lambda(dx)\right\}$$

according to Theorem 2.2.4. The relation between the moments of a random variable S_u and its characteristic function φ_{S_u} is as follows [14, section 15.4]: If $\mathbb{E}\left[|S_u|^k\right] < +\infty$, then φ_{S_u} is k times continuously differentiable and it holds

$$\mathbb{E}\left[S_u^j\right] = \frac{1}{i^j}\varphi_{S_u}^{(j)}(0) \qquad \text{for all } j = 0, 1, \dots, k.$$

Moreover, if φ_{S_u} is k times differentiable for some $k \in \mathbb{N}$ even, then

$$\mathbb{E}\left[S_u^k\right] = (-1)^{k/2}\varphi_{S_u}^{(k)}(0) < +\infty.$$

To find the derivatives of the characteristic function, we use a general formula for finding the derivatives of $e^f(x)$ [15]. It is

$$\left(e^{f(x)}\right)^{(n)} = e^{f(x)} \sum_{\{k | k \in \mathbb{N}^\infty, k \cdot s = n\}} b_k^n \left(f^{(s)}(x)\right)^k$$

with $s = (1, 2, 3, \dots)$ and coefficients

$$b_k^n = \frac{1}{\prod_{i=1}^n k_i!} \frac{n!}{\prod_{i=1}^n (s_i!)^{k_i}} = \frac{n!}{k!(s!)^k}.$$

We apply this to the function $f(t) = \int_{\mathbb{X}} (e^{itu(x)} - 1)\lambda(dx)$. For $j = 1, 2, \dots, N$, the j-th derivative of f is given by

$$f^{(j)}(t) = \int_{\mathbb{X}} i^j u^j(x) e^{itu(x)} \lambda(dx)$$

due to the assumed existence of the integrals $a_j = \int_{\mathbb{X}} u^j d\lambda$ and due to

$$\int_{\mathbb{X}} \left|e^{itu} - 1\right| d\lambda \le \int_{\mathbb{X}} 2 \wedge |tu| \, d\lambda \le |t| \int_{\mathbb{X}} |u| \, d\lambda < +\infty.$$

Hence, we find

$$\mathbb{E}\left[S_u^j\right] = \frac{1}{i^j}\varphi_{S_u}^{(j)}(0) = \sum_{\{k | k \in \mathbb{N}^\infty, k \cdot s = j\}} b_k^j \left(\prod_{i=1}^\infty (a_i)^{k_i}\right).$$

Thus, we arrive at

$$\begin{aligned}\mathbb{E}\left[(S_u - \mathbb{E}\left[S_u\right])^n\right] &= \sum_{l=0}^{n} \binom{n}{l} (-1)^{n-l} \mathbb{E}\left[S_u^l\right] \mathbb{E}\left[S_u\right]^{n-l} \\ &= \sum_{l=0}^{n} \sum_{k \cdot s = l} (-1)^{n-l} \frac{l!}{k!(s!)^k} \binom{n}{l} a^k a_1^{n-l}.\end{aligned}$$

We now want to show that the coefficients of all terms $a^k a_1^{n-l}$ with $k_1 + n - l > 0$ vanish in order to show the desired formula

$$\mathbb{E}\left[(S_u - \mathbb{E}\left[S_u\right])^n\right] = \sum_{k \cdot s = n, k_1 = 0} \frac{n!}{k!(s!)^k} a^k,$$

because the only terms satisfying $k_1 + n - l = 0$ are those with $k_1 = 0$ and $l = n$ which leads to the above formula.

We pick any $\tilde{k} \in \mathbb{N}^\infty$ and $0 \leq \tilde{l} \leq n$ such that $\tilde{k} \cdot s = \tilde{l}$ and $\tilde{k}_1 + n - \tilde{l} > 0$. The other pairs (k, l) contributing to the coefficient of $a^{\tilde{k}} a_1^{n-\tilde{l}}$ can be constructed from the following considerations. In order to result in the same term $a^{\tilde{k}} a_1^{n-\tilde{l}}$ it must hold

$$\tilde{k}_i = k_i \quad \text{for} \quad i \geq 2 \quad \text{and} \quad \tilde{k}_1 - \tilde{l} + n = k_1 - l + n \quad \Longleftrightarrow \quad l = k_1 - \tilde{k}_1 + \tilde{l}.$$

Thus, the smallest possible l (denoted by l_0) is attained for $k_1 = 0$ and from $\tilde{k} \cdot s = \tilde{l}$ we get

$$l_0 = \tilde{l} - \tilde{k}_1 \geq 0.$$

Moreover, it is $l_0 < n$ because if we had $l_0 = n$, we would end up with $n = l_0 = \tilde{l} - \tilde{k}_1 < n$, using the definition of the pair $(\tilde{k}, \tilde{l})$. The possible values for k_1 can now be constructed from l_0 via $k_1 = \tilde{k}_1 - \tilde{l} + l = l - l_0$, leading to

$$k^l = (l - l_0, \tilde{k}_2, \tilde{k}_3, \dots) \quad \text{for} \quad l_0 \leq l \leq n.$$

Therefore, the coefficients are obtained by

$$\begin{aligned}\operatorname{Coeff}(a^{\tilde{k}} a_1^{n-\tilde{l}}) &= \sum_{l=l_0}^{n} (-1)^{n-l} \binom{n}{l} \frac{l!}{k^l!(s!)^{k^l}} = \sum_{l=l_0}^{n} (-1)^{n-l} \binom{n}{l} \frac{l!}{(l-l_0)!l_0!} \frac{l_0! \tilde{k}_1!}{(s!)^{\tilde{k}} \tilde{k}!} \\ &= (-1)^n \frac{l_0! \tilde{k}_1!}{(s!)^{\tilde{k}} \tilde{k}!} \sum_{l=l_0}^{n} (-1)^l \binom{n}{l} \binom{l}{l_0}.\end{aligned}$$

We proceed by showing that the above sum vanishes for $0 \leq l_0 < n$ using induction. With the binomial theorem we get in the case $l_0 = 0$ for any n

$$\sum_{l=0}^{n} (-1)^l \binom{n}{l} \binom{l}{0} = \sum_{l=0}^{n} (-1)^l \binom{n}{l} = (-1+1)^n = 0.$$

This is the induction start for $n = 1$ as well. Now, suppose we have shown $\sum_{l=l_0}^{n} (-1)^l \binom{n}{l} \binom{l}{l_0} = 0$ for some $n \in \mathbb{N}$ and all $0 \leq l_0 < n$. The statement then follows for

$n+1$ and $1 \leq l_0 < n+1$ by observing

$$\begin{aligned}\sum_{l=l_0}^{n+1}(-1)^l\binom{n+1}{l}\binom{l}{l_0} &= -\sum_{l=l_0-1}^{n}(-1)^l\binom{n+1}{l+1}\binom{l+1}{l_0}\\ &= -\sum_{l=l_0-1}^{n}(-1)^l\frac{n+1}{l+1}\binom{n}{l}\frac{l+1}{l_0}\binom{l}{l_0-1}\\ &= -\frac{n+1}{l_0}\sum_{l=l_0-1}^{n}(-1)^l\binom{n}{l}\binom{l}{l_0-1} = 0\end{aligned}$$

as the last sum vanishes because of the induction hypothesis applied to n and $l_0 - 1 \geq 0$. The case $l_0 = 0$ was already done above.
Finally, the conditions $k \cdot s = n$ and $k_1 = 0$ imply that $|k| \leq \lfloor \frac{n}{2} \rfloor$ showing that

$$\mathbb{E}\left[(S_u - \mathbb{E}\left[S_u\right])^n\right] = \sum_{k\cdot s=n,k_1=0}\frac{n!}{k!(s!)^k}a^k$$

is a polynomial of degree at most $\lfloor \frac{n}{2} \rfloor$ in $a_1, a_2, \ldots, a_n$. □

2.3 Model Derivation

In this section we describe the mathematical model of the PET measurement and deduce our reconstruction formula for the inverse problem of determining the distribution of radioactive material from measured PET data. In essence, this data is obtained as follows. The radioactive material randomly emits a positron. Shortly after the emission the positron annihilates with an electron and two photons are emitted at 180 degrees to each other. These photons are then detected. Normally, the number of detected events is low. This means that reconstructing the underlying radioactive material distribution from measured PET data is an underdetermined and hence ill-posed problem which needs to be regularized by incorporating prior knowledge. Within our approach this will be done using optimal transport. The considerations in this section are based on [1].

2.3.1 Measurements and Material Distribution

The interior of the PET scanner is modeled by a compact and convex set $\Omega \subset \mathbb{R}^3$ and the measurement takes place during a time interval $[0, T]$ for a time horizon $T > 0$. The distribution ρ of the radioactive material is described by a nonnegative Radon measure on $[0, T] \times \Omega$

$$\rho \in \mathcal{M}_+([0, T] \times \Omega).$$

As we are using optimal transport regularization, the solutions to the optimization problem necessarily satisfy the continuity equation which results in a disintegration in time (see Section 2.1), i.e. we have

$$\rho = dt \otimes \rho_t$$

where dt is the Lebesgue measure on $[0, T]$.

The radioactive decay happens according to a Poisson point process with intensity

$$dt \otimes \lambda_t := dt \otimes \frac{\ln 2}{T_{1/2}} \rho_t$$

with $T_{1/2}$ being the radionuclid's half-life.
To describe the detection process, let $D = \{1, \dots, M\}$ be the set of photon detectors. Detector i covers a Borel measurable region $\Gamma_i \subset \partial\Omega$. The detectors are nonoverlapping, i.e. we have $\Gamma_i \cap \Gamma_j = \emptyset$ for $i \neq j$ and we assume each detector to have a piecewise C^1-boundary. Moreover, for each measurement we fix a temporal resolution Δt, leading to a partition of the time interval into N disjoint intervals $\tau_i = [(i-1)\Delta t, i\Delta t]$ for $i = 1, \dots, N = \frac{T}{\Delta t}$. Since we are using PET as the measurement method, each detected event comes with a photon detection in two different detectors. The actual measurement is then a map

$$E : \{1, \dots, N\} \times D \times D \to \mathbb{N}$$

where, for $k \leq l$, $E_{ikl} := E(i, k, l)$ is the number of photon pairs detected by detector pair (k, l) within the time period τ_i. We set $E_{ikl} = 0$ for $k > l$. Moreover, we choose Δt small enough to guarantee $E_{ikl} \in \{0, 1\}$. The measurement will be described by an operator

that is derived from physical considerations. This operator transforms the measure ρ_t to another measure κ_t that describes the distribution of photons at the detectors. The actual measurement is then generated by a Poisson point process on $[0,T] \times \partial\Omega \times \partial\Omega$ (the measurement process is defined on $\partial\Omega \times \partial\Omega$ due to our definition of the forward operator, see section 2.3.2) with intensity measure $dt \otimes \frac{\ln 2}{T_{1/2}} \kappa_t$, i.e. for $k \leq l$ the number of photons E_{ikl} is Poisson distributed with mean $\frac{\ln 2}{T_{1/2}} \int_{\tau_i} \kappa_t(\Gamma_{kl}) dt$ where Γ_{kl} describes the area covered by the detector pair kl. The sets Γ_{kl} do not necessarily cover the whole surface of the measurement volume. We set $\bigcup_{kl} \Gamma_{kl} =: \Gamma_{\texttt{tot}} \subset \partial\Omega \times \partial\Omega$.
Since we later want to analyze the behavior in the limit $T_{1/2} \to 0$ we explicitly emphasize the dependence of the measurement process on the half-life $T_{1/2}$. The density $\kappa_t \in \mathcal{M}_+(\partial\Omega \times \partial\Omega)$ can be derived from the material distribution by applying the forward operator introduced in the next section.

2.3.2 Forward operator

The forward operator transforms the intensity measure ρ_t of the PPP describing the decay of the radionuclide into the intensity measure κ_t of the measurement PPP. The detection process decomposes into three parts, attenuation, scattering and detection without substantial scattering. Therefore, we split the λ_t and ρ_t into

$$(\lambda_t^a, \lambda_t^s, \lambda_t^d) = \frac{\ln 2}{T_{1/2}} (\rho_t^a, \rho_t^s, \rho_t^d).$$

The probability of attenuation or scattering depends on the material that the photon passes. This gives us

$$\lambda_t^a = p_t^a \lambda_t, \; \lambda_t^s = p_t^s \lambda_t, \; \lambda_t^d = p_t^d \lambda_t$$

for some functions $p_t^a, p_t^s, p_t^d \colon \Omega \to [0,1], \; t \in [0,T]$, with $p_t^a + p_t^s + p_t^d = 1$. These functions depend on the material composition at time t. For simplicity, we assume these functions to be spatiotemporally constant, and we will write p^a, p^s and p^d in the following. We now model the three parts of the forward operator.

(1) *Attenuation*
The forward operator describing the attenuation simply sets all intensity to zero,

$$\begin{aligned} A^a \colon \mathcal{M}_+(\Omega) &\to \mathcal{M}_+(\partial\Omega \times \partial\Omega) \\ \lambda &\mapsto 0. \end{aligned}$$

(2) *Scattering*
Scattering changes the direction of the photon rays randomly in our model, meaning that the probability of a scattered photon pair to arrive at point $(x,y) \in \partial\Omega \times \partial\Omega$ is homogeneous. This leads to the operator

$$\begin{aligned} A^s \colon \mathcal{M}_+(\Omega) &\to \mathcal{M}_+(\partial\Omega \times \partial\Omega) \\ \lambda &\mapsto \frac{\lambda(\Omega)}{\mathcal{H}^2(\partial\Omega)^2} \mathcal{H}^2 \otimes \mathcal{H}^2 \llcorner (\partial\Omega \times \partial\Omega) \end{aligned}$$

where $\mathcal{H}^2$ is the two-dimensional Hausdorff measure and $\llcorner$ denotes the restriction of a measure to a set.

(3) *Detection*
The operator describing the detection process without scattering is composed of three linear operators

$$A^d = B_3 B_2 B_1$$

modeling the positron range, the direction of the scattered photon pairs and the detection.
Let $G_y : \Omega \to [0, +\infty)$ be a smooth convolution kernel describing the probability density of an annihilation of a positron emitted at y. The operator

$$\begin{aligned} B_1 \colon \mathcal{M}_+(\Omega) &\to \mathcal{M}_+(\Omega) \\ \lambda &\mapsto \int_\Omega G_y \lambda(dy) \end{aligned}$$

then models the positron range from emission to annihilation. In general, G_y depends on the surrounding material but we will only consider a spatially homogeneous Gaussian kernel

$$\frac{1}{\sqrt{8\pi^3}\varepsilon} \exp\left(-\frac{|x-y|^2}{2\varepsilon}\right)$$

for some fixed $\varepsilon > 0$.
When the positron annihilates, two photons are emitted in opposite directions. This is modeled by an operator transforming the intensity of photon emissions into a density on $\Omega \times G^{1,3}$ (a position with a direction), where $G^{1,3}$ denotes the Grassmannian manifold of one-dimensional subspaces in $\mathbb{R}^3$. On $G^{1,3}$ we have a natural probability measure $\mathrm{vol}_{G^{1,3}}$ [16, Chapter 3]. Since the direction of a photon pair after annihilation is distributed uniformly, the operator reads

$$\begin{aligned} B_2 \colon \mathcal{M}_+(\Omega) &\to \mathcal{M}_+(\Omega \times G^{1,3}) \\ \lambda &\mapsto \lambda \otimes \mathrm{vol}_{G^{1,3}}. \end{aligned}$$

Finally, the photon pair emitted at $x \in \Omega$ in direction $v \in G^{1,3}$ will be detected at positions $R(x, v)$ with

$$\begin{aligned} R \colon \Omega \times G^{1,3} &\to \partial\Omega \times \partial\Omega \\ (x, v) &\mapsto \partial\Omega \cap (x + v) \end{aligned}$$

where two-element subsets of $\partial\Omega$ are identified with a point in $\partial\Omega \times \partial\Omega$ with lexicographic ordering. With the function R we can write the operator B_3 describing the actual detection of a photon pair as

$$\begin{aligned} B_3 \colon \mathcal{M}_+(\Omega \times G^{1,3}) &\to \mathcal{M}_+(\partial\Omega \times \partial\Omega) \\ \lambda &\mapsto R_\# \lambda \end{aligned}$$

where $R_\#\lambda$ is the pushforward measure of λ. In total, the intensity $dt \otimes \frac{\ln 2}{T_{1/2}} \kappa_t$ inducing the measurement PPP is given by

$$\kappa_t = \underbrace{\kappa_t^a}_{=0} + \kappa_t^s + \kappa_t^d = \underbrace{p^a A^a \rho_t}_{=0} + p^s A^s \rho_t + p^d A^d \rho_t.$$

Remark 2.3.1 (Evaluation of the detection operator)**.** Let $\Gamma \subset \partial\Omega \times \partial\Omega$ and for $v \in G^{1,3}$ define

$$Z_v^\Gamma := \{x \in \Omega \mid (x, v) \in R^{-1}(\Gamma)\}.$$

According to Cavalieri's principle [17, Theorem 9.6.6] Z_v^Γ is measurable for every $v \in G^{1,3}$ and we have for $\rho_t \in \mathcal{M}_+(\Omega)$

$$\begin{aligned} A^d\rho_t(\Gamma) = B_3B_2B_1\rho_t(\Gamma) = (B_2B_1\rho_t)(R^{-1}(\Gamma)) &= \int_{G^{1,3}} (B_1\rho_t)(Z_v^\Gamma)\text{vol}_{G^{1,3}}(dv) \\ &= \int_{G^{1,3}} \int_{Z_v^\Gamma} (B_1\rho_t)(x)dx\text{vol}_{G^{1,3}}(dv) \\ &= \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega G_y(x)d\rho_t(y)dx\text{vol}_{G^{1,3}}(dv). \end{aligned}$$

Remark 2.3.2 (Absolute continuity w.r.t. $\mathcal{H}^2 \otimes \mathcal{H}^2$)**.** The measure κ_t is absolutely continuous with respect to the measure $\mathcal{H}^2 \otimes \mathcal{H}^2$ for every $t \in [0, T]$. For the scattering part, this follows immediately from the definition of the respective forward operator as $\mathcal{H}^2 \otimes \mathcal{H}^2$ restricted to $\partial\Omega \times \partial\Omega$ appears.
The absolute continuity of the detection part comes from the smoothing by convolution with the C^∞ function G. To see this, let $\Gamma \subset \partial\Omega \times \partial\Omega$ such that $\mathcal{H}^2 \otimes \mathcal{H}^2\,(\Gamma) = 0$. With Remark 2.3.1 we get

$$\begin{aligned} A^d\rho_t(\Gamma) &= \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega G_y(x)d\rho_t(y)dx\text{vol}_{G^{1,3}}(dv) \le \|G\|_\infty \rho_t(\Omega) \int_{G^{1,3}} \int_{Z_v^\Gamma} dx\text{vol}_{G^{1,3}}(dv) \\ &\lesssim \text{diam}(\Omega)\mathcal{H}^2 \otimes \mathcal{H}^2\,(\Gamma)\,\text{vol}_{G^{1,3}}(G^{1,3}) \lesssim \mathcal{H}^2 \otimes \mathcal{H}^2\,(\Gamma) = 0. \end{aligned}$$

2.3.3 Optimal transport regularization

To deal with the ill-posedness of our inverse problem we will apply a Bayesian approach and use a maximum a posteriori estimate. Within this framework we need to assign each possible material distribution a likelihood. This likelihood is obtained using optimal transport, thus we are going to use optimal transport regularization.
We use this framework to determine the kinetic motion energy at each time for all temporally evolving distributions and are going to assign a higher likelihood to a distribution the less kinetic energy is associated with the temporal evolution of it. To quantify the mass movement we introduce a vector valued Radon measure

$$\omega \in \mathcal{M}([0, T] \times \Omega)^3$$

describing the time dependent material flux leading to the temporal variation of ρ. Both measures ρ and ω need to be compatible which is expressed by the continuity equation

$$\partial_t\rho + \nabla_x \cdot \omega = 0 \tag{2.15}$$

that describes mass conservation. The equation is to be understood in the distributional sense, i.e. it holds

$$\int_{[0,T]\times\Omega} \partial_t\varphi d\rho + \int_{[0,T]\times\Omega} \langle\nabla_x\varphi, d\omega\rangle = 0 \quad \text{for all } \varphi \in C_c^1((0, T) \times \Omega).$$

Within our regularization approach ρ and ω can be disintegrated in time (see section 2.1 for details) and we can also write

$$\int_0^T \int_\Omega \partial_t \varphi d\rho_t dt + \int_0^T \int_\Omega \langle \nabla_x \varphi, d\omega_t \rangle dt = 0 \quad \text{for all } \varphi \in C_c^1((0,T) \times \Omega).$$

Using the continuity equation as above we assumed $T \ll T_{1/2}$ meaning that the amount of radioactive material approximately remains constant during the measurement time. Without these assumptions we could use unbalanced optimal transport to account for mass loss by introducing a decay term in the continuity equation.
For our Bayesian approach we need to assign a likelihood to each path (ρ, ω). This is done using a slightly modified version of the Benamou-Brenier functional $\mathcal{B}_2$ and that was already introduced in Section 2.1. We choose

$$\mathcal{S}(\rho, \omega) = \begin{cases} \int_0^T \int_\Omega \left(\frac{d\omega_t}{d\rho_t}\right)^2 d\rho_t dt & \text{if } \rho \geq 0,\ \omega \ll \rho \text{ and (2.15) holds,} \\ +\infty & \text{else,} \end{cases}$$

where $\frac{d\mu}{d\nu}$ denotes the Radon-Nikodym derivative of the measure μ with respect to ν. As already mentioned at the end of Section 2.1, we have $\mathcal{S}(\rho, \omega) = \mathcal{B}_2(\rho, \omega)$ as soon as $\mathcal{S}(\rho, \omega) < +\infty$. To finally quantify the likelihood of (ρ, ω) we assume a Boltzmann-type probability distribution meaning that we have

$$\mathbb{P}(\rho, \omega) \approx \exp(-\beta \mathcal{S}(\rho, \omega))$$

where β can be seen as a an inverse temperature and will be the regularization parameter. Physically, $\mathcal{S}(\rho, \omega)$ is the action of the path (ρ, ω) and $\int_\Omega \left(\frac{d\omega_t}{d\rho_t}\right)^2 d\rho_t$ can be seen as the kinetic energy of all particles in the system at a given time t. Hence, we use a kinetic regularization. As the minimum of the Benamou-Brenier functional for a transport from ρ_0 to ρ_T is proportional to the squared Wasserstein-2 distance between both measures, we assign a higher likelihood the less mass moves in total. This is another interpretation of the regularization which only takes into account the total mass movement but ignores the velocities that are associated to this mass movement. Note, that we make no assumptions about the spatial distribution of ρ.

2.3.4 Reconstruction method

In this paragraph we deduce the function to be minimized in order to obtain the reconstruction (ρ, ω) for a given measurement. We use a Bayesian approach to deal with the ill-posedness of the inverse reconstruction problem. Using Bayes' formula, the conditional probability of a pair (ρ, ω) given a measurement E reads

$$\mathbb{P}(\rho, \omega | E) = \frac{\mathbb{P}(E | \rho, \omega) \mathbb{P}(\rho, \omega)}{\mathbb{P}(E)}.$$

Let $\mathcal{P}_\lambda(k) = \lambda^k e^{-\lambda}/k!$ be the Poisson distribution with parameter λ and let

$$K_{ikl} = \frac{\ln 2}{T_{1/2}} \int_{\tau_i} \kappa_t(\underbrace{\Gamma_k \times \Gamma_l \cup \Gamma_l \times \Gamma_k}_{=:\Gamma_{kl}}) dt$$

be the expected number of photons in detector pair kl during the time interval τ_i. Note that we need both sets $\Gamma_k \times \Gamma_l$ and $\Gamma_l \times \Gamma_k$ to take into account the correct number of photons detected by the detector pair kl because we modeled the measurement process on $\partial\Omega \times \partial\Omega$. The measure κ_t is linked to the mass distribution via the forward operator described in the previous section. The conditional probability obtaining the measurement E for a given pair (ρ, ω) (actually only ρ is needed to apply the forward operator) is given by

$$\mathbb{P}\left(E|\rho,\omega\right) = \mathbb{P}\left(E|\rho\right) = \prod_{ikl} \mathcal{P}_{K_{ikl}}(E_{ikl})$$

where we used the independence property of PPPs on disjoint sets. Next, we determine the maximum a posteriori (MAP) estimate of (ρ, ω). Therefore, we maximize $\mathbb{P}\left(\rho,\omega|E\right)$ or equivalently minimize

$$\begin{aligned} -\log\left(\mathbb{P}\left(\rho,\omega|E\right)\right) &= -\log\left(\mathbb{P}\left(E|\rho,\omega\right)\right) - \log\left(\mathbb{P}\left(\rho,\omega\right)\right) + \mathbb{P}\left(E\right) \\ &\approx \sum_{ikl} -\log\left((K_{ikl})^{E_{ikl}} e^{-K_{ikl}}/(E_{ikl})!\right) + \beta\mathcal{S}(\rho,\omega) \\ &= \sum_{ikl} \left[K_{ikl} - E_{ikl}\log\left(K_{ikl}\right)\right] + \beta\mathcal{S}(\rho,\omega) \end{aligned}$$

where we neglected the constant $\mathbb{P}\left(E\right)$ and the constants coming from the factors of the model for $\mathbb{P}\left(\rho,\omega\right)$ as they will not matter in the following optimization. Moreover, we used the assumption $E_{ikl} \in \{0,1\}$ in order to not having the term $\sum_{ikl}\log\left(E_{ikl}!\right)$ in our functional. To point out all dependencies on the mass distribution, we set

$$A_{ikl}\rho = K_{ikl}.$$

This leads to the function

$$\hat{J}^E(\rho,\omega) = \sum_{ikl} \left[A_{ikl}\rho - E_{ikl}\log\left(A_{ikl}\rho\right)\right] + \beta\mathcal{S}(\rho,\omega)$$

to be minimized.

The MAP estimate resulting from the minimization of $\hat{J}$ fails to detect scattered photon pairs properly and instead declares most of the photon pairs to be unscattered. This is due to the part of $A_{ikl}\rho$ resulting from p^sA^s being very small compared to the other one coming from p^dA^d. This happens because A^s distributes the intensity evenly across all detectors whereas A^d concentrates intensity to a few suiting detectors only.

To reduce this problem, we split the measurement

$$E = E^s + E^d$$

into scattered and normally detected photon pairs and assume that we know the number K^s of scattered measurements. We now repeat the derivation of a MAP estimate but reconstruct (ρ, ω, E^s) this time for a fixed number K^s of scattered photon pairs. We have now

$$\mathbb{P}\left(\rho,\omega,E^s|E\right) = \frac{\mathbb{P}\left(E,E^s|\rho,\omega\right)\mathbb{P}\left(\rho,\omega\right)}{\mathbb{P}\left(E\right)}.$$

Assuming that scattering and normal detection happen independently, i.e. the probability of scattering is independent of the number of already scattered photons, and using the independence property of PPPs on disjoint sets, we find

$$\mathbb{P}\left(E, E^s|\rho,\omega\right) = \prod_{ikl} \mathcal{P}_{K^s_{ikl}}(E^s_{ikl})\mathcal{P}_{K^d_{ikl}}(E^d_{ikl})$$

with the definitions

$$K^{s/d}_{ikl} = \frac{\ln 2}{T_{1/2}} \int_{\tau_i} \kappa^{s/d}_t(\Gamma_{kl})dt = \frac{\ln 2}{T^{1/2}} \int_{\tau_i} p^{s/d} A^{s/d} \rho_t(\Gamma_{kl})dt.$$

Setting

$$A^{s/d}_{ikl}\rho = K^{s/d}_{ikl} \quad \text{and} \quad A_{ikl}\rho = A^s_{ikl}\rho + A^d_{ikl}\rho$$

to emphasize the dependence on the material distribution ρ, we arrive at the functional

$$\bar{J}^E(\rho,\omega,E^s) = \sum_{ikl} \left[A_{ikl}\rho - E^s_{ikl}\log\left(A^s_{ikl}\rho\right) - E^d_{ikl}\log\left(A^d_{ikl}\right)\right] + \beta\mathcal{S}(\rho,\omega)$$

by performing similar computations as in the above derivation of $\hat{J}$. This time, the functional $\bar{J}^E$ is minimized for ρ, ω and E^s taking into account the constraint

$$|E^s|_1 = \sum_{ikl} E^s_{ikl} = K^s.$$

Note, that we again assumed $E_{ikl} \in \{0,1\}$ in the derivation of $\bar{J}^E$. Two more modifications of our objective function are in order. First, the optimization of $\bar{J}^E$ is computationally costly as it involves the combinatorial optimization over all possible E^s. Therefore, we replace E^s by a tuning parameter $p > 0$ that modulates the importance of the scattering part of the forward operator. This yields the new functional

$$\tilde{J}^E(\rho,\omega,E^s) = \sum_{ikl} \left[A_{ikl}\rho - E_{ikl}\log\left(\max(pA^s_{ikl}\rho, A^d_{ikl}\rho)\right)\right] + \beta\mathcal{S}(\rho,\omega).$$

The interpretation of this simplification is as follows: Small values of p close to 0 result in the maximum $\max(pA^s_{ikl}\rho, A^d_{ikl}\rho)$ to be evaluated to the second term, i.e. the detection part, meaning that all events are declared as properly detected photon pairs. For large values of p the opposite interpretation of the detected events will be chosen. Thus, choosing an intermediate value of p results in a reasonable amount of detected scatter events.
Finally, we convexify our functional by replacing the log-part with its convex envelope. This leads to the final objective function

$$J^E(\rho,\omega) = \sum_{ikl} \left[A_{ikl}\rho - E_{ikl}\log\left(pA^s_{ikl}\rho + A^d_{ikl}\rho\right)\right] + \beta\mathcal{S}(\rho,\omega). \tag{2.16}$$

Therefore, we get the distribution of the radioactive material $\hat{\rho}$ from the PET measurement E by minimizing J^E over $\mathbb{M} = \mathcal{M}_+([0,T]\times\Omega) \times \mathcal{M}([0,T]\times\Omega)^3$, leading to

$$(\hat{\rho},\hat{\omega}) \in \operatorname*{argmin}_{(\rho,\omega)\in\mathbb{M}} J^E(\rho,\omega). \tag{MIN}$$

By considering $(\ln 2\rho, \ln 2\omega)$ instead of (ρ,ω) and $\frac{\beta}{\ln 2}$ instead of β, we can omit the constant factor $\ln 2$ in the definitions of A_{ikl} and $A^{s/d}_{ikl}$ in the following considerations.

Remark 2.3.3 (Time continuity of ρ)**.** Minimizers of the objective function J^E are satisfying the continuity equation due to the regularization term $\mathcal{S}$. Lemma 2.1.15 then guarantees the existence of a narrowly continuous representative $\tilde{\rho}_t$ of ρ_t. It holds for all open intervals $(a, b) \subset [0, T]$ and all open subsets $Z \subset \Omega$

$$\rho((a, b) \times Z) = \int_a^b \int_Z d\rho_t dt = \int_a^b \int_Z d\tilde{\rho}_t dt = \tilde{\rho}((a, b) \times Z).$$

Thus, both measures coincide on the generating system of $\mathcal{B}([0, T] \times \Omega)$ which means $\rho = \tilde{\rho}$ and we can assume that ρ_t is narrowly continuous in time.

2.4 Γ-Convergence

In the Sections 3.2 and 3.3 we are interested in the behavior of the reconstruction functional J^E when the half-life goes to zero, i.e. the intensity of the radioactive material goes to infinity. We are going to use the notion of Γ-convergence to analyze this limit process as the theory of Γ-convergence is well suited for investigating limit processes in minimization problems. Therefore, we define Γ-convergence and give an important result related to it in this section.

Definition 2.4.1 (Sequential Γ-convergence, [18])**.** *Let X be a topological space. Then the sequence $(f_i)_i$ of functions from X to $\mathbb{R} \cup \{+\infty\}$ sequentially Γ-converges to $f\colon X \mapsto \mathbb{R} \cup \{+\infty\}$ if we have*

(i) for every $x \in X$ and for every sequence $(x_i)_i$ converging to x in X it is

$$f(x) \leq \liminf_{i\to\infty} f_i(x_i);$$

(ii) for every $x \in X$ there exists sequence $(x_i)_i$ converging to x in X such that

$$f(x) \geq \limsup_{i\to\infty} f_i(x_i).$$

If the Γ-converging sequence of functions $(f_i)_i$ satisfies some additional properties, then we have that minimizers of the f_i converge to minimizers of the limit functional f.

Theorem 2.4.2 (Fundamental theorem of Γ-convergence, [18])**.** *Let X be a topological space and let $(f_i)_{i\in\mathbb{N}}$ be a sequence of equi-mildly coercive functions from X to $\mathbb{R} \cup \{+\infty\}$, i.e. there exists a countably compact set $K \subset X$ such that*

$$\inf_{x\in X} f_i(x) = \inf_{x\in K} f_i(x) \quad \textit{for all } i \in \mathbb{N}.$$

If $(f_i)_i$ sequentially Γ-converges to a function $f : X \to \mathbb{R} \cup \{+\infty\}$, then f has a minimizer in X and it holds

$$\min_{x\in X} f(x) = \min_{x\in K} f(x) = \lim_i \inf_{x\in X} f_i(x).$$

Moreover, if $(x_i)_i$ is a precompact sequence such that $\lim_i f_i(x_i) = \lim_i \inf_X f_i$, then every cluster point of the sequence $(x_i)_i$ is a minimizer of f.

Proof. Let K be a sequentially compact subset of X such that $\inf_{x\in X} f_i(x) = \inf_{x\in K} f_i(x)$ for all $i \in \mathbb{N}$. We start with showing two auxiliary inequalities. Let (x_i) be a sequence in K such that $\liminf_i \inf_{x\in K} f_i(x) = \liminf_i f_i(x_i)$. The existence of such a sequence is guaranteed by the following argument. Consider

$$\liminf_i \inf_{x\in K} f_i(x) = \sup_i \inf_{n\geq i} \inf_{x\in K} f_n(x) = \sup_i \inf_{n\geq i, x\in K} f_n(x).$$

There exists a sequence $(n_j, x_j)_j \subset \{n \in \mathbb{N} \mid n \geq i\} \times K$ such that $(f_{n_j}(x_{n_j}))_j$ is decreasing with limit $\inf_{n\geq i, x\in K} f_n(x)$. By assigning x_n an arbitrary element of K whenever $n \neq n_j$ we get

$$\inf_{n\geq i, x\in K} f_n(x) = \lim_j f_{n_j}(x_{n_j}) = \inf_j f_{n_j}(x_{n_j}) = \inf_{n\geq i} f_n(x_n)$$

and thus finally $\liminf_i \inf_{x\in K} f_i(x) = \liminf_i f_i(x_i)$. Further, let $(x_{i_j})_j$ be a subsequence such that $\lim_j f_{i_j}(x_{i_j}) = \liminf_i f_i(x_i)$. We can assume (x_{i_j}) to be convergent with limit $\bar{x}$ by using the sequential compactness of K. Next, define another sequence via

$$y_i = \begin{cases} x_{i_j} & \text{if } i = i_j \text{ for some } j, \\ \bar{x} & \text{if } i \neq i_j \text{ for all } j \end{cases}$$

that converges to $\bar{x}$ as well. Using the liminf condition it holds

$$\inf_{x\in K} f(x) \le f(\bar{x}) \le \liminf_i f_i(y_i) \le \liminf_j f_{i_j}(x_{i_j}) = \lim_j f_{i_j}(x_{i_j}) = \liminf_i \inf_{x\in K} f_i(x).$$

Thus, we get the first auxiliary inequality

$$\inf_{x\in K} f(x) \le \liminf_i \inf_{x\in K} f_i(x). \tag{2.17}$$

For the second inequality fix $\delta > 0$ and take $x \in X$ with $f(x) \le \inf_{x\in X} f(x) + \delta$. Further, let $(x_i)_i$ be a recovery sequence. Then we get

$$\inf_{x\in X} f(x) + \delta \ge f(x) \ge \limsup_i f_i(x_i) \ge \limsup_i \inf_{x\in X} f_i(x).$$

Since δ was arbitrary, it follows

$$\inf_{x\in X} f(x) \ge \limsup_i \inf_{x\in X} f_i(x). \tag{2.18}$$

With these results we get

$$\inf_{x\in X} f(x) \le \inf_{x\in K} f(x) \le \liminf_i \inf_{x\in K} f_i(x) = \liminf_i \inf_{x\in X} f_i(x) \le \limsup_i \inf_{x\in X} f_i(x) \le \inf_{x\in X} f(x),$$

implying that $\inf_{x\in X} f(x) = \inf_{x\in K} f(x) = \lim_i \inf_{x\in X} f_i(x)$ holds.
Next, take a precompact sequence $(x_i)_i$ such that $\lim_i f_i(x_i) = \lim_i \inf_X f_i$ and consider a converging subsequence $(x_{i_j})_j$ of $(x_i)_i$ with limit $\bar{x}$. Then we define the sequence

$$\tilde{x}_i = \begin{cases} x_{i_j} & \text{if } i = i_j, \\ \bar{x} & \text{if } i \neq i_j \text{ for all } j \end{cases}$$

which converges to $\bar{x}$ as well. By the liminf condition and with equations (2.17),(2.18) we get

$$\inf_{x\in X} f(x) \le f(\bar{x}) \le \liminf_i f_i(\tilde{x}_i) \le \liminf_j f_{i_j}(x_{i_j}) = \lim_i f_i(x_i) = \lim_i \inf_{x\in X} f_i(x) = \inf_{x\in X} f(x),$$

proving that $\bar{x}$ is a minimizer of f.
Finally, we use the last result to prove that there always is a minimizer of f. As the f_i are equi-mildly coercive, there is a precompact sequence $(x_i)_i \subset K$ with

$$f_i(x_i) \le \inf_{x\in X} f_i(x) + \frac{1}{i}$$

for all $i \in \mathbb{N}$ and such that $\lim_i f_i(x_i) = \lim_i \inf_X f_i$. The above result shows that $(x_i)_i$ has a cluster point that is a minimizer of f. This proves $\inf_{x\in K} f(x) = \min_{x\in K} f(x)$. □

CHAPTER 3

Existence of Minimizers and Γ-Convergence

In this chapter we analyze the minimization problem (MIN). We prove existence of minimizers of the objective function J^E and compute Γ-limits for the case of half-lifes going to zero (which means that the intensity of the radioactive material goes to infinity).
First, we reformulate the functional J^E in order to emphasize the stochastic character of this functional due to the randomness of the radioactive decay and hence the randomness of the PET measurement. We consider

$$\begin{aligned} J^E &: \mathbb{M} \to \mathbb{R} \\ (\rho,\omega) &\mapsto \sum_{ikl} \Big[A_{ikl}\rho - E_{ikl}\log(pA^s_{ikl}\rho + A^d_{ikl}\rho) \Big] + \beta\mathcal{S}(\rho,\omega). \end{aligned}$$

For the reformulation and for the next sections we introduce some abbreviation. We write

$$\kappa^p_t = p\kappa^s_t + \kappa^d_t$$

and

$$\kappa^{n,p}_t = p\kappa^{n,s}_t + \kappa^{n,d}_t$$

if the radioactive material distributions ρ^n depend on n. Moreover, we introduce piecewise constant functions for the scattering part, detection part and the above defined composition of the two parts. It is

$$\hat{\kappa}^{s/d} = \sum_{ikl} \mathbb{1}_{\tau_i \times \Gamma_{kl}} \int_{\tau_i} \kappa^{s/d}_t(\Gamma_{kl})dt \quad \text{and} \quad \hat{\kappa}^p = p\hat{\kappa}^s + \hat{\kappa}^d.$$

As above, $\hat{\kappa}^{n,s/d/p}$ indicates that the underlying material distributions ρ^n depend on n.
The first part of J^E can be rewritten as

$$\begin{aligned} \sum_{ikl} A_{ikl}\rho &= \sum_{kl}\sum_i \int_{\tau_i} \frac{1}{T_{1/2}}\kappa_t(\Gamma_{kl})dt = \sum_{kl}\int_0^T \frac{1}{T_{1/2}}\kappa_t(\Gamma_{kl})dt = \int_0^T \frac{1}{T_{1/2}}\kappa_t(\cup_{kl}\Gamma_{kl})dt \\ &= \frac{1}{T_{1/2}}\kappa([0,T]\times\Gamma_{\texttt{tot}}) \end{aligned}$$

where $\Gamma_{\texttt{tot}} = \cup_{kl}\Gamma_{kl}$ is the total area covered by detectors. The above expression gives us the expected number of detected photons during the time interval $[0,T]$. If we ignore the assumption $E_{ikl} \in \{0,1\}$, the second term can be expressed as

$$\begin{aligned} \sum_{ikl} E_{ikl}\log(pA^s_{ikl}\rho + A^d_{ikl}\rho) &= \sum_{ikl} E_{ikl}\log\left(\frac{1}{T^{1/2}}\int_{\tau_i}\kappa^p_t(\Gamma_{kl})dt\right) \\ &= \int \log\left(\frac{1}{T_{1/2}}\hat{\kappa}^p\right)dE \end{aligned}$$

with E now being a PPP (the measurement) with intensity measure $\frac{1}{T_{1/2}}\kappa_t^\dagger(x)dtd\phi$, $\phi = \mathcal{H}^2 \otimes \mathcal{H}^2$, where $\kappa^\dagger$ is th Radon-Nikodym derivative of the measure

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^\dagger(\Gamma)dt$$

w.r.t. $dt \otimes d\phi$ (it suffices to define the measure on the intersection stable generator $\{\tau \times \Gamma \mid \tau \in \mathcal{B}([0,T]),\ \Gamma \in \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta)\}$ [14, Lemma 1.42]).
Thus, we arrive at

$$J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}(\rho,\omega) = \frac{1}{T_{1/2}}\kappa([0,T] \times \Gamma_{\texttt{tot}}) - \int \log\left(\frac{1}{T_{1/2}}\hat{\kappa}^p\right) dE + \beta\mathcal{S}(\rho,\omega)$$

where we explicitly emphasize the dependence of the functional on the measurement process E, the temporal resolution Δt, the detectors $(\Gamma_k)_{kl}$, the regularization parameter β and the half-life $T_{1/2}$, because these parameters will be important when we are analyzing Γ-convergence later. Note, that the assumption $E_{ikl} \in \{0,1\}$ was only needed to derive the objective function. The function itself is valid for all non-negative values of E_{ikl}.

3.1 Existence of Minimizers

In this section we show that the functional J^E has a minimizer, i.e. our reconstruction method always produces a radioactive material distribution. We start by gathering some properties about the forward operator that we will need in the following.

Lemma 3.1.1 (Properties of the forward operators). *(a) The operators A_{ikl}^s and A_{ikl}^d are bounded, i.e. we have for $\rho \in \mathcal{M}_+([0,T] \times \Omega)$*

$$A_{ikl}^s\rho \leq \frac{p^s}{T_{1/2}}\|\rho\| \quad \textit{and} \quad A_{ikl}^d\rho \leq \frac{p^d}{T_{1/2}}\|\rho\|$$

(b) Let $\lambda^n \overset{*}{\rightharpoonup} \lambda$ in $\mathcal{M}_+(\Omega)$. Then $A^s\lambda^n \overset{*}{\rightharpoonup} A^s\lambda$ and $A^d\lambda^n \overset{*}{\rightharpoonup} A^d\lambda$ in $\mathcal{M}_+(\partial\Omega \times \partial\Omega)$.

(c) Let $\rho^n \overset{*}{\rightharpoonup} \rho$ in $\mathcal{M}_+([0,T] \times \Omega)$ and $\Gamma \subset \partial\Omega \times \partial\Omega$. Define the operators

$$\begin{aligned} A_\Gamma^s \colon \mathcal{M}_+([0,T] \times \Omega) &\to \mathcal{M}_+([0,T]) \\ \rho &\mapsto A_\Gamma^s\rho, \quad A_\Gamma^s\rho(S) = \int_S (A^s\rho_l)(\Gamma)dl \end{aligned}$$

and

$$\begin{aligned} A_\Gamma^d \colon \mathcal{M}_+([0,T] \times \Omega) &\to \mathcal{M}_+([0,T]) \\ \rho &\mapsto A_\Gamma^d\rho, \quad A_\Gamma^d\rho(S) = \int_S (A^d\rho_l)(\Gamma)dl. \end{aligned}$$

Then $A_\Gamma^s\rho^n \overset{*}{\rightharpoonup} A_\Gamma^s\rho$ and $A_\Gamma^d\rho^n \overset{*}{\rightharpoonup} A_\Gamma^d\rho$ in $\mathcal{M}([0,T])$. In particular, the operators A_{ikl}^s and A_{ikl}^d are weak- continuous.*

(d) *Let $(\rho, \omega) \in \mathbb{M}$ with $\mathcal{S}(\rho, \omega) < +\infty$ and let $\Gamma \subset \partial\Omega \times \partial\Omega$ with piecewise C^1-boundary. Then the functions*

$$[0, T] \to \mathbb{R},\ t \mapsto A^s \rho_t(\Gamma) = \frac{1}{p^s} \kappa_t^s(\Gamma)$$

and

$$[0, T] \to \mathbb{R},\ t \mapsto A^d \rho_t(\Gamma) = \frac{1}{p^d} \kappa_t^d(\Gamma)$$

are continuous.

(e) *Let $(\rho, \omega) \in \mathbb{M}$ with $\mathcal{S}(\rho, \omega) < +\infty$ and let $\Gamma \subset \partial\Omega \times \partial\Omega$. Then for every $t \in [0, T]$ it holds*

$$A^s \rho_t(\Gamma) \lesssim \|\rho\| \quad \textit{and} \quad A^d \rho_t(\Gamma) \lesssim \|\rho\| .$$

Proof. (a) For the scattering part we find

$$A^s_{ikl}\rho = \frac{1}{T_{1/2}} \int_{\tau_i} p^s \rho_t(\Omega) \frac{\mathcal{H}^2 \otimes \mathcal{H}^2 (\Gamma_{kl})}{\mathcal{H}^2 (\partial\Omega)^2} dt \le \frac{p^s}{T_{1/2}} \int_0^T \rho_t(\Omega) dt = \frac{p^s}{T_{1/2}} \|\rho\| .$$

In the case of the detection part of the forward operator we compute, using Remark 2.3.1,

$$\begin{aligned} A^d_{ikl}\rho &= \frac{1}{T_{1/2}} \int_{\tau_i} p^d \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega G_y(x) d\rho_t(y) dx \mathrm{vol}_{G^{1,3}}(dv) dt \\ &\le \frac{p^d}{T_{1/2}} \int_0^T \int_{G^{1,3}} \int_\Omega \int_\Omega G_y(x) d\rho_t(y) dx \mathrm{vol}_{G^{1,3}}(dv) dt \\ &\le \frac{p^d}{T_{1/2}} \int_0^T \rho_t(\Omega) dt = \frac{p^d}{T_{1/2}} \|\rho\| , \end{aligned}$$

where we used $\int_{G^{1,3}} \mathrm{vol}_{G^{1,3}}(dv) = 1$ and $\int_\Omega G_y(x) dx = 1$ for all $y \in \Omega$.

(b) Let $\lambda^n \overset{*}{\rightharpoonup} \lambda$ in $\mathcal{M}(\Omega)$ and $\phi \in C(\partial\Omega \times \partial\Omega)$. From the weak-* convergence and the compactness of Ω we get

$$\lambda^n(\Omega) = \int_\Omega 1 d\lambda^n \xrightarrow{n \to \infty} \int_\Omega 1 d\lambda = \lambda(\Omega).$$

It follows

$$\begin{aligned} \int_{\partial\Omega \times \partial\Omega} \phi(x) A^s \lambda^n(dx) &= \frac{\lambda^n(\Omega)}{\mathcal{H}^2(\partial\Omega \times \partial\Omega)^2} \int_{\partial\Omega \times \partial\Omega} \phi(x) \mathcal{H}^2 \otimes \mathcal{H}^2 (dx) \\ &\xrightarrow{n \to \infty} \frac{\lambda(\Omega)}{\mathcal{H}^2(\partial\Omega \times \partial\Omega)^2} \int_{\partial\Omega \times \partial\Omega} \phi(x) \mathcal{H}^2 \otimes \mathcal{H}^2 (dx) = \int_{\partial\Omega \times \partial\Omega} \phi(x) A^s \lambda(dx). \end{aligned}$$

Moreover we have

$$\begin{aligned} \lim_{n\to\infty} \int_{\partial\Omega \times \partial\Omega} \phi(x) A^d \lambda^n(dx) &= \lim_{n\to\infty} \int_{\Omega \times G^{1,3}} \phi(R(x, v)) \int_\Omega G_y(x) d\lambda^n(y) dx \mathrm{vol}_{G^{1,3}}(dv) \\ &= \int_{\Omega \times G^{1,3}} \phi(R(x, v)) \lim_{n\to\infty} \int_\Omega G_y(x) d\lambda^n(y) dx \mathrm{vol}_{G^{1,3}}(dv) \\ &= \int_{\Omega \times G^{1,3}} \phi(R(x, v)) \int_\Omega G_y(x) d\lambda(y) dx \mathrm{vol}_{G^{1,3}}(dv) \\ &= \int_{\partial\Omega \times \partial\Omega} \phi(x) A^d \lambda(dx) \end{aligned}$$

by dominated converges with $|\phi(R(x,v)) \int_\Omega G_y(x) d\lambda^n| \leq \|\phi\|_\infty \|G\|_\infty \lambda^n(\Omega) \lesssim 1$ because of $\lambda^n \overset{*}{\rightharpoonup} \lambda$.

(c) Let $\rho, \rho^n \in \mathcal{M}_+([0,T] \times \Omega)$ such that $\rho \overset{*}{\rightharpoonup} \rho$ and $\psi \in C([0,T])$. Then

$$A_\Gamma^s \rho^n(\psi) = \frac{1}{T_{1/2}} \int_0^T \psi(t)(A^s \rho_t)(\Gamma) dt = \frac{1}{T_{1/2}} \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \int_0^T \psi(t) \rho_t^n(\Omega) dt$$
$$\xrightarrow{n\to\infty} \frac{1}{T_{1/2}} \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \int_0^T \psi(t) \rho_t(\Omega) dt = A_\Gamma^s \rho(\psi),$$

using the weak-* convergence of the ρ^n applied to the function $(t,x) \mapsto \psi(t)$.
We proceed with A_Γ^d in a similar way. Using the definition of A^d (see Remark 2.3.1), dominated convergence and weak-* convergence of ρ^n we get for $\psi \in C([0,T])$

$$\begin{aligned}
\lim_{n\to\infty} A_\Gamma^d \rho^n(\psi) &= \lim_{n\to\infty} \int_0^T \psi(t) \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega G_y(x) d\rho_t^n(y) dx \mathrm{vol}_{G^{1,3}}(dv) dt \\
&= \int_{G^{1,3}} \int_{Z_v^\Gamma} \lim_{n\to\infty} \int_0^T \int_\Omega \psi(t) G_y(x) d\rho_t^n(y) dx dt \mathrm{vol}_{G^{1,3}}(dv) \\
&= \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_0^T \int_\Omega \psi(t) G_y(x) d\rho_t(y) dx dt \mathrm{vol}_{G^{1,3}}(dv) = A_\Gamma^d \rho(\psi).
\end{aligned}$$

We could use Fubini and dominated convergence because we only have finite measure spaces and continuous functions as integrands. This shows the weak-* continuity of A_Γ^s and A_Γ^d.
The weak-* continuity of A_{ikl}^s and A_{ikl}^d follows from Portemanteau theorem [14, Theorem 13.16] as the time intervals' boundaries $\partial\tau_i$ have Lebesgue measure zero and the measures $A_\Gamma^{s/d}\rho$ are absolutely continuous w.r.t. the Lebesgue measure.

(d) Let $t^n \to t$. Since ρ satisfies the continuity equation, we know from Remark 2.3.3 that $(\rho_t)_t$ is a narrowly continuous curve in $\mathcal{M}(\Omega)$ and hence $\rho_{t^n} \overset{*}{\rightharpoonup} \rho_t$ in $\mathcal{M}(\Omega)$ as well. Part (b) yields $A^s \rho_{t^n} \overset{*}{\rightharpoonup} A^s \rho_t$ and $A^d \rho_{t^n} \overset{*}{\rightharpoonup} A^s \rho_t$. Using again the Portemanteau theorem we get

$$A^s \rho_{t^n}(\Gamma) \xrightarrow{n\to\infty} A^s \rho_t(\Gamma) \quad \text{and} \quad A^d \rho_{t^n}(\Gamma) \xrightarrow{n\to\infty} A^d \rho_t(\Gamma).$$

To see that the Portemanteau theorem can be applied, we consider the following. In remark 2.3.2 we have shown the absolute continuity of κ_t w.r.t. $\mathcal{H}^2 \otimes \mathcal{H}^2$. Since $\partial\Gamma$ is assumed to have C^1-boundary, it is rectifiable. This means we have $\mathcal{H}^1(\partial\Omega) < +\infty$ and hence $\mathcal{H}^2(\partial\Omega) = 0$ [16, Chapter 4]. This way we get $A^{s/d}\rho_t(\Gamma) = 0$ allowing us to apply the Portemanteau theorem.

(e) For the proof we use the continuity in time of ρ_t and mass conservation from Lemma 2.1.15. We find

$$A^s \rho_t(\Gamma) = \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \rho_t(\Omega) = \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \frac{1}{T} \int_0^T \rho_s(\Omega) ds \lesssim \|\rho\|$$

and

$$\begin{aligned} A^d\rho_t(\Gamma) &= \int_{G^{1,3}}\int_{Z_v^\Gamma}\int_\Omega G_y(x)d\rho_t(y)dx\mathrm{vol}_{G^{1,3}}(dv) \\ &\leq \int_{G^{1,3}}\int_\Omega\int_\Omega G_y(x)dxd\rho_t(y)\mathrm{vol}_{G^{1,3}}(dv) \\ &\lesssim \frac{1}{T}\int_0^T \rho_s(\Omega)ds \lesssim \|\rho\|\,. \end{aligned}$$

□

Theorem 3.1.2 (Existence of minimizers)**.** *The functional J^E has a minimizer in $\mathbb{M}$. Moreover, the (stochastic) functional $J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}$ almost surely has a minimizer in $\mathbb{M}$.*

Proof. We write J for either J^E or $J_{\beta,T_{1/2}}^{E,\Delta t,(\Gamma_{kl})_{kl}}$ in the following.
Note, that our measurement E is modeled by a finite Poisson point process because it is

$$\mathbb{P}\left(E([0,T]\times\partial\Omega\times\partial\Omega)<+\infty\right)=\mathbb{P}\left(\bigcup_{n=0}^{\infty}\{E([0,T]\times\partial\Omega\times\partial\Omega)=n\}\right)=\sum_{n=0}^{\infty}e^{-\lambda}\frac{\lambda^n}{n!}=1$$

with the mean number of points $\lambda=\int_0^T\kappa_t^\dagger(\partial\Omega\times\partial\Omega)$. This means we have $\sum_{ikl}E_{ikl}<+\infty$ almost surely and can thus prove existence of minimizers for both functionals simultaneously because the proof relies on $\max_{ikl}(E_{ikl})$ being finite.
We have $(\mathcal{L},0)\in\mathbb{M}$ with $\mathcal{L}$ being the Lebesgue measure on $[0,T]\times\Omega$. The pair $(\mathcal{L},0)$ satisfies the continuity equation in the distributional sense. This is established by

$$\int_{[0,T]\times\Omega}\partial_t\phi d\mathcal{L}=\int_\Omega\int_0^T\partial_t\phi dtdx=\int_\Omega(\phi(T,\cdot)-\phi(0,\cdot))dx$$

for all $\phi\in\mathcal{C}^1([0,T]\times\Omega)$. Moreover, we have $\mathcal{S}(\mathcal{L},0)=0$ and hence $J(\mathcal{L},0)<+\infty$.
Now let $(\rho^n,\omega^n)_{n\in\mathbb{N}}\subset\mathbb{M}$ be a minimizing sequence of J. By possibly extracting a subsequence we may assume w.l.o.g. that $J(\rho^n,\omega^n)\leq J(\mathcal{L},0)$. Next, we find uniform bounds on the norms of ρ^n and ω^n. It holds with $\bar{p}=\max(p,1)$ and $E_{\max}=\max(1,\max_{ikl}(E_{ikl}))$

$$\begin{aligned} J(\rho^n,\omega^n) &= \sum_{ikl}\left[A_{ikl}\rho^n-E_{ikl}\log(pA_{ikl}^s\rho^n+A_{ikl}^d\rho^n)\right]+\beta\mathcal{S}(\rho^n,\omega^n) \\ &\geq \sum_{ikl}\left[A_{ikl}\rho^n-E_{ikl}\log(\bar{p}A_{ikl}\rho^n)\right]+\beta\mathcal{S}(\rho^n,\omega^n) \\ &\geq \sum_{ikl}\left[A_{ikl}\rho^n\chi_{\{\bar{p}A_{ikl}\rho^n<1\}}+(A_{ikl}\rho^n-E_{\max}\log(\bar{p}A_{ikl}\rho^n))\chi_{\{\bar{p}A_{ikl}\rho^n\geq 1\}}\right] \\ &\quad+\beta\mathcal{S}(\rho^n,\omega^n). \end{aligned}$$

To continue, let f be an affine strictly increasing function such that $m-E_{\max}\log(\bar{p}m)\geq f(m)$ and $m\geq f(m)$. We get

$$\begin{aligned} J(\rho^n,\omega^n) &\geq \sum_{ikl}f(A_{ikl}\rho^n)+\beta\mathcal{S}(\rho^n,\omega^n) \\ &\geq \sum_{ikl}f(A_{ikl}^s\rho^n)+\beta\mathcal{S}(\rho^n,\omega^n) \\ &= \sum_{ikl}f(c_{kl}\rho^n(\tau_i\times\Omega))+\beta\mathcal{S}(\rho^n,\omega^n) \end{aligned}$$

for positive constants c_{kl} which can be calculated from the definition of the scattering operator A^s. The c_{kl} read

$$c_{kl} = \frac{p^s}{T^{1/2}} \frac{\mathcal{H}^2 \otimes \mathcal{H}^2 \left(\Gamma_{kl}\right)}{\mathcal{H}^2 \otimes \mathcal{H}^2 \left(\partial\Omega \times \partial\Omega\right)}.$$

Since f is affine and strictly increasing and $J(\rho^n, \omega^n)$ is uniformly bounded in n, we also have uniform boundedness of $\|\rho^n\| = \sum_i \rho^n(\tau_i \times \Omega)$. In addition, we deduce, using Hölder's inequality,

$$\|\omega^n\| = \left\|\frac{d\omega^n}{d\rho^n}\right\|_{L^1(\rho^n)} \leq \left\|\frac{d\omega^n}{d\rho^n}\right\|_{L^2(\rho^n)} \|1\|_{L^2(\rho^n)} = \underbrace{\left\|\frac{d\omega^n}{d\rho^n}\right\|_{L^2(\rho^n)}}_{=\mathcal{S}(\rho^n,\omega^n)} \rho^n([0,T] \times \Omega)^{\frac{1}{2}}$$

which shows the unifrom boundedness of $\|\omega^n\|$. By Prokhorov's theorem there exists a subsequence (still indexed by n) and $(\rho, \omega) \in \mathbb{M}$ such that $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$. Since the operators A_{ikl}, A^s_{ikl}, A^d_{ikl} are weak-* continuous (Lemma 3.1.1) and the functional $\mathcal{S}$ is weak-* lower semi-continuous (Lemma 2.1.14), we finally arrive at

$$J(\rho, \omega) \leq \liminf_{n\to\infty} J(\rho^n, \omega^n) \leq J(\bar{\rho}, \bar{\omega})$$

for any $(\bar{\rho}, \bar{\omega}) \in \mathbb{M}$, which shows that (ρ, ω) minimizes J.

□

3.2 Γ-Convergence I: Fixed Detector Sizes

In this section we prove Γ-convergence for the case of a sequence of half-lifes $(T^n)_{n\in\mathbb{N}}$ tending to zero (which means we have an intensity of radiation tending to infinity). Within this limit process, we change the size of our temporal resolution Δt^n but keep the detector sizes fixed. In the following section we will also include changes in the detector sizes.

Lemma 3.2.1. *Let (ρ^n, ω^n) with $\mathcal{S}(\rho^n, \omega^n) < +\infty$. Then the function $t \mapsto \kappa^{n,d}_t(\Gamma)$ is in $W^{1,1}((0,T))$ with weak derivative*

$$t \mapsto p^d \int_{G^{1,3}} \int_{Z^\Gamma_v} \int_\Omega \langle \nabla_y G_y(x), d\omega^n_t(y) \rangle dx vol_{G^{1,3}}(dv).$$

Proof. From $\mathcal{S}(\rho^n, \omega^n) < +\infty$ we get that (ρ^n, ω^n) satisfies the continuity equation and is disintegrable in time. Let $\varphi \in C^\infty_c((0,T))$. Using the definition of the forward operator (see Remark 2.3.1) and the continuity equation, we find

$$\begin{aligned}\int_0^T \varphi'(t)\kappa_t^n(\Gamma)dt &= p^d \int_0^T \varphi'(t) \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega G_y(x) d\rho_t^n(y) dx \mathrm{vol}_{G^{1,3}}(dv) dt \\ &= p^d \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_0^T \int_\Omega \varphi'(t) G_y(x) d\rho_t^n(y) dt dx \mathrm{vol}_{G^{1,3}}(dv) \\ &= -p^d \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_0^T \int_\Omega \langle \varphi(t) \nabla_y G_y(x), d\omega_t^n(y)\rangle dt dx \mathrm{vol}_{G^{1,3}}(dv) \\ &= -p^d \int_0^T \varphi(t) \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega \langle \nabla_y G_y(x), d\omega_t^n(y)\rangle dx \mathrm{vol}_{G^{1,3}}(dv) dt\end{aligned}$$

where Fubini's theorem is applicable due to the finiteness of all measures and the smoothness of the involved functions. For the same reason the weak derivative is in $L^1((0,T))$. □

Lemma 3.2.2. *Let $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$ with $\mathcal{S}(\rho^n, \omega^n) \leq M$ for every n and $\mathcal{S}(\rho, \omega) < +\infty$. Moreover, let $\Gamma \subset \partial\Omega \times \partial\Omega$ with piecewise C^1-boundary. Then the sequence of functions*

$$[0,T] \to [0, +\infty), \ t \mapsto \kappa_t^{n,p}(\Gamma)$$

converges (up to a subsequence) uniformly to the function $t \mapsto \kappa_t^p(\Gamma)$.
Analogous results hold for the scattering and detection part of κ_t as well.

Proof. We start with the scattering part. Using mass conservation it is

$$\kappa_t^{n,s}(\Gamma) = p^s \frac{\rho_t^n(\Omega)}{\mathcal{H}^2(\partial\Omega)^2} \mathcal{H}^2 \otimes \mathcal{H}^2 (\Gamma) = p^s \frac{\mathcal{H}^2 \otimes \mathcal{H}^2 (\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \frac{1}{T} \int_0^T \rho_l^n(\Omega) dl.$$

With this we get for any $\varepsilon > 0$ and $t \in [0,T]$

$$|\kappa_t^{n,s}(\Gamma) - \kappa_t^s(\Gamma)| = p^s \frac{\mathcal{H}^2 \otimes \mathcal{H}^2 (\Gamma)}{\mathcal{H}^2(\partial\Omega)^2} \frac{1}{T} \left| \int_0^T \rho_l^n(\Omega) dl - \int_0^T \rho_l(\Omega) dl \right| < \varepsilon$$

for n large enough due to $\rho^n \overset{*}{\rightharpoonup} \rho$.
For the detection part we want to apply the Arzelà-Ascoli theorem. Therefore, we need to show uniform boundedness and equicontinuity of the family of functions $(\kappa_t^{n,d}(\Gamma))_n$.
From Lemma 3.1.1 we get the boundedness

$$0 \leq \kappa_t^{n,d}(\Gamma) \lesssim \|\rho^n\| = \int_0^T \rho_l^n(\Omega) dl$$

with the last term being uniformly bounded in t due to the weak-* convergence $\rho^n \overset{*}{\rightharpoonup} \rho$.
For the equicontinuity, we use Lemma 3.2.1 to apply the fundamental theoreom of calculus which holds since the function $t \mapsto \kappa_t^{n,d}(\Gamma)$ is continuous ([19, Theorem 8.2] and Lemma

3.1.1). This yields for $s < t$

$$
\begin{aligned}
\left|\kappa_t^{n,d}(\Gamma) - \kappa_s^{n,d}(\Gamma)\right| &\leq \left| p^d \int_s^t \int_{G^{1,3}} \int_{Z_v^\Gamma} \int_\Omega \langle \nabla_y G_y(x), d\omega_l(y)\rangle dx \mathrm{vol}_{G^{1,3}}(dv) dl \right| \\
&\leq p^d R_\#(dx \otimes \mathrm{vol}_{G^{1,3}})(\Gamma) \left\|\nabla_y G\right\|_\infty \int_s^t \int_\Omega \left|\frac{d\omega_l^n}{d\rho_l^n}\right| d\rho_l^n dl \\
&\overset{\star}{\lesssim} (t-s)^{1/2} \left(\int_s^t \left(\int_\Omega \left|\frac{d\omega_l^n}{d\rho_l^n}\right| d\rho_l^n \right)^2 dl \right)^{1/2} \\
&\overset{\star\star}{\leq} (t-s)^{1/2} \left(\int_s^t \rho_l^n(\Omega)^2 \left(\int_\Omega \left|\frac{d\omega_l^n}{d\rho_l^n}\right|^2 d\rho_l^n / \rho_l^n(\Omega) \right) dl \right)^{1/2} \\
&\overset{\star\star\star}{\lesssim} (t-s)^{1/2} \left(\rho^n([0,T]\times\Omega)\right)^{1/2} \left(\int_0^T \left(\int_\Omega \left|\frac{d\omega_l^n}{d\rho_l^n}\right|^2 d\rho_l^n \right) dl \right)^{1/2} \\
&\lesssim (t-s)^{1/2} M
\end{aligned}
$$

where we used Hölder's inequality in ($\star$), Jensen's inequality in ($\star\star$), mass conservation in ($\star\star\star$) and $\rho^n \overset{*}{\rightharpoonup} \rho$ in the last step to find a uniform estimate. This establishes the equicontinuity. The Arzelà-Ascoli theorem then gives us a subsequence (again indexed by n) such that $\kappa_t^{n,d}(\Gamma) \to \kappa_t^{d,\infty}(\Gamma)$ uniformly. We are left to show that $\kappa_t^{d,\infty}(\Gamma) = \kappa_t^d(\Gamma)$. This is established by the weak-* continuity of the forward operator. Indeed, we have for $E \subset [0,T]$

$$
\int_E \kappa_t^{d,\infty}(\Gamma) dt = \lim_{n\to\infty} \int_E \kappa_t^{n,d}(\Gamma) dt = \int_E \kappa_t^d(\Gamma) dt
$$

for all $E \subset [0,T]$, where we used dominated convergence in the first step (applied to $\kappa_t^{n,d}(\Gamma) \lesssim \frac{1}{T}\int_0^T \rho_l^n(\Omega) dl \lesssim 1$) and weak-* convergence in the second step (since the Lebesgue measure of ∂E is zero for every E and $(E \mapsto \int_E \kappa_t^d(\Gamma) dt) \ll dt$, the Portemanteau theorem [14] can be applied). This shows $\kappa_t^\infty(\Gamma) = \kappa_t(\Gamma)$ for almost all $t \in [0,T]$ ([14, Lemma 1.42]) and hence all t as the function is continuous by Lemma 3.1.1. □

Remark 3.2.3 (L^p-regularization)**.** We use some sort of L^2-regularization for our regularization term

$$
\mathcal{S}(\rho,\omega) = \int_0^T \int_\Omega \left|\frac{d\omega_t}{d\rho_t}\right|^2 d\rho_t dt.
$$

This way, we penalize mass movement or kinetic motion energy which is a realistic idea for finding a good reconstruction.
Mathematically, L^p-regularizations of higher order are also possible without any change of arguments because Hölder's and Jensen's inequality is still applicable in the above proof. Thus, we could use as well

$$
\mathcal{S}_p(\rho,\omega) = \int_0^T \int_\Omega \left|\frac{d\omega_t}{d\rho_t}\right|^p d\rho_t dt \quad \text{for } p \geq 2
$$

as the regularization term. The assumed uniform boundedness $\mathcal{S}_p(\rho^n,\omega^n) \lesssim 1$ in Lemma 3.2.2 will be fulfilled for an infimizing sequence in the framework of Γ-convergence and it is still independent of the choice of the order p.

Lemma 3.2.4. *Let $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$ with $\mathcal{S}(\rho^n, \omega^n) \le M$ for every n and $\mathcal{S}(\rho, \omega) < +\infty$. For $t \in [0, T]$ let $(\tau^n)_n$ be a sequence of intervals $\tau^n \subset [0, T]$ such that $\tau^n \xrightarrow{n\to\infty} t$. Then we have for each $\Gamma \subset \partial\Omega \times \partial\Omega$ with piecewise C^1-boundary*

$$\lim_{n\to\infty} \frac{1}{|\tau^n|} \int_{\tau^n} \kappa_t^{n,p}(\Gamma)dt \to \kappa_t^p(\Gamma). \tag{3.1}$$

Proof. It suffices to show the result for a subsequence. Indeed, let $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$. Then $(\rho^{n_j}, \omega^{n_j}) \overset{*}{\rightharpoonup} (\rho, \omega)$ for every subsequence indexed by n_j and we get

$$\lim_{l\to\infty} \frac{1}{|\tau^{n_{j_l}}|} \int_{\tau^{n_{j_l}}} \kappa_t^{n_{j_l},p}(\Gamma)dt \to \kappa_t^p(\Gamma) \tag{3.2}$$

for a further subsequence n_{j_l}. Thus, every subsequence has a further subsequence converging to $\kappa_t^p(\Gamma)$, implying that the whole sequence converges to this value.
Now both, the scattering and the detection part are continuous in time by Lemma 3.1.1. Thus, by the mean value theorem we get

$$\frac{1}{|\tau^n|} \int_{\tau^n} \kappa_t^{n,p}(\Gamma)dt = \kappa_{t^n}^{n,p}(\Gamma)$$

for a sequence $(t^n)_n \subset [0, T]$ with $t^n \in \tau^n$ and $t^n \xrightarrow{n\to\infty} t$. According to Lemma 3.2.2, there exists a subsequence (again indexed by n) such that $t \mapsto \kappa_t^{n,p}(\Gamma)$ converges uniformly to $t \mapsto \kappa_t^p(\Gamma)$. Using this and the continuity of $t \mapsto \kappa_t^p(\Gamma)$, we have

$$\begin{aligned} |\kappa_{t^n}^{n,p}(\Gamma) - \kappa_t^p(\Gamma)| &\le |\kappa_{t^n}^{n,p}(\Gamma) - \kappa_{t^n}^p(\Gamma)| + |\kappa_{t^n}^p(\Gamma) - \kappa_t^p(\Gamma)| \\ &\le \sup_{l\in[0,T]} \left|\kappa_l^{n,p}(\Gamma) - \kappa_l^p(\Gamma)\right| + |\kappa_{t^n}^p(\Gamma) - \kappa_t^p(\Gamma)| \xrightarrow{n\to\infty} 0. \end{aligned}$$

□

The measurement process is modeled stochastically. Therefore, we have a different functional to be minimized for each possible measurement. To cope with this stochastic behavior, we need to analyze convergence in a stochastic framework, meaning that we only expect to have almost sure convergence of our functionals. Let $\mathcal{E}^n(\rho, \omega, \xi)$ be the stochastic functionals that should Γ-converge to $\mathcal{E}^\infty(\rho, \omega, \xi)$. By Γ-convergence of $\mathcal{E}^n$ to $\mathcal{E}^\infty$ almost surely we mean the following. For every $(\rho, \omega) \in \mathbb{M}$ and every sequence $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$, it holds

$$\mathbb{P}\left(\xi \in X \mid \liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n, \xi) \ge \mathcal{E}^\infty(\rho, \omega, \xi)\right) = 1$$

and there exists a sequence $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$ such that

$$\mathbb{P}\left(\xi \in X \mid \limsup_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n, \xi) \le \mathcal{E}^\infty(\rho, \omega, \xi)\right) = 1.$$

Next, we prove Γ-convergence for temporal resolutions Δt^n converging to zero while the intensity of radiation of the radioactive material tends to infinity. The PPP E^n producing the measurement for a given half-life T^n (with $T^n \to 0$) has the intensity measure $\frac{1}{T^n}\kappa_t^\dagger(x)dtd\phi$ where $\kappa^\dagger$ is the Radon-Nikodym derivative of the measure

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega \times \partial\Omega) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^\dagger(\Gamma)dt.$$

w.r.t. $d\phi \otimes dt$ where $\phi = \mathcal{H}^2 \otimes \mathcal{H}^2$.

Theorem 3.2.5 (Γ-convergence)**.** *Let $(T^n)_n$ be a sequence of half-lifes with $\sum_n (T^n)^m < +\infty$ for some $m \in \mathbb{N}$, $T^n\beta^n \to \beta > 0$, let E^n be a Poisson point process with intensity measure $\frac{1}{T^n}\kappa_t^\dagger(x)dtd\phi$ and $\Delta t^n \to 0$ for a decreasing sequence of temporal resolutions $(\Delta t^n)_n$. We set*

$$C^n = T^n \log\left(\frac{\Delta t^n}{T^n}\right) E^n([0,T] \times \Gamma_{\texttt{tot}}).$$

Then, with respect to the weak- convergence on $\mathbb{M}$,*

$$\Gamma - \lim_{n\to\infty} T^n J_{\beta^n,T^n}^{E^n,\Delta t^n,(\Gamma_{kl})_{kl}} + C^n = \mathcal{E}^\infty$$

almost surely. The limit functional reads

$$\begin{aligned}\mathcal{E}^\infty(\rho,\omega) &= \sum_{kl} \int_0^T \left[\kappa_t(\Gamma_{kl}) - \log\left(\kappa_t^p(\Gamma_{kl})\right)\kappa_t^\dagger(\Gamma_{kl})\right] dt + \beta\mathcal{S}(\rho,\omega) \\ &= \kappa([0,T]\times\Gamma_{\texttt{tot}}) - \sum_{kl}\int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right)\kappa_t^\dagger(\Gamma_{kl})dt + \beta\mathcal{S}(\rho,\omega).\end{aligned}$$

Proof. During the proof we state different things almost surely. Since we are doing this finitely often only, the union of all null sets on which the statements do not hold is again a null set. This means that all statements together still hold almost surely.
For the actual proof we define

$$\begin{aligned}\mathcal{E}^n(\rho,\omega) &= T^n J_{\beta^n,T^n}^{E^n,\Delta t^n,(\Gamma_{kl})_{kl}}(\rho,\omega) + C^n \\ &= \kappa([0,T]\times\Gamma_{\texttt{tot}}) - T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^p\right) dE^n + T^n\beta^n\mathcal{S}(\rho,\omega).\end{aligned}$$

First, we prove the liminf condition. Let $(\rho^n,\omega^n) \overset{*}{\rightharpoonup} (\rho,\omega)$. If (ρ,ω) does not satisfy the continuity equation, then only finitely many (ρ^n,ω^n) do satisfy it. To see this, let $\varphi \in C_c^1((0,T)\times\Omega)$ such that

$$\int_0^T \int_\Omega \partial_t\varphi d\rho_t dt + \int_0^T \int_\Omega \langle \nabla_x\varphi, d\omega_t\rangle dt \neq 0,$$

i.e. (ρ,ω) is not satisfying the continuity equation. Suppose infinitely many (ρ^n,ω^n) do satisfy the equation, then for those it is

$$\begin{aligned}0 &< \left|\int_0^T \int_\Omega \partial_t\varphi d\rho_t dt + \int_0^T \int_\Omega \langle \nabla_x\varphi, d\omega_t\rangle dt\right| \qquad (3.3)\\ &\le \left|\int_0^T \int_\Omega \partial_t\varphi d\rho_t dt - \int_0^T \int_\Omega \partial_t\varphi d\rho_t^n dt\right| + \left|\int_0^T \int_\Omega \langle \nabla_x\varphi, d\omega_t\rangle dt - \int_0^T \int_\Omega \langle \nabla_x\varphi, d\omega_t^n\rangle dt\right| \to 0\end{aligned}$$

for $n \to \infty$ due to the weak-* convergence of (ρ^n, ω^n). Since the above argument shows a contradiction, only finitely many (ρ^n, ω^n) satisfy the continuity equation and so $\mathcal{E}^\infty(\rho, \omega) \leq \liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n)$ is trivially fulfilled in the case where (ρ, ω) does not satisfy the continuity equation (because of $\mathcal{S}(\rho, \omega) = +\infty$ in this case).
Furthermore, if $\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) = +\infty$, the liminf condition is again trivially fulfilled and thus we can w.l.o.g. restrict ourselves to a subsequence of (ρ^n, ω^n) satisfying the continuity equation (since otherwise it would be $\mathcal{E}^n(\rho^n, \omega^n) = +\infty$ for those pairs not satisfying the continuity equation) and additionally assume $\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) < +\infty$.
In the proof of Lemma 3.2.8 we will show that

$$C\mathcal{S}(\rho^n, \omega^n) + f(0) \leq \mathcal{E}^n(\rho^n, \omega^n)$$

almost surely for n large enough, a constant $C > 0$ and an affine, strictly increasing function f. This implies $\mathcal{S}(\rho^{n_k}, \omega^{n_k}) \leq M$ for all n_k along the infimizing subsequence $(n_k)_k$. To ensure this uniform bound on the regularization part, we will restrict ourselves to the infimizing subsequence in the following (which will again be indexed by n) for each realization of the random variable $\mathcal{E}^n$. The convergence results of the stochastic part of the random variable will be independent of the extraction of a specific subsequence which means that we can choose a different subsequence for each realization of $\mathcal{E}^n$. Note, that we need to assume $\beta > 0$ in order to guarantee that the constant C is nonzero.

We start the actual proof of the liminf condition with the stochastic part of the functionals and consider the error term

$$\begin{aligned}
&\left|\sum_{kl} \int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right) \kappa_t^\dagger(\Gamma_{kl}) dt - \int T^n \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) dE^n\right| \\
=&\left|\sum_{kl}\left[\int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right) \kappa_t^\dagger(\Gamma_{kl}) dt - \int T^n \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n\right]\right| \\
\leq&\sum_{kl}\left|\int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right) \kappa_t^\dagger(\Gamma_{kl}) dt - T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n\right|.
\end{aligned}$$

We will show convergence for each detector pair *kl* separately. It is

$$\begin{aligned}
&\left|\int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right) \kappa_t^\dagger(\Gamma_{kl}) dt - T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n\right| \\
\leq&\left|\int_0^T \left(\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right) - \log\left(\kappa_t^p(\Gamma_{kl})\right)\right) \kappa_t^\dagger(\Gamma_{kl}) dt\right| \qquad (3.4)\\
&+\left|T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n - \int_0^T \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right) \kappa_t^\dagger(\Gamma_{kl}) dt\right|.
\end{aligned}$$

We would like to apply dominated convergence on the first part of (3.4). Thus, we show that $\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)$ is uniformly bounded. Using continuity of the κ functions in time and the mean value theorem, it is for $t \in \tau_i^n$ and with Lemma 3.1.1

$$\begin{aligned}\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right) &= \log\left(p\kappa^{n,s}_{t_i^n}(\Gamma_{kl}) + \kappa^{n,d}_{t_i^n}(\Gamma_{kl})\right)\\ &\le \log\left(C\int_0^T \rho_t^n(\Omega)dt\right).\end{aligned}$$

For the lower bound we get with mass conservation

$$\begin{aligned}\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right) &\ge \log\left(p\kappa^{n,s}_{t_i^n}(\Gamma_{kl})\right)\\ &= \log\left(pp^s\frac{\mathcal{H}^2(\Gamma_{kl})}{\mathcal{H}^2(\partial\Omega\times\partial\Omega)}\rho^n_{t_i^n}(\Omega)\right)\\ &= \log\left(\frac{pp^s}{T}\frac{\mathcal{H}^2(\Gamma_{kl})}{\mathcal{H}^2(\partial\Omega\times\partial\Omega)}\int_0^T\rho_t^n(\Omega)dt\right).\end{aligned}$$

Using the convergence $\int_0^T \rho_t^n(\Omega)dt \xrightarrow{n\to\infty} \int_0^T \rho_t(\Omega)dt > 0$ (unless $\rho^n \overset{*}{\rightharpoonup} 0$ which will be handled separately) we get $\left|\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\right| \lesssim 1$ for n large enough. This allows us to apply dominated convergence on the first term and thus this part converges to zero according to Lemma 3.2.4.
For the second term in (3.4) we will apply Markov's inequality. Therefore, we compute

$$\begin{aligned}&\mathbb{E}\left[T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n - \int_0^T\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\kappa_t^\dagger(\Gamma_{kl})dt\right]\\ =&T^n\int_0^T\int_{\partial\Omega\times\partial\Omega}\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\frac{1}{T^n}\kappa_t^\dagger(x)\mathbb{1}_{\Gamma_{kl}}(x)d\phi(x)dt\\ &\qquad - \int_0^T\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\kappa_t^\dagger(\Gamma_{kl})dt\\ =&0.\end{aligned}$$

By Markov's inequality and Lemma 2.2.5 we get for every $\varepsilon > 0$

$$\begin{aligned}&\mathbb{P}\left(\left|T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n - \int_0^T\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\kappa_t^\dagger(\Gamma_{kl})dt\right| > \varepsilon\right)\\ \le&\frac{1}{\varepsilon^{2m}}\mathbb{E}\left[\left(T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n - \mathbb{E}\left[T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n\right]\right)^{2m}\right]\\ =&\frac{(T^n)^{2m}}{\varepsilon^{2m}}\sum_{k\cdot s=2m,\ k_1=0}\frac{(2m)!}{k!(s!)^k}a^k\\ \le&\frac{(T^n)^m}{\varepsilon^{2m}}\sum_{k\cdot s=2m,\ k_1=0}\frac{(2m)!}{k!(s!)^k}(T^n a)^k\end{aligned}$$

with a_i defined as in Lemma 2.2.5 applied to the function $t \mapsto \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}$ and the intensity measure $\lambda = \frac{1}{T^n}\kappa^\dagger$. Keeping in mind that

$$\mathbb{E}\left[\left(T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n - \mathbb{E}\left[T^n\int\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}}dE^n\right]\right)^{2m}\right]$$

is of degree at most m in the a_i, the last inequality only holds for n large enough such that $T^n < 1$. As $\left|\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\right|$ is uniformly bounded in n by C, we can estimate

$$\begin{aligned}\left|(T^n a)^k\right| &\leq \prod_{i=1}^{\infty}\left(\int_0^T \left|\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\right|^i \kappa_t^\dagger(\Gamma_{kl})dt\right)^{k_i} \\ &\leq \max(1,C)^m \max\left(1, \int_0^T \kappa_t^\dagger(\Gamma_{kl})dt\right)^m \lesssim 1\end{aligned}$$

for all k satisfying $k \cdot s = 2m$, $k_1 = 0$ and $\sum_{i=1}^{\infty} k_i \leq m$. Thus, we have

$$\begin{aligned}&\mathbb{P}\left(\left|T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}} dE^n - \int_0^T \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\kappa_t^\dagger(\Gamma_{kl})dt\right| > \varepsilon\right) \\ &\lesssim (T^n)^m \xrightarrow{n\to\infty} 0\end{aligned}$$

which establishes convergence in probability. Using the assumed convergence of the series $\sum_n (T^n)^m < +\infty$, the above result implies almost sure convergence [14, Theorem 6.12]

$$T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)\mathbb{1}_{\Gamma_{kl}} dE^n - \int_0^T \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{[0,T]\times\Gamma_{kl}}\right)\kappa_t^\dagger(\Gamma_{kl})dt \xrightarrow{\text{a.s.}} 0$$

and hence almost sure convergence of the stochastic part

$$\int T^n \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) dE^n \xrightarrow{\text{a.s.}} \sum_{kl}\int_0^T \log\left(\kappa_t^p(\Gamma_{kl})\right)\kappa_t^\dagger(\Gamma_{kl})dt.$$

For the regularization part $\mathcal{S}$ we have, due to the lower semi-continuity of the Benamou-Brenier functional w.r.t. the weak-* convergence by Lemma 2.1.14,

$$\liminf_{n\to\infty} T^n \beta^n \mathcal{S}(\rho^n, \omega^n) \geq \beta \mathcal{S}(\rho, \omega). \tag{3.5}$$

From the convergence $\rho^n \overset{*}{\rightharpoonup} \rho$, the weak-* continuity of $\rho \mapsto A_{\Gamma_{kl}}^{s/d}\rho$ by Lemma 3.1.1 and the Portemanteau theorem we get

$$\kappa^n([0,T]\times\Gamma_{\mathtt{tot}}) = \sum_{kl}\int_0^T \kappa_t^n(\Gamma_{kl})dt \xrightarrow{n\to\infty} \sum_{kl}\int_0^T \kappa_t(\Gamma_{kl})dt = \kappa([0,T]\times\Gamma_{\mathtt{tot}}).$$

Thus, we have shown the liminf condition unless $\rho^n \overset{*}{\rightharpoonup} 0$. This case will be considered now. Setting $\mathcal{E}^\infty(0,0) = +\infty$, we need to show $\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) = +\infty$. The crucial part of this computation is the estimation of the log-part. With $N^n = \frac{1}{\Delta t^n} = \frac{1}{|\tau_i^n|}$ and with applying the mean value theorem to get $\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}|_{\tau_i^n\times\Gamma_{kl}} = \kappa^{n,p}_{t_i^n}(\Gamma_{kl})$, we deduce

$$\begin{aligned}
&- T^n \int_0^T \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n \\
=& - T^n \int_0^T \sum_{i=1}^{N^n} \mathbb{1}_{\tau_i^n} \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n \\
=& - T^n \sum_{i=1}^{N^n} \log\left(\frac{1}{\Delta t^n}\int_{\tau_i^n} \kappa_t^{n,p}(\Gamma_{kl})dt\right) E^n(\tau_i^n \times \Gamma_{kl}) \\
=& - T^n \sum_{i=1}^{N^n} \log\left(\kappa^{n,p}_{t_i^n}(\Gamma_{kl})\right) E^n(\tau_i^n \times \Gamma_{kl}) \\
\geq& - T^n \log\left(\sup_{t\in[0,T]} \kappa_t^{n,p}(\Gamma_{kl})\right) \sum_{i=1}^{N^n} E^n(\tau_i^n \times \Gamma_{kl}) \\
=& - T^n E^n([0,T]\times\Gamma_{kl}) \log\left(\sup_{t\in[0,T]} \kappa_t^{n,p}(\Gamma_{kl})\right). \qquad (3.6)
\end{aligned}$$

Because of

$$\mathbb{P}\left(\left|T^n E^n([0,T]\times\Gamma_{kl}) - \int_0^T \kappa_t^{\dagger}(\Gamma_{kl})dt\right| > \varepsilon\right) \leq \frac{\mathbb{E}\left[\left(T^n E^n([0,T]\times\Gamma_{kl}) - \int_0^T \kappa_t^{\dagger}(\Gamma_{kl})dt\right)^{2m}\right]}{\varepsilon^{2m}}$$

$$= \frac{(T^n)^{2m}}{\varepsilon^{2m}} \sum_{k\cdot s=2m,\ k_1=0} \frac{(2m)!}{k!(s!)^k}(a)^k \lesssim \frac{(T^n)^m}{\varepsilon^{2m}} \max(1, \int_0^T \kappa_t^{\dagger}(\Gamma_{kl})dt)^m \xrightarrow{n\to\infty} 0,$$

the first factor of (3.6) converges in probability to some positive constant. As above, this implies almost sure convergence because of $\sum_n (T^n)^m < +\infty$. For the second factor we use the uniform convergence $\kappa_t^{n,p} \to 0$ established by Lemma 3.2.2 to get

$$\liminf_{n\to\infty} -T^n \int_0^T \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n \gtrsim \liminf_{n\to\infty} -\log\left(\sup_{t\in[0,T]} \kappa_t^{n,p}(\Gamma_{kl})\right) = +\infty.$$

Thus,

$$\begin{aligned}
&\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) \\
=& \liminf_{n\to\infty} \left(\kappa^n([0,T]\times\Gamma_{\texttt{tot}}) - T^n \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) dE^n + T^n \beta^n \mathcal{S}(\rho^n, \omega^n)\right) \\
\geq& \liminf_{n\to\infty} -T^n \sum_{kl} \int \log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right) \mathbb{1}_{\Gamma_{kl}} dE^n = +\infty
\end{aligned}$$

almost surely for $\rho^n \overset{*}{\rightharpoonup} 0$. This finally proves the liminf condition almost surely.

For the limsup condition we can take $(\rho^n, \omega^n) = (\rho, \omega)$ for all n and do the same computations as above to get

$$\limsup_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) \leq \mathcal{E}^\infty(\rho, \omega)$$

almost surely. □

Remark 3.2.6. In Theorem 3.2.5 we assumed $\beta^n T^n \to \beta > 0$, i.e. that the regularization does not vanish in the Γ-limit. This is an important assumption as it ensures the uniform boundedness of the sequence $\mathcal{S}(\rho^n, \omega^n)$ which was a crucial condition in order to prove Lemma 3.2.4, i.e. $\frac{1}{\Delta t^n}\hat{\kappa}^{n,p} \to \kappa_t^p(\Gamma_{kl})$. This convergence was a key ingredient for showing the Γ-convergence.

Remark 3.2.7. Unlike for deterministic convergence, in the stochastic framework we do not have that a sequence converges to a limit l if and only if every subsequence has a further subsequence converging to l. This equivalence is replaced by the following [14, corollary 6.13]. A sequence of random variables $(X_n)_n$ converges to the random variable X in probability if and only if every subseqeunce of $(X_n)_n$ has a further subseqeunce converging almost surely to X. Since in general convergence in probability does not imply convergence almost surely (see [14, Remark 6.6]), this equivalence is not replacing the one we have for deterministic sequences. This means that we need the additional assumption $\sum_n (T^n)^m < +\infty$ in order to deduce almost sure convergence from convergence in probability of the stochastic part of the functionals $\mathcal{E}^n$.

Lemma 3.2.8. *Under the assumptions of Theorem 3.2.5 we have that the sequence of functions $\mathcal{E}^n$ is almost surely equi-mildly coercive, i.e. there exists a countably compact set $\mathbb{K} \subset \mathbb{M}$ such that*

$$\inf_{(\rho,\omega)\in\mathbb{M}} \mathcal{E}^n(\rho,\omega) = \inf_{(\rho,\omega)\in\mathbb{K}} \mathcal{E}^n(\rho,\omega)$$

for all $n \in \mathbb{N}$.

Proof. We show that the set of minimizers of the $\mathcal{E}^n$ is uniformly bounded and thus weakly-* precompact. Again, this needs to be done in a stochastic way and the proof will be given as follows. We denote the minimizers of $\mathcal{E}^n(\xi)$ by $(\rho^n(\xi), \omega^n(\xi))$ which exist for almost every $\xi \in X$, i.e. for almost every realization of the random variable $\mathcal{E}^n$. Since our sequence of functionals $(\mathcal{E}^n)_n$ is countable, there exists a set $A \subset X$ with probability one such that for every $\xi \in A$ the functionals $\mathcal{E}^n(\xi)$ have a minimizer $(\rho^n(\xi), \omega^n(\xi))$ for every n. Then, for every $\xi \in A$, $\mathbb{K}(\xi) = \overline{(\rho^n(\xi), \omega^n(\xi))_n}$ is the desired countably compact set.
First, we show that the functions $\mathcal{E}^n$ at the respective minimum are uniformly bounded. Indeed, for the Lebesgue measure $\mathcal{L}$ on $[0,T] \times \Omega$ it holds (see proof of existence, Theorem 3.1.2)

$$\min_{(\rho,\omega)\in\mathbb{M}} \mathcal{E}^n(\rho,\omega) \leq \mathcal{E}^n(\mathcal{L}, 0) < +\infty$$

and the last expression is almost surely uniformly bounded. To see this boundedness, consider

$$\mathcal{E}^n(\mathcal{L}, 0) = \kappa([0,T] \times \Gamma_{\texttt{tot}}) - T^n \int \log\left(\frac{1}{\Delta t^n}(p\hat{\kappa}^s + \hat{\kappa}^d)\right) dE^n.$$

We have to control the last part. As in the proof of Theorem 3.2.5 we have $\left|\log\left(\frac{1}{\Delta t^n}(p\hat{\kappa}^s + \hat{\kappa}^d)\right)\right| \lesssim 1$, thus

$$\left|T^n \int \log\left(\frac{1}{\Delta t^n}(p\hat{\kappa}^s + \hat{\kappa}^d)\right) dE^n\right| \lesssim T^n \int dE^n$$

and for any $\varepsilon > 0$

$$\mathbb{P}\left(\left|T^n \int dE^n - \int_0^T \kappa_t^\dagger(\partial\Omega \times \partial\Omega)dt\right| > \varepsilon\right) \leq \frac{(T^n)^m}{\varepsilon^{2m}} \sum_{k\cdot s=2m,\ k_1=0} \frac{(2m)!}{k!(s!)^k}(T^n a)^k \xrightarrow{n\to\infty} 0,$$

implying that $T^n \int dE^n \xrightarrow{n\to\infty} \int_0^T \kappa_t^\dagger(\partial\Omega \times \partial\Omega)dt$ almost surely with the condition $\sum_n (T^n)^m < +\infty$. This means that $T^n \int dE^n$ is almost surely bounded, which gives us almost surely a uniform bound on $\mathcal{E}^n(\mathcal{L}, 0)$. Next, we continue in a similar way as in the proof of Theorem 3.1.2 but need to take into account the dependence on n this time. The minimizers (ρ^n, ω^n) of $\mathcal{E}^n$ satisfy $\mathcal{S}(\rho^n, \omega^n) < +\infty$. Hence, by Lemma 3.1.1, we have that $t \mapsto A\rho_t(\Gamma_{kl})$ is continuous and $A\rho_t(\Gamma_{kl}) \lesssim \|\rho\|$. Using this we find with $\overline{p} = \max(1, p)$

$$\begin{aligned}
\mathcal{E}^n(\rho^n, \omega^n) &\geq T^n \sum_{i^n kl} \left[A_{ikl}^n \rho^n - E_{ikl}^n \log\left(\overline{p}\frac{T^n}{\Delta t^n} A_{ikl}^n \rho^n\right)\right] + T^n \beta^n \mathcal{S}(\rho^n, \omega^n) \\
&\geq T^n \sum_{i^n kl} \left[A_{ikl}^{n,s} \rho^n - E_{ikl}^n \log\left(\frac{\overline{p}}{\Delta t^n} \int_{\tau_i^n} A\rho_t^n(\Gamma_{kl})dt\right)\right] + T^n \beta^n \mathcal{S}(\rho^n, \omega^n) \\
&\geq T^n \sum_{i^n kl} \left[\frac{1}{T^n} \int_{\tau_i^n} \rho_t^n(\Omega)dt \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma_{kl})}{\mathcal{H}^2(\partial\Omega)^2} - E_{ikl}^n \log\left(\overline{p} A\rho_{t_i^n}^n(\Gamma_{kl})\right)\right] \\
&\qquad + T^n \beta^n \mathcal{S}(\rho^n, \omega^n) \\
&\geq \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma_{\mathtt{tot}})}{\mathcal{H}^2(\partial\Omega)^2} \|\rho^n\| - T^n \sum_{i^n kl} E_{ikl}^n \log(C\|\rho^n\|) + T^n \beta^n \mathcal{S}(\rho^n, \omega^n)
\end{aligned}$$

where the n in $A_{ikl}^n \rho^n$ stresses the fact that the half-lifes are now depending on n and the notation i^n means that the sum runs from one to $N^n = \frac{1}{\Delta t^n}$. Next, we write

$$T^n \sum_{i^n kl} E_{ikl}^n = T^n \int \mathbb{1}_{\Gamma_{\mathtt{tot}}} dE^n \leq T^n \int dE^n.$$

As in the proof of Theorem 3.2.5 we have $T^n \int dE^n \xrightarrow{n\to\infty} \int_0^T \kappa_t^\dagger(\partial\Omega \times \partial\Omega)dt$ almost surely, hence $(T^n \int dE^n)_n$ is almost surely bounded by a (probabilistic) constant C'. Moreover, $(T^n\beta^n)_n$ is bounded from below by $C'' > 0$ for n large enough because of $T^n\beta^n \xrightarrow{n\to\infty} \beta > 0$. With this we can further estimate for $C\|\rho^n\| > 1$ (for $C\|\rho^n\| \leq 1$ there is nothing to show)

$$\begin{aligned}
\mathcal{E}^n(\rho^n, \omega^n) &\geq \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma_{\mathtt{tot}})}{\mathcal{H}^2(\partial\Omega)^2} \|\rho^n\| - \left(T^n \int dE^n\right) \log(C\|\rho^n\|) + T^n \beta^n \mathcal{S}(\rho^n, \omega^n) \\
&\geq \frac{\mathcal{H}^2 \otimes \mathcal{H}^2(\Gamma_{\mathtt{tot}})}{\mathcal{H}^2(\partial\Omega)^2} \|\rho^n\| - C' \log(C\|\rho^n\|) + C'' \mathcal{S}(\rho^n, \omega^n) \\
&\geq f(\|\rho^n\|) + C'' \mathcal{S}(\rho^n, \omega^n)
\end{aligned}$$

for a strictly increasing affine function f, see proof of Theorem 3.1.2. Similar to this proof it now follows that the set of minimizers $\mathbb{K} := (\rho^n, \omega^n)_n$ is uniformly bounded, hence precompact in the weak-* topology. Because of $\inf_{\mathbb{M}} \mathcal{E}^n = \inf_{\mathbb{K}} \mathcal{E}^n$ for all n, the family $(\mathcal{E}^n)_n$ is equi-mildly coercive. □

Remark 3.2.9 (Convergence of minimizers of $\mathcal{E}^n$)**.** For an equi-mildy coercive sequence of functions that Γ-converge we have that cluster points of a minimizing sequence are minimizers of the limit functional. This is stated by Theorem 2.4.2. In our stochastic framework we can show that the functions $\mathcal{E}^n$ are almost surely equi-mildly coercive (see Lemma 3.2.8). Unfortunately we cannot use this to deduce that cluster points of a minimizing sequence of the $\mathcal{E}^n$ are minimizers of $\mathcal{E}^\infty$. The Problem is the following. We have shown separately for every sequence $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$ that the liminf inequality almost surely holds, i.e. for every $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$ we have

$$\mathbb{P}\Big(\xi \in X \mid \liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n, \xi) \geq \mathcal{E}^\infty(\rho, \omega, \xi)\Big) = 1.$$

To be able to apply Theorem 2.4.2 we would have needed to show

$$\mathbb{P}\Big(\xi \in X \mid \liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n, \xi) \geq \mathcal{E}^\infty(\rho, \omega, \xi) \text{ for every } (\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)\Big) = 1.$$

In our situation we can extract for almost every realization $\mathcal{E}^n(\zeta)$, $\zeta \in X$, of the sequence of random variables a precompact sequence of minimizers $(\rho^n(\zeta), \omega^n(\zeta))$ (see proof of Lemma 3.2.8) but we do not know if it holds

$$\liminf_{n\to\infty} \mathcal{E}^n(\rho^n(\zeta), \omega^n(\zeta), \zeta) \geq \mathcal{E}^\infty(\rho(\zeta), \omega(\zeta))$$

for this specific realization. We only know that

$$\liminf_{n\to\infty} \mathcal{E}^n(\rho^n(\zeta), \omega^n(\zeta), \xi) \geq \mathcal{E}^\infty(\rho(\zeta), \omega(\zeta))$$

holds for almost every $\xi \in X$ but we cannot tell from the proof of Theorem 3.2.5 for which ξ this inequality holds. This remains to be shown.

3.3 Γ-Convergence II: Variable Detector Sizes

In the sequence of functionals in Theorem 3.2.5 we only changed the size of the time intervals with increasing intensity of the radioactive material. One could also think about changing the detector sizes as well. This leads to a slightly different limit functional as in this case integrals with respect to time and space are involved. The arguments of the proof of Γ-convergence in this scenario are similar to the ones proving Theorem 3.2.5 but we need some additional assumptions on our mathematical model. These are going to be motivated now.

In the proof of Theorem 3.2.5 we used dominated convergence at some point. Therefore, we showed uniform boundedness of $\log\left(\frac{1}{\Delta t^n}\hat{\kappa}^{n,p}\right)$ (which was used later again when concluding almost sure convergence from Markov's inequality). Thus, to be able to follow along the steps of the proof of Theorem 3.2.5 we now need boundedness of $\log\left(\frac{1}{\Delta t^n |\Gamma^n_{kl}|}\hat{\kappa}^{n,p}\right)$. Parts of showing this can be adopted from the previous considerations but showing boundedness of the expression

$$\frac{R_\#(dx \otimes \mathrm{vol}_G)(\Gamma^n_{kl})}{|\Gamma^n_{kl}|} \tag{3.7}$$

that arises when estimating $\kappa^n_{t^n}(\Gamma^n_{kl})$ is more complicated. To be able to bound this expression, we need to make another assumption regarding the domain that the measurement takes place in. We assume that the detectors are located on the boundary of a convex and compact set Ω_δ and that the radioactive material is located in $\Omega \subsetneq \Omega_\delta$. The parameter δ quantifies the minimal distance between Ω and $\partial\Omega_\delta$, i.e. it is

$$\delta \leq \inf_{x\in\Omega,\ y\in\partial\Omega_\delta} \|x-y\|.$$

With this new configuration we must adjust our definition of the forward operator. The scattering part now reads

$$\begin{aligned} A^s : \mathcal{M}_+(\Omega) &\to \mathcal{M}_+(\partial\Omega_\delta \times \partial\Omega_\delta) \\ \lambda &\mapsto \frac{\lambda(\Omega)}{\mathcal{H}^2(\partial\Omega_\delta \times \partial\Omega_\delta)^2}\mathcal{H}^2 \otimes \mathcal{H} \llcorner (\partial\Omega_\delta \times \partial\Omega_\delta). \end{aligned}$$

We also need to modify the R-function from the definition of A^d. It is now given by

$$\begin{aligned} R_\delta : \Omega \times G^{1,3} &\to \partial\Omega_\delta \times \partial\Omega_\delta \\ (x,v) &\mapsto \partial\Omega_\delta \cap (x+v). \end{aligned}$$

This way of defining the function guarantees that the distance between two detectors being able to detect an unscattered photon pair is sufficiently large for small enough detector sizes. This condition will be needed to bound the expression (3.7) without making further assumptions on the regularity of $\partial\Omega$.

The second kind of assumptions are due to convergences of the kind

$$\lim_{(\tau^n,\Gamma^n)\to(t,x)} \frac{\int_{\tau^n} \kappa_t^p(\Gamma^n)dt}{|\tau^n|\,|\Gamma^n|} = \kappa_t^p(x) \tag{3.8}$$

for almost every $(t,x) \in [0,T] \times \partial\Omega_\delta \times \partial\Omega_\delta$. To guarantee existence of the limit (3.8), we need the collection of all detector pairs $(\Gamma_{kl}^n)_{kl,n}$ to be a ϕ vitali relation and the collection $(\tau_i^n)_{i,n} \times (\Gamma_{kl}^n)_{kl,n}$ to be a $dt \otimes d\phi$ vitali relation[20, Theorem 2.9.7]. This is true if we construct the sequence Γ_{kl}^n in such a way that each Γ_{kl}^{n+1} is contained in exactly one Γ_{kl}^n meaning that we receive the sequence by subdividing the existent detector pairs in each step. The same should hold for the time intervals τ_i^n, i.e. it must hold $\Delta t^{n+1} = \frac{1}{k^n}\Delta t^n$ for $k^n \in \mathbb{N}$. Additionally, we need the assumption $\cup_{kl}\Gamma_{kl}^n = \partial\Omega_\delta \times \partial\Omega_\delta$ in order to cover the whole surface with detectors such that the limit (3.8) can be computed almost everywhere on $\partial\Omega_\delta \times \partial\Omega_\delta$. Moreover, we need our detector sizes to converge to zero, we want the detectors to have a comparable size in each step n and we want the area of each detector to be proportional to its squared diameter. These demands on the shape of the detectors are stated in the following assumption.

Assumption 3.3.1. We assume that there exists a positive function f from $\mathbb{N}$ to $\mathbb{R}$ and a constant $c > 0$ satisfying

$$0 < cf(n) \leq \operatorname{diam}(\Gamma_k^n) \leq f(n)$$

for all $n \in \mathbb{N}$ and every detector k with $f(n) \xrightarrow{n\to\infty} 0$. Moreover, we assume that positive constants c' and c'' exist such that

$$c'\,|\Gamma_k^n| \leq \operatorname{diam}(\Gamma_k^n)^2 \leq c''\,|\Gamma_k^n|.$$

Lemma 3.3.2. *If we construct our sequence of detectors* $(\Gamma_{kl}^n)_{kl,n}$ *in the way described above and under Assumption 3.3.1, we receive a* ϕ *Vitali relation which is given by*

$$\{(x,S) \mid x \in S \in (\Gamma_{kl}^n)_{kl,n}\}$$

and for $\Delta t^{n+1} = \frac{1}{k^n}\Delta t^n$, $k^n \in \mathbb{N}$, *a* $\phi \otimes dt$ *Vitali relation*

$$\{(x,t,\tau,S) \mid (x,t) \in S \times \tau \in (\tau_i^n)_{i,n} \times (\Gamma_{kl}^n)_{kl,n}\}.$$

Proof. The proof follows from [20, Theorem 2.8.19]. □

Example 3.3.3 (Explicit setup)**.** In this example we explicitly give a setup meeting the above assumptions, i.e. a setup for which the following considerations of this section will hold.
Let Ω be a cube with edge length e and Ω_δ a cube with edge length $e_\delta = e + \delta$ for some $\delta > 0$. We subdivide each face of Ω_δ into n^2 squares (the detectors Γ_k^n) each of which has an edge length of e_δ/n. Then for every n it holds $\bigcup_{kl} \Gamma_{kl}^n = \partial\Omega_\delta \times \partial\Omega_\delta$ and for each detector Γ_k^n we have

$$\operatorname{diam}(\Gamma_k^n)^2 = \left(\frac{\sqrt{2}e_\delta}{n}\right)^2 = 2\,|\Gamma_k^n|\,.$$

One crucial point in proving Γ-convergence was Lemma 3.2.4. We now expand this result and include a decreasing family of detector pairs.

Lemma 3.3.4. *Let* $(\rho^n, \omega^n) \xrightharpoonup{*} (\rho, \omega)$, *let* $(\Gamma_{kl}^n)_{kl,n}$ *be a* ϕ *Vitali relation,* $(\tau_i^n)_{i,n} \times (\Gamma_{kl}^n)_{kl,n}$ *a* $dt \otimes \phi$ *Vitali relation, let* $(\tau_t^n)_n \subset (\tau_i^n)_{i,n}$ *be a sequence of intervals with* $\tau_t^n \to t$ *and let* $(\Gamma_x^n)_n \subset (\Gamma_{kl}^n)_{kl,n}$ *such that* $\Gamma_x^n \to x$. *Moreover, we assume the uniform boundedness* $\mathcal{S}(\rho^n,\omega^n) \le M$ *and* $\mathcal{S}(\rho,\omega) < +\infty$. *Then we have for* $dt \otimes \phi$ *almost every* (t,x) *meeting the above conditions*

$$\lim_{n\to\infty} \frac{1}{|\tau^n|\phi(\Gamma_x^n)} \int_{\tau^n} \kappa_t^{n,p}(\Gamma_x^n)dt \to \kappa_t^p(x) \tag{3.9}$$

where $\kappa_t^p(x)$ *is the Radon-Nikodym derivative of the measure defined by*

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^p(\Gamma)dt$$

with respect to the measure $dt \otimes \phi$.
The convergence (3.9) *holds for the scattering and detection part of* κ *separately.*

Proof. We start with showing (3.9) for the scattering part of κ. It is

$$\begin{aligned}\frac{1}{|\tau^n|\,\phi(\Gamma_x^n)} \int_{\tau^n} \kappa_t^{n,s}(\Gamma_x^n)dt &= \frac{1}{|\tau^n|\,\phi(\Gamma_x^n)} \int_{\tau^n} \kappa_t^s(\Gamma_x^n) + [\kappa_t^{n,s}(\Gamma_x^n) - \kappa_t^s(\Gamma_x^n)]\,dt \\ =&\frac{1}{|\tau^n|\,\phi(\Gamma_x^n)} \frac{p^s\phi(\Gamma_x^n)}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)} \int_{\tau^n} \rho_t(\Omega)dt + \frac{1}{|\tau^n|\,\phi(\Gamma_x^n)} \frac{p^s\phi(\Gamma_x^n)}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)} \int_{\tau^n} \rho_t^n(\Omega) - \rho_t(\Omega)dt.\end{aligned}$$

By mass conservation the first part is the Radon-Nikodym derivative of the measure

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^s(\Gamma)dt$$

w.r.t. $dt \otimes d\phi$ which is given by

$$(t,x) \mapsto \frac{p^s}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\rho_t(\Omega).$$

Using the mean value theorem, mass conservation and weak-* convergence, the second part converges to zero. We find

$$\begin{aligned}&\frac{1}{|\tau^n|\,\phi(\Gamma_x^n)}\frac{p^s\phi(\Gamma_x^n)}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\int_{\tau^n}\rho_t^n(\Omega) - \rho_t(\Omega)dt \\ =&\frac{p^s}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\frac{1}{T}\int_0^T \rho_t^n(\Omega) - \rho_t(\Omega)dt \xrightarrow{n\to\infty} 0.\end{aligned}$$

Next, we prove the result for the detection part. Just like in the proof of Lemma 3.2.4, it suffices to show the result for a subsequence.
We write

$$\frac{1}{|\tau^n|\,\phi(\Gamma_x^n)}\int_{\tau^n}\kappa_t^{n,d}(\Gamma_x^n)dt = \frac{1}{|\tau^n|\,\phi(\Gamma_x^n)}\int_{\tau^n}\kappa_t^{d}(\Gamma_x^n)dt + \frac{1}{|\tau^n|\,\phi(\Gamma_x^n)}\int_{\tau^n}\kappa_t^{n,d}(\Gamma_x^n) - \kappa_t^{d}(\Gamma_x^n)dt.$$

The first term converges to the Radon-Nikodym derivative of the measure

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^d(\Gamma)dt$$

with respect to the measure $dt \otimes \phi$ [20, Theorem 2.9.7]. The limit exists for $dt \otimes \phi$ almost all $(t,x) \in [0,T] \times (\partial\Omega_\delta \times \partial\Omega_\delta)$. Thus, we are left to show that the second term goes to zero. We find

$$\begin{aligned}&\left|\frac{1}{|\tau^n|\,\phi(\Gamma_x^n)}\int_{\tau^n}\kappa_t^{n,d}(\Gamma_x^n) - \kappa_t^{d}(\Gamma_x^n)dt\right| \\ =&\left|\frac{p^d}{|\tau^n|\,\phi(\Gamma_x^n)}\int_{\tau^n}\int_{G^{1,3}}\int_{Z_v^{\Gamma_x^n}}\left(\int_\Omega G_y(x)d\rho_t^n(y) - G_y(x)d\rho_t(y)\right)dx\mathrm{vol}_{G^{1,3}}(dv)dt\right| \\ \lesssim&\frac{R_{\delta\#}(dx\otimes\mathrm{vol}_{G^{1,3}})(\Gamma_x^n)}{\phi(\Gamma_x^n)}\sup_x\left|\frac{1}{|\tau^n|}\int_{\tau^n}\left(\int_\Omega G_y(x)d\rho_t^n(y) - G_y(x)d\rho_t(y)\right)dt\right| \\ \leq&\frac{R_{\delta\#}(dx\otimes\mathrm{vol}_{G^{1,3}})(\Gamma_x^n)}{\phi(\Gamma_x^n)}\sup_{x,s}\left|\left(\int_\Omega G_y(x)d\rho_s^n(y) - G_y(x)d\rho_s(y)\right)\right|.\end{aligned}$$

The first part $\frac{R_{\delta\#}(dx\otimes\mathrm{vol}_{G^{1,3}})(\Gamma_x^n)}{\phi(\Gamma_x^n)}$ converges to the Radon-Nikodym derivative of $R_{\delta\#}(dx \otimes \mathrm{vol}_{G^{1,3}})$ with respect to ϕ for ϕ almost all $x \in \partial\Omega_\delta \times \partial\Omega_\delta$. The second part converges to zero which will be shown in the following. To this end we prove uniform convergence of

$$(s,x) \mapsto \int_\Omega G_y(x)d\rho_s^n(y)$$

and finally show that the limit coincides with $\int_\Omega G_y(x)d\rho_s(y)$.
We want to apply the Arzelà-Ascoli theorem to establish uniform convergence. Therefore,

we need a uniform bound on and equicontinuity of the family of functions $\left((s,x) \mapsto \int_\Omega G_y(x) d\rho_s^n(y)\right)_n$. It is

$$\left| \int_\Omega G_y(x) d\rho_s^n(y) \right| \lesssim \rho_s^n(\Omega) \lesssim 1,$$

establishing that the family of functions is uniformly bounded. Further, we find

$$\begin{aligned} &\left| \int_\Omega G_y(x) d\rho_s^n(y) - \int_\Omega G_y(\tilde{x}) d\rho_{\tilde{s}}^n(y) \right| \\ \leq &\underbrace{\left| \int_\Omega (G_y(x) - G_y)(\tilde{x}) d\rho_s^n(y) \right|}_{I} + \underbrace{\left| \int_\Omega G_y(\tilde{x})(d\rho_s^n(y) - d\rho_{\tilde{s}}^n(y)) \right|}_{II}. \end{aligned}$$

We will now show $I \lesssim |x - \tilde{x}|$ and $II \lesssim |s - \tilde{s}|^{1/2}$ which gives us the equicontinuity. For I, consider the function

$$x \mapsto \int_\Omega G_y(x) d\rho_s^n(y)$$

for fixed $s \in [0,T]$. From $\rho^n \overset{*}{\rightharpoonup} \rho$ and mass conservation we get $\rho_s^n(\Omega) \lesssim \int_0^T \rho_t^n(\Omega) dt \lesssim 1$. It then holds

$$\begin{aligned} \left| \int_\Omega G_y(x) d\rho_s^n(y) - \int_\Omega G_y(\tilde{x}) d\rho_s^n(y) \right| &\leq \sup_{y \in \Omega} |G_y(x) - G_y(\tilde{x})| \, \rho_s^n(\Omega) \\ &\lesssim \sup_{y \in \Omega} |G_y(x) - G_y(\tilde{x})| \,. \end{aligned}$$

The set Ω is compact and it is $G \in C^\infty(\Omega^2)$. Then, using the mean value theorem, we have (for some $c \in (0,1)$)

$$|G_y(x) - G_y(\tilde{x})| = |\nabla_x G_y((1-c)\tilde{x} - cx)(x - \tilde{x})| \lesssim |x - \tilde{x}| \,,$$

i.e. $I \lesssim |x - \tilde{x}|$. The estimate of II is derived similar to the estimation in the proof of Lemma 3.2.2. The weak derivative of

$$s \mapsto \int_\Omega G_y(x) d\rho_s^n(y)$$

reads

$$s \mapsto \int_\Omega \langle \nabla_y G_y(x), d\omega_s^n(y) \rangle$$

and the function is continuous which can be seen from $\rho_{t^k}^n \overset{*}{\rightharpoonup} \rho_s^n$ for $t^k \to t$ (proof of Lemma 3.1.1 part (d)). Since $G_y \in C(\Omega)$ we get $\int_\Omega G_y(x) d\rho_{t^k}^n(y) \to \int_\Omega G_y(x) d\rho_t^n(y)$. With this we can derive

$$\begin{aligned} \left| \int_\Omega G_y(\tilde{x}) d\rho_s^n(y) - \int_\Omega G_y(\tilde{x}) d\rho_{\tilde{s}}^n(y) \right| &\leq \|\nabla_y G_y(\tilde{x})\|_\infty \int_s^{\tilde{s}} \int_\Omega \left| \frac{d\omega_t^n}{d\rho_t^n} \right| d\rho_t^n dt \\ &\lesssim |s - \tilde{s}|^{1/2} \int_0^T \left\| \frac{d\omega_t^n}{d\rho_t^n} \right\|_{L^2(\rho_t^n)} dt \\ &\lesssim |s - \tilde{s}|^{1/2} \,. \end{aligned}$$

As we have shown uniform boundedness and equicontinuity of

$$(s,x) \mapsto \int_\Omega G_y(x) d\rho_s^n(y) =: f^n(x,t),$$

the Arzelà-Ascoli theorem now shows that there exists a subsequence (still indexed by n) such that the above function f converges uniformly. We need to verify that this uniform limit f^∞ coincides with $\int_\Omega G_y(x) d\rho_t(y)$. Using dominated convergence and weak-* convergence we find for every $x \in \Omega$ and every $\psi \in C([0,T])$

$$\begin{aligned}\int_0^T \psi(t) f^\infty(x,t) dt = \lim_{n\to\infty} \int_0^T \psi(t) f^n(x,t) dt = \lim_{n\to\infty} \int_0^T \int_\Omega \psi(t) G_y(x) d\rho_t^n(y) dt \\ = \int_0^T \int_\Omega \psi(t) G_y(x) d\rho_t dt.\end{aligned}$$

Thus, $f^\infty(x,t) = \int_\Omega G_y(x) d\rho_t$ for all x and dt almost all t and continuity yields equality for all t. □

Lemma 3.3.5. *It is*

$$\frac{R_{\delta_\#}(dx \otimes vol_{G^{1,3}})(\Gamma_{kl}^n)}{|\Gamma_{kl}^n|} \lesssim 1$$

in the given setting of this section.

Proof. To establish this estimate, we will show

$$R_{\delta_\#}(dx \otimes \mathrm{vol}_{G^{1,3}})(\Gamma_{kl}^n) \lesssim |\Gamma_{kl}^n|\,.$$

Therefore, we need to quantify "how many points" and "how many directions" possibly contribute to the detections in the given detector pair Γ_{kl}^n.
We start with quantifying the "number of directions". Note, that we can assume the two considered detectors to have a minimal distance of δ from each other due to the new definitions of the domains and forward operator. Since our detector sizes converge to zero and by Assumption 3.3.1, their maximal size is smaller than $\delta/2$ for all n larger than some $n_0 \in \mathbb{N}$. Due to the assumption

$$\delta \leq \inf_{x\in\Omega,\ y\in\partial\Omega_\delta} |x-y|$$

on the domains, the distance between two detection points on $\partial\Omega_\delta \times \partial\Omega_\delta$ is at least 2δ. Thus, for all $n \geq n_0$ we have that the distance between two detectors contributing to measurements is at least δ.

We are now ready to estimate the number of lines (actually the number of directions of those lines) connecting the two detectors. To do so, it suffices to consider two detectors shaped as circles, each with a radius of $R_{kl}^n := \max(\mathrm{diam}(\overline{\Gamma_k^n}), \mathrm{diam}(\overline{\Gamma_l^n}))$, that are directed opposite to each other at a distance of $d_{kl}^n := \min_{x\in\overline{\Gamma_k^n}, y\in\overline{\Gamma_l^n}} |x-y| \geq \delta$. To see this, we choose two points

$$p_k^n \in \operatorname*{argmin}_{x\in\overline{\Gamma_k^n}} \min_{y\in\overline{\Gamma_l^n}} |x-y| \quad \text{and} \quad p_l^n \in \operatorname*{argmin}_{y\in\overline{\Gamma_l^n}} \min_{x\in\overline{\Gamma_k^n}} |x-y|$$

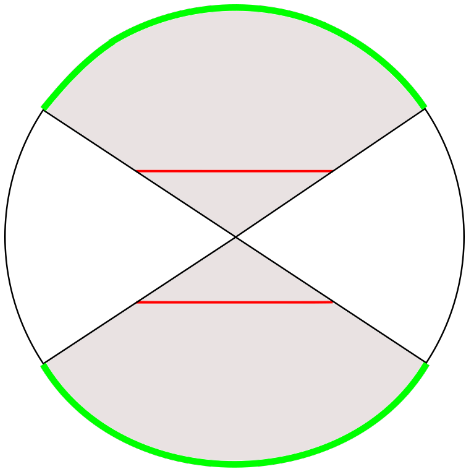

Figure 3.1: Two dimensional depiction of how to measure the directions of lines connecting two circles (the circles are lines in 2d). The two circles (red lines) are projected onto the sphere and the green part of the sphere measures the number of directions of lines connecting the two circles.

and consider the planes $P^n_{k/l}$ having $p^n_k - p^n_l$ as normal vector and $p^n_{k/l}$ as support vectors. We then consider the two circles $A^n_{k/l}$ with centers $p^n_{k/l}$ and radius R^n_{kl} located in the respective plane $P^n_{k/l}$. Due to the choice of the size of the circles we have that the orthogonal projection of $\Gamma^n_{k/l}$ onto the planes $P^n_{k/l}$ results in the circles $A^n_{k/l}$. Since in addition the faces of both circles are pointed to one another, any vector (i.e. direction) $(a-b)$ with $a \in \Gamma^n_k$ and $b \in \Gamma^n_l$ passes through both circles $A^n_{k/l}$. Therefore, we have the estimate

$$\begin{aligned} &\mathrm{vol}_{G^{1,3}}(v \in G^{1,3} \mid (x+v) \cap \Gamma^n_k \neq \emptyset \text{ and } (x+v) \cap \Gamma^n_l \neq \emptyset \text{ for some } x \in \Omega) \\ \leq &\mathrm{vol}_{G^{1,3}}(\underbrace{v \in G^{1,3} \mid (x+v) \cap A^n_k \neq \emptyset \text{ and } (x+v) \cap A^n_l \neq \emptyset \text{ for some } x \in \Omega}_{:=V^n_{kl}}), \end{aligned}$$

proving that it suffices to consider two circles faced to one another instead of the actual detectors. Further we can assume w.l.o.g. that the two circles are oriented symmetrically around the origin.

Next, we quantify the number of lines (actually the number of their directions) connecting the two circles $A^n_{k/l}$. The measure $\mathrm{vol}_{G^{1,3}}$ on $G^{1,3}$ can be seen as a probability measure on S^2 that measures the number of lines through the origin piercing the unit sphere S^2 [16, Section 3.9]. The most extreme directions that we are interested in are lines through the origin connecting two points from the boundaries of $A^n_{k/l}$ (the other directions can be derived from those). These directions form a cone. This means that we are interested in the radial projection of the $A^n_{k/l}$ onto S^2 in order to measure the number of lines that connect the two circles. The situation is depicted in Figure 3.1 for the two-dimensional case. The two circles (red lines) are projected onto the sphere and the green part of the sphere measures the number of directions of lines connecting the two circles. As both circles are of the same shape, one detector can (after possibly reorientating the configuration) be parametrized as

$$\begin{aligned}\Big\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le (R^n_{kl})^2\Big\} &\to \mathbb{R}^3\\ (x,y) &\mapsto (x,y,e^n_{kl})\end{aligned}$$

with a distance of $e^n_{kl} = d^n_{kl}/2$ from the origin. It suffices to consider this detector only because of the symmetry of the configuration. The projection onto S^2 then reads

$$\begin{aligned}\Phi^n_{kl}\colon \Big\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le (R^n_{kl})^2\Big\} &\to \mathbb{R}^3\\ (x,y) &\mapsto \frac{(x,y,e^n_{kl})}{\sqrt{x^2+y^2+(e^n_{kl})^2}}\end{aligned}$$

with Jacobian

$$(x^2+y^2+(e^n_{kl})^2)D\Phi^n_{kl} = \begin{bmatrix} \sqrt{x^2+y^2+(e^n_{kl})^2}-\frac{x^2}{\sqrt{x^2+y^2+(e^n_{kl})^2}} & -\frac{xy}{\sqrt{x^2+y^2+(e^n_{kl})^2}}\\ -\frac{xy}{\sqrt{x^2+y^2+(e^n_{kl})^2}} & \sqrt{x^2+y^2+(e^n_{kl})^2}-\frac{y^2}{\sqrt{x^2+y^2+(e^n_{kl})^2}}\\ -\frac{xe^n_{kl}}{\sqrt{x^2+y^2+(e^n_{kl})^2}} & -\frac{ye^n_{kl}}{\sqrt{x^2+y^2+(e^n_{kl})^2}}\end{bmatrix}.$$

We then have

$$\begin{aligned}\det\Big((D\Phi^n_{kl})^T D\Phi^n_{kl}\Big) &= \frac{1}{\Big(x^2+y^2+(e^n_{kl})^2\Big)^2}\left|\begin{bmatrix} y^2+(e^n_{kl})^2 & -xy\\ -xy & x^2+(e^n_{kl})^2\end{bmatrix}\right|\\ &= \frac{y^2x^2-x^2y^2}{\Big(x^2+y^2+(e^n_{kl})^2\Big)^2}+\frac{(e^n_{kl})^2}{x^2+y^2+(e^n_{kl})^2}\le 1\end{aligned}$$

and

$$\mathrm{vol}_{G^{1,3}}(V^n_{kl}) = 2\frac{1}{4\pi}\int_{\Big\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le (R^n_{kl})^2\Big\}}\sqrt{\det\left((D\Phi^n_{kl})^T(x,y)D\Phi^n_{kl}(x,y)\right)}d(x,y)$$

where the factor two comes from the fact that we have two circles. This integral can be estimated as

$$\begin{aligned}\mathrm{vol}_{G^{1,3}}(V^n_{kl}) &\le \max_{(x,y)\in\Big\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le (R^n_{kl})^2\Big\}}\left(\sqrt{\det\left((D\Phi^n_{kl})^T(x,y)D\Phi^n_{kl}(x,y)\right)}\right)\pi\frac{1}{2\pi}(R^n_{kl})^2\\ &\lesssim (R^n_{kl})^2\,.\end{aligned}$$

We now estimate the number of points possibly contribute to a detection at detector pair Γ^n_{kl}. These points are described by

$$N^n_{kl} = \Big\{x\in\Omega \mid \exists v\in G^{1,3} \text{ with } (x+v)\cap\Gamma^n_k\ne\emptyset \text{ and } (x+v)\cap\Gamma^n_l\ne\emptyset\Big\}.$$

The largest possible number of points in Ω that can lead to a detection of a photon pair at Γ_{kl}^n is less than or equal to $\pi \min((R_k^n)^2, (R_l^n)^2) \sup_{x\in\Gamma_k^n, y\in\Gamma_l^n} \|x-y\|$, where $R_{k/l}^n$ is the diameter of the respective detector. We can further estimate

$$\sup_{x\in\Gamma_k^n, y\in\Gamma_l^n} |x-y| \leq \operatorname{diam}(\Omega)$$

which gives us

$$|N_{kl}^n| \leq \pi \operatorname{diam}(\Omega) \min((R_k^n)^2, (R_l^n)^2) \lesssim \min(|\Gamma_k^n|, |\Gamma_l^n|).$$

In total, we have

$$\begin{aligned}\frac{R_{\delta_\#}(dx \otimes \operatorname{vol}_{G^{1,3}})(\Gamma_{kl}^n)}{|\Gamma_{kl}^n|} &\lesssim \frac{\min(|\Gamma_k^n|, |\Gamma_l^n|) \max((R_k^n)^2, (R_l^n)^2)}{|\Gamma_k^n|\,|\Gamma_l^n|}\\ &\lesssim \frac{\min(|\Gamma_k^n|, |\Gamma_l^n|) \max(|\Gamma_k^n|, |\Gamma_l^n|)}{|\Gamma_k^n|\,|\Gamma_l^n|} = 1,\end{aligned}$$

showing the desired estimate. □

Next, we prove Γ-convergence for temporal resolutions Δt^n and detector pairs Γ_{kl}^n converging to zero while the intensity of radiation of the radioactive material tends to infinity. The PPP E^n producing the measurement for a given half-life T^n (with $T^n \to 0$) has the intensity measure $\frac{1}{T^n}\kappa_t^\dagger(x) dt d\phi$ where $\kappa^\dagger$ is the Radon-Nikodym derivative of the measure

$$\mathcal{B}([0,T]) \times \mathcal{B}(\partial\Omega_\delta \times \partial\Omega_\delta) \to [0,+\infty], \quad \tau \times \Gamma \mapsto \int_\tau \kappa_t^\dagger(\Gamma) dt$$

w.r.t. $dt \otimes d\phi$.

Theorem 3.3.6 (Γ-convergence general case)**.** *Let $(T^n)_n$ be a sequence of half-lifes with $\sum_n (T^n)^m < +\infty$ for some $m \in \mathbb{N}$, $T^n\beta^n \to \beta$, let E^n be a PPP with intensity measure $\frac{1}{T^n}\kappa_t(x)^\dagger dt d\phi$ and $\Delta t^n \to 0$ for a decreasing sequence of temporal resolutions $(\Delta t^n)_n$. Moreover, we assume $(\Gamma_{kl}^n)_{kl,n}$ to be a ϕ Vitali relation such that $\bigcup_{kl} \Gamma_{kl}^n = \partial\Omega_\delta \times \partial\Omega_\delta$ for all n, $(\tau_i^n)_{i,n} \times (\Gamma_{kl}^n)_{kl,n}$ to be a $dt \otimes \phi$ Vitali relation such that $\cup_{ikl} \tau_i^n \times \Gamma_{kl}^n = [0,T] \times \partial\Omega_\delta \times \partial\Omega_\delta$ for all n and Assumption 3.3.1 is satisfied. We set*

$$C^n = T^n \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \log\left(\frac{\Delta t^n\,|\Gamma^n|}{T^n}\right) dE^n.$$

With respect to the weak- convergence on $\mathbb{M}$ we then have*

$$\Gamma - \lim_{n\to\infty} T^n J_{\beta^n, T^n}^{E^n, \Delta t^n, (\Gamma_{kl}^n)_{kl}} + C^n = \mathcal{E}^\infty$$

almost surely with the limit functional

$$\mathcal{E}^\infty(\rho, \omega) = \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \left(\kappa_t(x) - \log\left(\kappa_t^p(x)\right) \kappa_t^\dagger(x)\right) dt d\phi(x) + \beta \mathcal{S}(\rho, \omega).$$

Proof. We define

$$\begin{aligned}\mathcal{E}^n(\rho,\omega) &= T^n J^{E,\Delta t^n,(\Gamma^n_{kl})_{kl}}_{\beta^n,T^n}(\rho,\omega) + C^n \\ &= \kappa([0,T]\times\partial\Omega_\delta\times\partial\Omega_\delta) - T^n \int \log\left(\frac{1}{\Delta t^n\,|\Gamma^n|}\hat{\kappa}^p\right) dE^n + T^n\beta^n\mathcal{S}(\rho,\omega),\end{aligned}$$

where $|\Gamma^n| = \sum_{(kl)^n} \mathbb{1}_{\Gamma^n_{kl}}\phi(\Gamma^n_{kl})$ is a function that is constant on every detector pair Γ^n_{kl} with value $\phi(\Gamma^n_{kl})$.

We start with the liminf condition. Let $(\rho^n,\omega^n) \overset{*}{\rightharpoonup} (\rho,\omega)$. As in the proof of Theorem 3.2.5 (and by considerations in the proof of the following Lemma 3.3.7), we can assume w.l.o.g. that $\liminf_{n\to\infty}\mathcal{E}^n(\rho^n,\omega^n) < +\infty$ and $\mathcal{S}(\rho^n,\omega^n) \leq M$ for every n as well as $\mathcal{S}(\rho,\omega) < +\infty$. For the regularization part we have due to the lower semi-continuity of the Benamou-Brenier functional w.r.t. the weak-* convergence established by Lemma 2.1.14

$$\liminf_{n\to\infty} T^n\beta^n\mathcal{S}(\rho^n,\omega^n) \geq \beta\mathcal{S}(\rho,\omega).$$

Additionally, weak-* convergence $\rho^n \overset{*}{\rightharpoonup} \rho$ and continuity of the forward operator w.r.t. this convergence yields

$$\begin{aligned}\kappa^n([0,T]\times\partial\Omega_\delta\times\partial\Omega_\delta) &= \int_0^T \kappa^n_t(\partial\Omega_\delta\times\partial\Omega_\delta)dt \\ \xrightarrow{n\to\infty} \int_0^T \kappa_t(\partial\Omega_\delta\times\partial\Omega_\delta)dt &= \int_0^T\int_{\partial\Omega_\delta\times\partial\Omega_\delta}\kappa_t(x)d\phi(x)dt.\end{aligned}$$

Thus, we are left to prove the liminf condition for the log part. We consider the error term

$$Z^n = \left|\int_{\partial\Omega_\delta\times\partial\Omega_\delta}\int_0^T \log\left(\kappa^p_t(x)\right)\kappa^\dagger_t(x)dtd\phi(x) - \int T^n\log\left(\frac{1}{\Delta t^n\,|\Gamma^n|}\hat{\kappa}^{n,p}\right)dE^n\right|$$

with E^n being a PPP on $[0,T]\times(\partial\Omega_\delta\times\partial\Omega_\delta)$ with intensity measure $\frac{1}{T^n}\kappa^\dagger_t(x)dtd\phi(x)$. We get

$$\begin{aligned}Z^n \leq &\left|\int_{\partial\Omega_\delta\times\partial\Omega_\delta}\int_0^T\left[\log\left(\frac{1}{\Delta t^n\,|\Gamma^n|}\hat{\kappa}^{n,p}\right) - \log\left(\kappa^p_t(x)\right)\right]\kappa^\dagger_t(x)dtd\phi(x)\right| \qquad (3.10)\\ &+\left|T^n\int\log\left(\frac{1}{\Delta t^n\,|\Gamma^n|}\hat{\kappa}^{n,p}\right)dE^n - \int_{\partial\Omega_\delta\times\partial\Omega_\delta}\int_0^T\log\left(\frac{1}{\Delta t^n\,|\Gamma^n|}\hat{\kappa}^{n,p}\right)\kappa^\dagger_t(x)dtd\phi(x)\right|.\end{aligned}$$

We would like to use dominated convergence on the first part of (3.10). Thus, we show that $\log\left(\frac{1}{\Delta t^n|\Gamma^n|}\hat{\kappa}^{n,p}\right)$ is uniformly bounded in n. Using mass conservation, Lemma 3.3.5

and weak-* convergence, it is for $(t,x) \in \tau_i^n \times \Gamma_{kl}^n$

$$\begin{aligned}
\log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right)(t,x) &= \log\left(\frac{1}{\Delta t^n \phi(\Gamma_{kl}^n)}\left(p\int_{\tau_i^n}\kappa_t^{n,s}(\Gamma_{kl}^n)dt + \int_{\tau_i^n}\kappa_t^{n,d}(\Gamma_{kl}^n)dt\right)\right)\\
&\le \log\left(\frac{pp^s}{T\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\int_0^T \rho_t^n(\Omega)dt\right.\\
&\qquad\left. + \frac{p^d}{\Delta t^n \phi(\Gamma_{kl}^n)}\int_{\tau_i^n}\int_{G^{1,3}}\int_{Z_v^{\Gamma_{kl}^n}}\int_\Omega G_y(x)\rho_t^n(dy)dx\mathrm{vol}_{G^{1,3}}(dv)dt\right)\\
&\le \log\left(\frac{pp^s}{T\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\int_0^T \rho_t^n(\Omega)dt + p^d\frac{R_{\delta\#}(dx\otimes \mathrm{vol}_G)(\Gamma_{kl}^n)}{\phi(\Gamma_{kl}^n)}\frac{\|G_y(x)\|_\infty}{T}\int_0^T \rho_t^n(\Omega)dt\right)\\
&\le \log\left(C\int_0^T \rho_t^n(\Omega)dt\right).
\end{aligned}$$

For the lower bound we get

$$\begin{aligned}
\log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right)(t,x) &\ge \log\left(\frac{p}{\Delta t^n\phi(\Gamma_{kl}^n)}\int_{\tau_i^n}\kappa_t^{n,s}(\Gamma_{kl}^n)\right)\\
&= \log\left(\frac{pp^s}{T\phi(\partial\Omega_\delta \times \partial\Omega_\delta)}\int_0^T \rho_t^n(\Omega)dt\right).
\end{aligned}$$

Using the convergence $\int_0^T \rho_t^n(\Omega)dt \xrightarrow{n\to\infty} \int_0^T \rho_t(\Omega)dt > 0$ (unless $\rho^n \overset{*}{\rightharpoonup} 0$ which will be handled separately), we get with the above estimates $\left|\log\left(\frac{1}{\Delta t^n |\Gamma_{kl}^n|}\hat{\kappa}^{n,p}\right)\right| \lesssim 1$. This allows us to apply dominated convergence on the first term of (3.10) and thus this part converges to zero according to Lemma 3.3.4.
For the second term in (3.10) we will apply Markov's inequality to show convergence to zero. We find

$$\begin{aligned}
&\mathbb{E}\left[T^n\int\log\left(\frac{1}{\Delta t^n|\Gamma^n|}\hat{\kappa}^{n,p}\right)dE^n - \int_{\partial\Omega_\delta\times\partial\Omega_\delta}\int_0^T\log\left(\frac{1}{\Delta t^n|\Gamma^n|}\hat{\kappa}^{n,p}\right)\kappa_t^\dagger(x)dtd\phi(x)\right]\\
=&T^n\int_0^T\log\left(\frac{1}{\Delta t^n|\Gamma^n|}\hat{\kappa}^{n,p}\right)\frac{1}{T^n}\kappa_t^\dagger(x)dtd\phi(x)\\
&\qquad - \int_{\partial\Omega_\delta\times\partial\Omega_\delta}\int_0^T\log\left(\frac{1}{\Delta t^n|\Gamma^n|}\hat{\kappa}^{n,p}\right)\kappa_t^\dagger(x)dtd\phi(x)\\
=&0.
\end{aligned}$$

By Markov's inequality and Lemma 2.2.5 we get for every $\varepsilon > 0$

$$\begin{aligned}
&\mathbb{P}\left(\left|T^n \int \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n - \int_0^T \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) \kappa_t^\dagger(x) dt d\phi(x)\right| > \varepsilon\right) \\
\leq &\frac{1}{\varepsilon^{2m}}\mathbb{E}\left[\left(T^n \int \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n - \mathbb{E}\left[T^n \int \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n\right]\right)^{2m}\right] \\
\leq &\frac{(T^n)^{2m}}{\varepsilon^{2m}} \sum_{k\cdot s = 2m,\ k_1 = 0} \frac{(2m)!}{k!(s!)^k} a^k \\
\leq &\frac{(T^n)^{m}}{\varepsilon^{2m}} \sum_{k\cdot s = 2m,\ k_1 = 0} \frac{(2m)!}{k!(s!)^k} (T^n a)^k
\end{aligned}$$

for n large enough such that $T^n < 1$ and with a_i defined as in Lemma 2.2.5 applied to the function $t \mapsto \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right)$ and the intensity measure $\lambda = \frac{1}{T^n}\kappa_t^\dagger(x) dt d\phi$. As $\left|\log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right)\right|$ is uniformly bounded in n by C, we can estimate

$$\begin{aligned}
\left|(T^n a)^k\right| &= \prod_{i=1}^{\infty}\left(\int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \left|\log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right)\right|^i \kappa_t^\dagger(x) dt d\phi(x)\right)^{k_i} \\
&\leq \max(1, C)^m \max(1, \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \kappa_t^\dagger(x) dt d\phi(x))^m \lesssim 1
\end{aligned}$$

for all k satisfying $k \cdot s = 2m$, $k_1 = 0$ and $\sum_{i=1}^{\infty} k_i \leq m$. Thus we have

$$\begin{aligned}
&\mathbb{P}\left(\left|T^n \int \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n - \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) \kappa_t^\dagger(x) dt d\phi(x)\right| > \varepsilon\right) \\
\lesssim &(T^n)^m \xrightarrow{n\to\infty} 0,
\end{aligned}$$

establishing convergence to zero in probability. Using the property $\sum_n (T^n)^m < +\infty$, the above result implies the almost sure convergence [14, Theorem 6.12]

$$T^n \int \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n - \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) \kappa_t^\dagger(x) dt d\phi(x) \xrightarrow{\text{a.s.}} 0,$$

i.e. $Z^n \xrightarrow{\text{a.s.}} 0$. This finally implies the almost sure convergence

$$\int T^n \log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}\right) dE^n \xrightarrow{\text{a.s.}} \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \int_0^T \log\left(\kappa_t^p(x)\right) \kappa_t^\dagger(x) dt d\phi(x)$$

of the stochastic part of the functionals. Thus, in total we have shown the liminf condition unless $\rho^n \overset{*}{\rightharpoonup} \rho$.
Finally, the case $\rho^n \overset{*}{\rightharpoonup} 0$ will be considered now. Setting $\mathcal{E}^\infty(0,0) = +\infty$ we need to show $\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) = +\infty$. The crucial part of this computation is the estimation of the log-part. With N^n being the number of time intervals and M^n being the number of detector pairs and with applying the mean value theorem to get $\frac{1}{\Delta t^n}\int_{\tau_i^n} \kappa_t^{n,p}(\Gamma_{kl}^n) dt = \kappa_{t_i^n}^{n,p}(\Gamma_{kl}^n)$, we

deduce

$$
\begin{aligned}
& - T^n \int \log\left(\frac{1}{\Delta t^n \, |\Gamma^n|} \hat{\kappa}^{n,p}\right) dE^n \\
=& - T^n \int \sum_{i=1}^{N^n} \sum_{kl=1}^{M^n} \mathbb{1}_{\tau_i^n} \mathbb{1}_{\Gamma_{kl}^n} \log\left(\frac{1}{\Delta t^n \phi(\Gamma_{kl}^n)} \int_{\tau_i^n} \kappa_t^{n,p}(\Gamma_{kl}^n) dt\right) dE^n \\
=& - T^n \sum_{i=1}^{N^n} \sum_{kl=1}^{M^n} E^n(\tau_i^n \times \Gamma_{kl}^n) \log\left(\frac{1}{\phi(\Gamma_{kl}^n)} \kappa_{t_i^n}^{n,p}(\Gamma_{kl}^n)\right) \\
=& - T^n \sum_{i=1}^{N^n} \sum_{kl=1}^{M^n} E^n(\tau_i^n \times \Gamma_{kl}^n) \log\Bigg(\frac{pp^s}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)} \rho_{t_i^n}^n(\Omega) \\
& \qquad + \frac{p^d}{\phi(\Gamma_{kl}^n)} \int_{G^{1,3}} \int_{Z_v^{\Gamma_{kl}^n}} \int_\Omega G_y(x) \rho_{t_i^n}^n(dy) dx \mathrm{vol}_{G^{1,3}}(dv)\Bigg) \\
\geq& - T^n \sum_{i=1}^{N^n} \sum_{kl=1}^{M^n} E^n(\tau_i^n \times \Gamma_{kl}^n) \log\Bigg(\frac{pp^s}{\phi(\partial\Omega_\delta \times \partial\Omega_\delta)} \rho_{t_i^n}^n(\Omega) \\
& \qquad + \underbrace{p^d \frac{R_{\delta\#}(dx \otimes \mathrm{vol}_{G^{1,3}})(\Gamma_{kl}^n)}{\phi(\Gamma_{kl}^n)}}_{\lesssim 1} \|G\|_\infty \rho_{t_i^n}^n(\Omega)\Bigg) \\
\geq& - T^n \sum_{i=1}^{N^n} \sum_{kl=1}^{M^n} E^n(\tau_i^n \times \Gamma_{kl}^n) \log\left(\frac{C}{T} \int_0^T \rho_t^n(\Omega) dt\right) \\
=& \underbrace{- \log\left(\frac{C}{T} \int_0^T \rho_t^n(\Omega) dt\right)}_{I} \underbrace{T^n E^n([0,T] \times \partial\Omega_\delta \times \partial\Omega_\delta)}_{II}.
\end{aligned}
$$

Factor I goes to infinity as we have $\int_0^T \rho_t^n(\Omega) dt \xrightarrow{n\to\infty} 0$ due to $\rho^n \overset{*}{\rightharpoonup} 0$. Because of

$$
\mathbb{E}\left[T^n E^n([0,T] \times (\partial\Omega_\delta \times \partial\Omega_\delta))\right] = \int_0^T \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \kappa_t^\dagger(x) d\phi(x) dt
$$

we can apply Markov's inequality and get

$$
\begin{aligned}
& \mathbb{P}\left(\left| T^n E^n([0,T] \times (\partial\Omega_\delta \times \partial\Omega_\delta)) - \int_0^T \int_{\partial\Omega_\delta \times \partial\Omega_\delta} \kappa_t^\dagger(x) d\phi(x) dt \right| > \varepsilon\right) \\
\leq& \frac{\mathbb{E}\left[\left(T^n E^n([0,T] \times (\partial\Omega_\delta \times \partial\Omega_\delta)) - \int_0^T \kappa_t^\dagger(\partial\Omega_\delta \times \partial\Omega_\delta) dt\right)^{2m}\right]}{\varepsilon^{2m}} \\
=& \frac{(T^n)^{2m}}{\varepsilon^{2m}} \sum_{k \cdot s = 2m,\ k_1 = 0} \frac{(2m)!}{k!(s!)^k} (a)^k \lesssim \frac{(T^n)^m}{\varepsilon^{2m}} \max(1, \int_0^T \kappa_t^\dagger(\partial\Omega_\delta \times \partial\Omega_\delta) dt)^m \xrightarrow{n\to\infty} 0.
\end{aligned}
$$

This shows that factor II converges in probability to some positive constant. As above, this implies almost sure convergence because of $\sum_n (T^n)^m < +\infty$. Thus, $\liminf_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) = +\infty$ almost surely for $\rho^n \overset{*}{\rightharpoonup} 0$ as the other parts of the functional are bounded from below by zero.

This finally proves the liminf condition almost surely.

For the limsup inequality we can take $(\rho^n, \omega^n) = (\rho, \omega)$ for all n and do the same computations as above to get

$$\limsup_{n\to\infty} \mathcal{E}^n(\rho^n, \omega^n) \leq \mathcal{E}^\infty(\rho, \omega)$$

almost surely. □

Lemma 3.3.7. *Under the assumptions of Theorem 3.3.6 we have that the sequence of functions $\mathcal{E}^n$ is almost surely equi-mildly coercive, i.e. there exists a countably compact set $\mathbb{K} \subset \mathbb{M}$ such that*

$$\inf_{(\rho,\omega)\in\mathbb{M}} \mathcal{E}^n(\rho, \omega) = \inf_{(\rho,\omega)\in\mathbb{K}} \mathcal{E}^n(\rho, \omega)$$

for all $n \in \mathbb{N}$.

Proof. The proof is basically the same as in the case of fixed detector sizes that is covered by Lemma 3.2.8. Therefore, we are only going to repeat the most important and different steps.
Again, the proof relies on showing that the set of minimizers $(\rho^n, \omega^n)_n$ of the functionals $\mathcal{E}^n$ is precompact. We have

$$\min_{(\rho,\omega)\in\mathbb{M}} \mathcal{E}^n(\rho, \omega) \leq \mathcal{E}^n(\mathcal{L}, 0) < +\infty$$

and the last expression is almost surely uniformly bounded. For this proof, the uniform boundedness relies on the uniform boundedness of $\left|\log\left(\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^p\right)\right|$ established in the proof of Theorem 3.3.6 when using $(\mathcal{L}, 0) \overset{*}{\rightharpoonup} (\mathcal{L}, 0)$ instead of $(\rho^n, \omega^n) \overset{*}{\rightharpoonup} (\rho, \omega)$.
Further, we can use the same estimate of the proof of Theorem 3.3.6 (this time applied to the set of minimizers $(\rho^n, \omega^n)_n$) to conclude for $(t, x) \in \tau_i^n \times \Gamma_{kl}^n$

$$\frac{1}{\Delta t^n |\Gamma^n|}\hat{\kappa}^{n,p}(t, x) \lesssim \frac{1}{\Delta t^n \phi(\gamma_{kl}^n)} \int_{\tau_i^n} A\rho_t^n(\Gamma_{kl}^n)dt \lesssim \|\rho^n\|\,.$$

For $\overline{p}\,\|\rho^n\| > 1$ this gives us the estimate, using $\cup_{(kl)^n}\Gamma_{kl}^n = \partial\Omega_\delta \times \partial\Omega_\delta$ for all n,

$$\begin{aligned}
\mathcal{E}^n(\rho^n, \omega^n) &\geq T^n \sum_{(ikl)^n} \left[A_{ikl}^n \rho^n - E_{ikl}^n \log\left(\frac{\overline{p}}{\Delta t^n \phi(\Gamma_{kl}^n)} \int_{\tau_i^n} A\rho_t^n(\Gamma_{kl}^n)dt\right)\right] + T^n\beta^n \mathcal{S}(\rho^n, \omega^n) \\
&\geq \|\rho^n\| - T^n \sum_{(ikl)^n} E_{ikl}^n \log\left(C\,\|\rho^n\|\right) + T^n\beta^n\mathcal{S}(\rho^n, \omega^n) \\
&\geq f(\|\rho^n\|) + C''\mathcal{S}(\rho^n, \omega^n),
\end{aligned}$$

showing that the set of minimizers is precompact. □

Remark 3.3.8 (Convergence of minimizers)**.** Again, we cannot deduce from the almost sure equi-mildly coerciveness of the sequence of functions $(\mathcal{E}^n)_n$ that cluster points of a minimizing sequence of the $\mathcal{E}^n$ are minimizers of $\mathcal{E}^\infty$. The problem is the same as described in Remark 3.2.9.

CHAPTER 4

Conclusion and Outlook

We introduced a dynamic image reconstruction method for PET data. The intended use of the method is tracking single or small numbers of radioactively labeled cells. We used a maximum a posteriori estimate to determine the temporal evolution of the radioactive material distribution. Since the problem of determining this distribution from PET data is ill-posed, we had to introduce a regularization in the MAP estimation. We chose optimal transport regularization which penalizes high kinetic energies of the considered particles. Our modeling approach leads to the minimization problem

$$\min_{(\rho,\omega)} J^E(\rho,\omega) \quad \text{with} \quad J^E(\rho,\omega) = \sum_{ikl} \left[A_{ikl}\rho - E_{ikl} \log\left(pA^s_{ikl}\rho + A^d_{ikl}\rho\right)\right] + \beta\mathcal{S}(\rho,\omega)$$

where $A_{ikl} = \frac{1}{T_{1/2}} \int_{\tau_i} \kappa_t(\Gamma_{kl})dt = \frac{1}{T_{1/2}} \int_{\tau_i} (A\rho_t)(\Gamma_{kl})dt$ for the forward operator A that was derived from physical considerations of the measurement process.
The measurement E was modeled as a Poisson point process with intensity measure $\frac{1}{T_{1/2}}\kappa^\dagger_t(x)d\phi(x)dt$ which led to a stochastic formulation of the objective function J^E which we denoted by $J^{E,\Delta t,(\Gamma_{kl})_{kl}}_{\beta,T_{1/2}}$. We showed that the stochastic functional $J^{E,\Delta t,(\Gamma_{kl})_{kl}}_{\beta,T_{1/2}}$ almost surely has a minimizer in $\mathbb{M} = \mathcal{M}([0,T]\times\Omega)\times\mathcal{M}([0,T]\times\Omega)^3$. This existence could only be guaranteed almost surely since Poisson point processes are not finitely valued, but we have $\mathbb{P}(E_{ikl} < \infty) = 1$. The property $\max_{ikl} E_{ikl} < +\infty$ was important for deriving uniform norm bounds on a minimizing sequence of $J^{E,\Delta t,(\Gamma_{kl})_{kl}}_{\beta,T_{1/2}}$ and thus for showing the existence of minimizers.

The main part of the mathematical analysis of the objective function $J^{E,\Delta t,(\Gamma_{kl})_{kl}}_{\beta,T_{1/2}}$ was computing two Γ-limits. We used Γ-convergence to analyze the behavior of $J^{E,\Delta t,(\Gamma_{kl})_{kl}}_{\beta,T_{1/2}}$ when the half-life T^n converges to zero, i.e. the intensity of radiation goes to infinity. We examined two different scenarios. First, we only changed the size Δt^n of the time intervals τ^n_i. This resulted in the Γ-convergence

$$\Gamma - \lim_{n\to\infty} T^n J^{E^n,\Delta t^n,(\Gamma_{kl})_{kl}}_{\beta^n,T^n} + C^n = \mathcal{E}^\infty$$

almost surely where the Γ-limit reads

$$\mathcal{E}^\infty(\rho,\omega) = \sum_{kl} \int_0^T \left[\kappa_t(\Gamma_{kl}) - \log\left(p\kappa^s_t(\Gamma_{kl}) + \kappa^d_t(\Gamma_{kl})\right)\kappa^\dagger_t(\Gamma_{kl})\right] dt + \beta\mathcal{S}(\rho,\omega).$$

The scaling of the objective function is due to the factor $\frac{1}{T^n}$ in the definition of $A_{ikl}\rho$ and the (stochastic) constants C^n are needed to account for a different convergence behavior of

T^n and Δt^n. In the limit functional we see that the discrete character of the term

$$E_{ikl} \log\left(pA^s_{ikl}\rho + A^d_{ikl}\rho\right)$$

has disappeared and the term was replaced by

$$\int_0^T \log\left(p\kappa^s_t(\Gamma_{kl}) + \kappa^d_t(\Gamma_{kl})\right) \kappa^\dagger_t(\Gamma_{kl})dt$$

which shows a continuous behavior in the time variable. Summation with respect to the measurement process E was replaced by integration w.r.t. the intensity measure of the process.
For the second Γ-limit we additionally changed the size of the detectors. This resulted in a similar Γ-convergence behavior

$$\Gamma - \lim_{n\to\infty} T^n J^{E^n,\Delta t^n,(\Gamma^n_{kl})_{kl,n}}_{\beta^n,T^n} + C^n = \mathcal{E}^\infty$$

almost surely with a limit functional

$$\mathcal{E}^\infty(\rho,\omega) = \int_{\partial\Omega_\delta\times\partial\Omega_\delta} \int_0^T \left(\kappa_t(x) - \log\left(\kappa^p_t(x)\right)\kappa^\dagger_t(x)\right) dt d\phi(x) + \beta\mathcal{S}(\rho,\omega).$$

This time, the constants C^n account for a different convergence behavior of Δt^n and $|\Gamma^n_{kl}|$ compared to T^n and the limit functional now shows continuous behavior in the time and space variable.
Some topics were not addressed in this thesis. One could have also considered the following aspects.

- For the modeling of the forward operator we focused on the most important physical aspects and neglected a more detailed consideration. One could include anatomical information (e.g. from PET-MR or PET-CT scanners) and time of flight information. The additional a priori knowledge helps to find a good solution of the ill-posed problem.
- We used optimal transport regularization and penalized kinetic energy with the functional

  $$\mathcal{S}(\rho,\omega) = \int_0^T \int_\Omega \left(\frac{d\omega_t}{d\rho_t}\right)^2 d\rho_t dt.$$

  Other regularization terms might be more realistic from a biological point of view and thus might lead to a better reconstruction because more suitable prior knowledge would be incorporated.
- Within our stochastic framework of Γ-convergence we were not able to deduce that cluster points of a sequence of minimizers of the $\mathcal{E}^n$ are minimizers of the limit functional $\mathcal{E}^\infty$ (see Remark 3.2.9 and 3.3.8). This remains to be analyzed.
- For the analysis of the limit behavior $T^n \to 0$ we used Γ-convergence. In a further step other notions of convergence could be used. In [1] the proposed metric to compare the reconstructed distribution from simulated PET data with the true underlying distribution is the Wasserstein-Fisher-Rao metric. This metric accounts for mass localization errors as well as errors in determining the correct amount of mass.

List of Symbols

Related to Forward Operator

$\hat{\kappa}^{n,s/d/p}$ Piecewise constant, averaged functions. It is $\hat{\kappa}^{n,s/d} = \sum_{ikl} \int_{\tau_i} \kappa_t^{n,s/d}(\Gamma_{kl}) dt \mathbb{1}_{\tau_i \times \Gamma_{kl}}$ and $\hat{\kappa}^{n,p} = p\hat{\kappa}^{n,s} + \hat{\kappa}^{n,d}$. The Γ_{kl} and τ_i might as well depend on n

$\hat{\kappa}^{s/d/p}$ Piecewise constant, averaged functions. It is $\hat{\kappa}^{s/d} = \sum_{ikl} \int_{\tau_i} \kappa_t^{s/d}(\Gamma_{kl}) dt \mathbb{1}_{\tau_i \times \Gamma_{kl}}$ and $\hat{\kappa}^{p} = p\hat{\kappa}^{s} + \hat{\kappa}^{d}$

κ_t, κ_t^n Spatial part of the intensity measure of the measurement process resulting from a material distribution ρ after applying the forward operator. It depends on n if ρ does so

κ_t^p Composition of the scattering and detection part of κ_t: $\kappa_t^p = p\kappa_t^s + \kappa_t^d$

$\kappa_t^{n,p}$ Composition of the scattering and detection part of κ_t: $\kappa_t^p = p\kappa_t^{n,s} + \kappa_t^{n,d}$

$\mathbb{M}$ Domain of the objective function. $\mathbb{M} = \mathcal{M}_+([0,T] \times \Omega) \times \mathcal{M}([0,T] \times \Omega)^3$

$\mathcal{M}_+$, $\mathcal{M}$, $\mathcal{M}^d$ Nonnegative Radon measures, Radon measutes and d dimensional vector measures

ϕ $\phi = \mathcal{H}^2 \otimes \mathcal{H}^2$

A^a, A^s, A^d Forward operator describing attenuation, scattering and normal detection

E_{ikl} (Stocastic) number of detected photons in detector pair kl during i-th time interval

$G^{1,3}$ Grassmannian manifold of all one dimensional subspaces in $\mathbb{R}^3$

R Function describing the detection process of two photons

Geometry

Δt, Δt^n Temporal resolution that may depend on n

$\Gamma_{\texttt{tot}}$ Total area of the surface covered by the detectors Γ_k

Γ_{kl}, Γ_{kl}^n Area covered by the detector pair kl with a possible dependence on n, i.e. with a changing size. Subset of $\partial\Omega \times \partial\Omega$ or $\partial\Omega_\delta \times \partial\Omega_\delta$

Γ_k Area covered by the k-th detector. Subset of $\partial\Omega$ with piecewise C^1-boundary

Ω Interior of the PET scanner where the radioactive material is located in. Compact and convex subset of $\mathbb{R}^3$

Ω_δ Interior of the PET scanner whose boundary has a distance of at least δ to Ω. Compact and convex subset of $\mathbb{R}^3$ with $\Omega \subset \Omega_\delta$

$\tau_i,\ \tau_i^n$ i-th time interval that may depend on n

T Length of the considered time interval

$T_{1/2},\ T^n$ Half-life of the radionuclide that may depend on n

Other Symbols

$\lesssim,\ \gtrsim$ These symbols mean $\leq$ or $\geq$ up to a constant

$\wedge,\ \vee$ Minimum and maximum of two values

Bibliography

[1] Bernhard Schmitzer, Klaus P. Schäfers, and Benedikt Wirth. Dynamic cell imaging in PET with optimal transport regularization. *CoRR*, abs/1902.07521, 2019.

[2] Theresa Whiteside, Sandra Demaria, Maria Rodriguez-Ruiz, Hassane Zarour, and Ignacio Melero. Emerging opportunities and challenges in cancer immunotherapy. *Clinical Cancer Research*, 22:1845–1855, 04 2016.

[3] Keum Lee, Tae Jin Kim, and Guillem Pratx. Single-cell tracking with pet using a novel trajectory reconstruction algorithm. *IEEE transactions on medical imaging*, 34, 11 2014.

[4] Yu Ouyang, Tae Jin Kim, and Guillem Pratx. Evaluation of a bgo-based pet system for single-cell tracking performance by simulation and phantom studies. *Molecular Imaging*, 15, 05 2016.

[5] Luigi Ambrosio et al. *Gradient Flows In Metric Spaces and in the Space of Probability Measures.* Birkhäuser Basel, 2008.

[6] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling.* Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, 2015.

[7] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* 01 2000.

[8] Benedetto Piccoli and Francesco Rossi. Generalized wasserstein distance and its application to transport equations with source. *Archive for Rational Mechanics and Analysis*, 211(1):335–358, Jan 2014.

[9] L. Chizat, G. Peyré, and B. Schmitzer et al. An interpolating distance between optimal transport and fisher–rao metrics. *Foundations of Computational Mathematics*, 18(1), 2018.

[10] Lenaic Chizat. *Unbalanced Optimal Transport : Models, Numerical Methods, Applications.* Theses, PSL Research University, 2017.

[11] Günter Last and Mathew Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[12] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability.* The Clarendon Press Oxford University Press, New York, 1993.

[13] Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, 1997.

[14] A. Klenke. *Wahrscheinlichkeitstheorie.* Masterclass. Springer Berlin Heidelberg, 2008.

[15] Konstantinos Drakakis. On a closed formula for the derivatives of $e^{f(x)}$ and related financial applications. *International Mathematical Forum*, 2009.

[16] Pertti Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and*

Rectifiability. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1995.

[17] K.D. Schmidt. *Maß und Wahrscheinlichkeit*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2011.

[18] G.D. Maso. *An Introduction to* Γ*-convergence*. Progress in nonlinear differential equations and their applications. Birkhäuser, 1993.

[19] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer New York Dordrecht Heidelberg London, 2010.

[20] H. Federer. *Geometric Measure Theory*. Classics in Mathematics. Springer Berlin Heidelberg, 1969.

Declaration of Academic Integrity

I hereby confirm that this thesis on *Analysis of a Dynamic Cell Imaging Model in Positron Emission Tomography* is solely my own work and that I have used no sources or aids other than the ones stated. All passages in my thesis for which other sources, including electronic media, have been used, be it direct quotes or content references, have been acknowledged as such and the sources cited.

Münster, January 17, 2020

(Marco Jonas Mauritz)

I agree to have my thesis checked in order to rule out potential similarities with other works and to have my thesis stored in a database for this purpose.

Münster, January 17, 2020

(Marco Jonas Mauritz)