A Variational Framework for Region-Based Segmentation Incorporating Physical Noise Models

Alex Sawatzky · Daniel Tenbrinck · Xiaoyi Jiang · Martin Burger

Abstract Image segmentation is one of the fundamental problems in computer vision and image processing. In the recent years mathematical models based on partial differential equations and variational methods have led to superior results in many applications, e.g., medical imaging. A majority of works on image segmentation implicitly assume the given image to be biased by additive Gaussian noise, for instance the popular Mumford-Shah model. Since this assumption is not suitable for a variety of problems, we propose a region-based variational segmentation framework to segment also images with non-Gaussian noise models. Motivated by applications in biomedical imaging, we discuss the cases of Poisson and multiplicative speckle noise intensively. Analytical results such as the existence of a solution are verified and we investigate the use of different regularization functionals to provide a-priori information regarding the expected solution. The performance of the proposed framework is illustrated by experimental results on synthetic and real data.

1 Introduction

The task of automated image segmentation has become increasingly important in the last decade, due to a fast expanding field of applications, e.g., in biomedical imaging. The main goal of image segmentation is to recover an object-of-interest from a given dataset by partitioning it into disjoint compartments. In general, one can distinguish between edge-based (cf. e.g., [15, 44, 55]) and region-based (cf. e.g., [20, 22, 79]) segmentation methods. In this paper we will concentrate on the latter one, since our work is motivated by segmentation tasks in biomedical imaging, where we have to segment continuous objects-of-interest.

Recently, mathematical tools such as level sets, active contours, and variational methods led to significant improvements in automated image segmentation. One famous framework, which also allows to incorporate a-priori knowledge into the process of segmentation, is the popular Mumford-Shah (MS) model [55]. Based on this framework the frequently used Chan-Vese segmentation method [20] was developed, which simplifies the MS segmentation model to the case of piecewise constant approximations of the image intensity.

1.1 Motivation

Despite its high level of awareness in the segmentation community, the MS formulation has not yet been investigated in

A. Sawatzky (✉) · M. Burger
Institut für Numerische und Angewandte Mathematik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, 48149 Münster, Germany
e-mail: alex.sawatzky@wwu.de

M. Burger
e-mail: martin.burger@wwu.de

D. Tenbrinck · X. Jiang
Institut für Informatik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, 48149 Münster, Germany

D. Tenbrinck
e-mail: daniel.tenbrinck@wwu.de

X. Jiang
e-mail: xjiang@uni-muenster.de

a more general context of physical noise modeling. This is a crucial part in image denoising, since the image noise naturally has to be covered by the denoising method in order to produce satisfying results. Some exemplary literature on image denoising based on statistical methods can be found in [6, 19, 47]. Furthermore, only few publications considered the effect of a specific noise model on the results of image segmentation [23, 53]. Since the field of applications for automated image segmentation grows steadily, a lot of segmentation problems need a suitable noise model, e.g., synthetic aperture radar, positron emission tomography or medical ultrasound imaging. Especially for data with poor statistics, i.e., with a low signal-to-noise ratio, it is important to consider the impact of the present noise model in the process of segmentation as we will show in later sections.

Besides the application of different noise models, the selection of appropriate regularizers has a remarkable influence on the results of a segmentation approach. By incorporating a-priori knowledge into the mathematical model one is able to reduce the set of possible solutions to a subset of segmentations satisfying certain constraints. This effect can be seen best in the work of Chan-Vese [20], which restricts the set of solutions of the segmentation problem to the subset of piecewise constant functions. However, it is possible to think about situations in which the original signal is not constant at all, e.g., inhomogeneous intensity distributions in magnetic resonance imaging [81] or medical ultrasound imaging [85]. This motivates the possibility to adjust the incorporation of a-priori knowledge with the help of suitable regularizers.

In the following we describe a general segmentation framework for different physical noise models, which also allows the incorporation of a-priori knowledge by using different regularization terms. In Sect. 2 we introduce a statistical formulation of the region-based segmentation problem and present the general segmentation framework. We study our statistical formulation for three different noise models, i.e., additive Gaussian noise, Poisson noise, and multiplicative speckle noise. The framework is analyzed extensively in Sect. 3. In particular, we investigate the existence of a solution and discuss different regularizers for the presented variational segmentation formulation, which allow us to incorporate a-priori knowledge about solutions to the segmentation problem. The relationship to the classical formulations of Chan-Vese and Mumford-Shah is discussed in detail in Sect. 4. In particular, we show that the region-based variant of the popular MS model is a special case of our framework for additive Gaussian noise. Moreover, we consider a natural extension of the Chan-Vese method to non-Gaussian noise models. Section 5 describes the numerical realization of the proposed segmentation formulation. We discuss the possibilities of global convex minimization and describe how to implement the corresponding optimization schemes efficiently. We present experimental results in Sect. 6 and validate our methods on both synthetic and real data from medical imaging with challenging image attributes, e.g., high noise level and intensity inhomogeneities. Finally, this paper is ended by discussion in Sect. 8.

2 Region-Based Segmentation Framework

The main idea of the region-based segmentation framework proposed in this paper is based on the fact that a wide range of noise types is present in real-life applications, particularly including noise models that are fundamentally different from additive Gaussian noise. To formulate a segmentation framework for different noise models and thus for a large set of imaging modalities, we use tools from statistics. First, we introduce some preliminary definitions to describe our model accurately.

Let $\Omega \subset \mathbb{R}^d$ be the image domain (we consider the typical cases $d \in \{2, 3\}$) and let f be the given (noisy) image we want to segment. The segmentation problem consists in separation of the image domain Ω into an "optimal" partition $\mathcal{P}_m(\Omega)$ of pairwise disjoint regions Ω_i, $i = 1, \ldots, m$, i.e.,

$$\mathcal{P}_m(\Omega) \in \left\{ (\Omega_1, \ldots, \Omega_m) : \overline{\Omega} = \bigcup_{i=1}^{m} \overline{\Omega_i} \text{ and} \right.$$
$$\left. \Omega_i \cap \Omega_j = \emptyset \text{ for all } i \neq j \right\}. \quad (1)$$

Naturally, the partition $\mathcal{P}_m(\Omega)$ is meant to be done with respect to the given image information induced by f, e.g., separation into an object-of-interest and background for $m = 2$. Finally, we remark that in most cases the specific order of the Ω_i in (1) does not matter.

In this work we are not only interested in the partition $\mathcal{P}_m(\Omega)$ of the image domain but also in the simultaneous restoration of the given data f as an approximation of the original noise free image. For this purpose we follow the idea proposed in [12] and [84, Sect. 4.4.4] and compute a smooth function u_i for each subregion Ω_i of $\mathcal{P}_m(\Omega)$, where the smoothness of u_i is not only enforced in Ω_i but on the entire image domain Ω. Using this approach an approximation u of the noise free image is given by

$$u = \chi_{\Omega_1} u_1 + \cdots + \chi_{\Omega_m} u_m, \quad (2)$$

where χ_{Ω_i} denotes the indicator function of Ω_i and u_i is a global smooth function induced by Ω_i and the given data f, i.e.,

$$\chi_{\Omega_i}(x) = \begin{cases} 1, & \text{if } x \in \Omega_i, \\ 0, & \text{else,} \end{cases} \quad (3)$$

and

$$u_i = \begin{cases} \text{restoration of } f \text{ in } \Omega_i, \\ \text{appropriate extension in } \Omega \setminus \Omega_i. \end{cases} \quad (4)$$

2.1 Statistical Formulation of Region-Based Segmentation

In this section we provide a general region-based segmentation framework from the viewpoint of statistical (Bayesian) modeling. Following [25, 61] the partition $\mathcal{P}_m(\Omega)$ of the image domain Ω can be computed via a maximum a-posteriori probability (MAP) estimation, i.e., by maximizing the a-posteriori probability density $p(\mathcal{P}_m(\Omega)|f)$ using Bayes' theorem. However, since we also want to restore an approximation u of the original noise free image, we have to maximize a modified a-posteriori probability density as discussed below. In order to give precise statements on probability densities we use a discrete formulation with N denoting the number of pixels (or voxels) and expressing the dependency on N by a superscript in the functions (to be interpreted as piecewise constant on pixels and identified with the finite-dimensional vector of coefficients in a suitable basis) and partitions (any subdomain restricted to be a union of a finite number of pixels). As a last step we consider the formal limit $N \to \infty$ to obtain the variational model. Since this serves as a motivation only, we will not treat the challenging problem of analyzing the continuum limit, which in the case of hierarchical Bayesian priors related to the standard Mumford-Shah model was already carried out in [40].

In the following, we have to maximize an a-posteriori probability density $p(u^N, \mathcal{P}_m^N(\Omega) \mid f^N)$, which can be rewritten as

$$p(u^N, \mathcal{P}_m^N(\Omega) \mid f^N)$$
$$\propto p(\mathcal{P}_m^N(\Omega)) p(u^N \mid \mathcal{P}_m^N(\Omega)) p(f^N \mid u^N, \mathcal{P}_m^N(\Omega)). \quad (5)$$

The main advantage of the statistical formulation in (5) is the possibility to separate geometric properties of the partition of Ω (first term) from image-based features (second and third term). In addition, the densities on the right-hand side of (5) are often easier to model than the a-posteriori probability density $p(u^N, \mathcal{P}_m^N(\Omega) \mid f^N)$ itself. For the sake of completeness, we note that the probability densities $p(\mathcal{P}_m^N(\Omega))$ and $p(u^N \mid \mathcal{P}_m^N(\Omega))$ in (5) allow to incorporate a-priori information with respect to the desired partition $\mathcal{P}_m^N(\Omega)$ and the restoration u^N in (2) into the segmentation process. Moreover, the posterior distribution $p(f^N \mid u^N, \mathcal{P}_m^N(\Omega))$ in (5) solely depends on the noise model in the given data f^N, i.e., on the image formation process of the imaging device.

In order to characterize the a-priori probability density $p(\mathcal{P}_m^N(\Omega))$ in (5), we consider a geometric prior which is most frequently used in segmentation problems. This prior

provides a regularization constraint favoring smallness of the edge set

$$\Gamma^N = \bigcup_{i \neq j} (\partial \Omega_i^N \cap \partial \Omega_j^N \setminus \partial \Omega)$$

in the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$, i.e.,

$$p(\mathcal{P}_m^N(\Omega)) \propto e^{-\beta \mathcal{H}_N^{d-1}(\Gamma^N)}, \quad \beta > 0. \quad (6)$$

Note that in order to avoid unwanted grid effects one should use an appropriate approximation $\mathcal{H}_N^{d-1}$ of the Hausdorff measure which also guarantees a correct limit as $N \to \infty$.

For characterization of the two remaining densities $p(u^N \mid \mathcal{P}_m^N(\Omega))$ and $p(f^N \mid u^N, \mathcal{P}_m^N(\Omega))$ in (5) we assume in the following that the functions u_i^N in (2) are uncorrelated and independent with respect to the partition $\mathcal{P}_m^N(\Omega)$. This is natural since the segmentation should exactly separate the parts with different behavior of u^N. Hence, due to the composition of u^N by functions u_i^N and the pairwise disjoint partition of Ω^N by Ω_i^N in (1), we obtain simplified expressions of the form

$$p(u^N \mid \mathcal{P}_m^N(\Omega)) = \prod_{i=1}^m p(u_i^N \mid \Omega_i^N) \quad (7)$$

and

$$p(f^N \mid u^N, \mathcal{P}_m^N(\Omega)) = \prod_{i=1}^m p(f^N \mid u_i^N, \Omega_i^N), \quad (8)$$

where $p(u_i^N \mid \Omega_i^N)$ and $p(f^N \mid u_i^N, \Omega_i^N)$ denote for a region-of-interest Ω_i^N the probability of observing an image u_i^N and f^N, respectively.

First, we discuss the densities $p(u_i^N \mid \Omega_i^N)$ from (7), which can be reduced to a-priori probability density functions $p(u_i^N)$. The most frequently used a-priori densities, in analogy to statistical mechanics, are Gibbs functions [31, 32] of the form

$$p(u_i^N) \propto e^{-\alpha_i R_i^N(u_i^N)}, \quad \alpha_i > 0, \quad (9)$$

where R_i^N is a discretized version of a non-negative (and usually convex) energy functional R_i. Using these a-priori densities (7) is then given by

$$p(u^N \mid \mathcal{P}_m^N(\Omega)) \propto \prod_{i=1}^m e^{-\alpha_i R_i^N(u_i^N)}. \quad (10)$$

To characterize the densities $p(f^N \mid u_i^N, \Omega_i^N)$ in (8) we assume that each value $f^N_{\mid Px}$ (with $Px \subset \Omega^N$ being a pixel) describes a realization of a random variable and all random variables are pairwise independent and identically distributed in the same corresponding subregion Ω_i^N. Consequently, it is possible to replace the probability $p(f^N \mid$

u_i^N, Ω_i^N) by a joint a-posteriori probability $p_i(f^N \mid u_i^N)$ in Ω_i^N, i.e., the expression in (8) (with Px denoting a pixel) reads as

$$p(f^N \mid u^N, \mathcal{P}_m^N(\Omega))$$

$$\propto \prod_{i=1}^m \prod_{Px \subset \Omega_i^N} p_i(f^N_{\mid Px} \mid u^N_{i \mid Px}). \tag{11}$$

As mentioned above, we use the MAP estimator to determine an approximation of the unknown image u and a partition of the image domain $\mathcal{P}_m(\Omega)$. Following this approach, we maximize the a-posteriori probability (5), respectively minimize its negative logarithm, i.e.,

$$(u^N, \mathcal{P}_m^N(\Omega))_{MAP} \in \operatorname*{arg\,min}_{u^N, \mathcal{P}_m^N(\Omega)} \left\{ -\log p(f^N \mid u^N, \mathcal{P}_m^N(\Omega)) \right.$$

$$\left. -\log p(u^N \mid \mathcal{P}_m^N(\Omega)) - \log p(\mathcal{P}_m^N(\Omega)) \right\}.$$

By inserting the a-priori constraints (6) and (10) for the geometric and image terms, respectively, as well as the region-based image term (11), we consequently minimize the following energy functional,

$$E^N(u_1^N, \ldots, u_m^N, \Omega_1^N, \ldots, \Omega_m^N)$$

$$= \sum_{i=1}^m \sum_{Px \subset \Omega_i^N} -\log p_i(f^N_{\mid Px} \mid u^N_{i \mid Px})$$

$$+ \sum_{i=1}^m \alpha_i R_i^N(u_i^N) + \beta \mathcal{H}_N^{d-1}(\Gamma^N). \tag{12}$$

We already stated above that a suitable selection of probability densities $p_i(f^N \mid u_i^N)$ depends on the underlying physical noise model in the given data f^N and the subregion Ω_i^N. For the cases of additive Gaussian, Poisson, and multiplicative speckle noise we present the corresponding form of $p_i(f^N \mid u_i^N)$ in Sect. 2.2.

The variational problem (12) for the MAP estimate has a formal continuum limit (with α_i and β rescaled by the pixel volume), which we shall consider as the basis of the proposed variational framework in the following:

$$E(u_1, \ldots, u_m, \Omega_1, \ldots, \Omega_m)$$

$$= \beta \mathcal{H}^{d-1}(\Gamma)$$

$$+ \sum_{i=1}^m \left(\int_{\Omega_i} -\log p_i(f(x) \mid u_i(x)) dx + \alpha_i R_i(u_i) \right). \tag{13}$$

Finally, we add that in the context of inverse problems the functionals R_i in (13) and the Gibbs a-priori density in (9)

are related to regularization functionals, whereas the resulting functionals $\int_{\Omega_i} -\log p_i(f(x) \mid u_i(x)) dx$ are related to data fidelity terms for each subregion Ω_i.

The main advantage of the proposed region-based segmentation framework (13) is the ability to handle the information, i.e., the occurring type of noise and the desired smoothness conditions, in each subregion Ω_i of the image domain Ω separately. For example, it is possible to choose different smoothing functionals R_i if subregions of different characteristics are expected. Moreover, the framework in (13) describes a direct generalization of the Chan-Vese method and the region-based version of the Mumford-Shah segmentation model to non-Gaussian noise problems, which is further discussed in Sect. 4.

2.2 (Non-)Gaussian Noise Models

As mentioned above, the choice of probability densities $p_i(f \mid u_i)$ in (13) solely depends on the noise occurring in the data f and the subregion Ω_i. Typical examples for probability densities $p_i(f \mid u_i)$ are exponentially distributed raw data f (see e.g., [23, 53]). In most cases it is (often implicitly) assumed that the data is perturbed by additive Gaussian noise. However, there are many real-life applications in which different types of noise occur, especially signal-dependent ones, which form the main interest of this paper. In this work we focus the attention on Poisson and multiplicative speckle noise. However, we note that there are also other non-Gaussian noise models which can be of particular interest, e.g., salt-and-pepper noise or different types of multiplicative noise like Gamma noise [6], multiplicative Gaussian noise [67], or Rayleigh-distributed noise [68].

In the following we discuss the selection of probability densities $p_i(f \mid u_i)$ for additive Gaussian noise, Poisson noise, and multiplicative speckle noise. To illustrate the different characteristics of these noise forms we show in Fig. 1 a synthetic 1D signal corrupted by the mentioned three noise models. We can observe that the appearance of Poisson and speckle noise is in general stronger compared to the additive Gaussian noise. Hence, an appropriate choice of probability densities is required to handle the perturbation effects of different noise models accurately. For the sake of simplicity and since we are only interested in the formulation in (13), we will write $p_i(f(x) \mid u_i(x))$ in the following. However, this term has to be interpreted as the value of pixels in the sense of a correct modeling.

2.2.1 Additive Gaussian Noise

One of the most commonly used noise models in computer vision and mathematical image processing is the additive Gaussian noise model (see Fig. 1(b)) of the form $f = u + \eta$, where η is a Gaussian-distributed random variable with expectation 0 and variance σ^2. This kind of noise

Fig. 1 Illustration of three physical noise models in one dimension. (**a**) Noise free 1D signal. (**b**) Signal biased by additive Gaussian noise with $\sigma = 5$ (see Sect. 2.2.1). (**c**) Signal biased by Poisson noise (see Sect. 2.2.2). (**d**) Signal biased by speckle noise with $\sigma = 5$ (see Sect. 2.2.3). Note, that we cut off a few values in case of the speckle noise model to maintain a comparable vertical scale for each signal. We observe that the Poisson and multiplicative speckle noise is much stronger than the classical additive Gaussian noise

(a) original 1D signal

(b) add. Gaussian noise

(c) Poisson noise

(d) mult. speckle noise
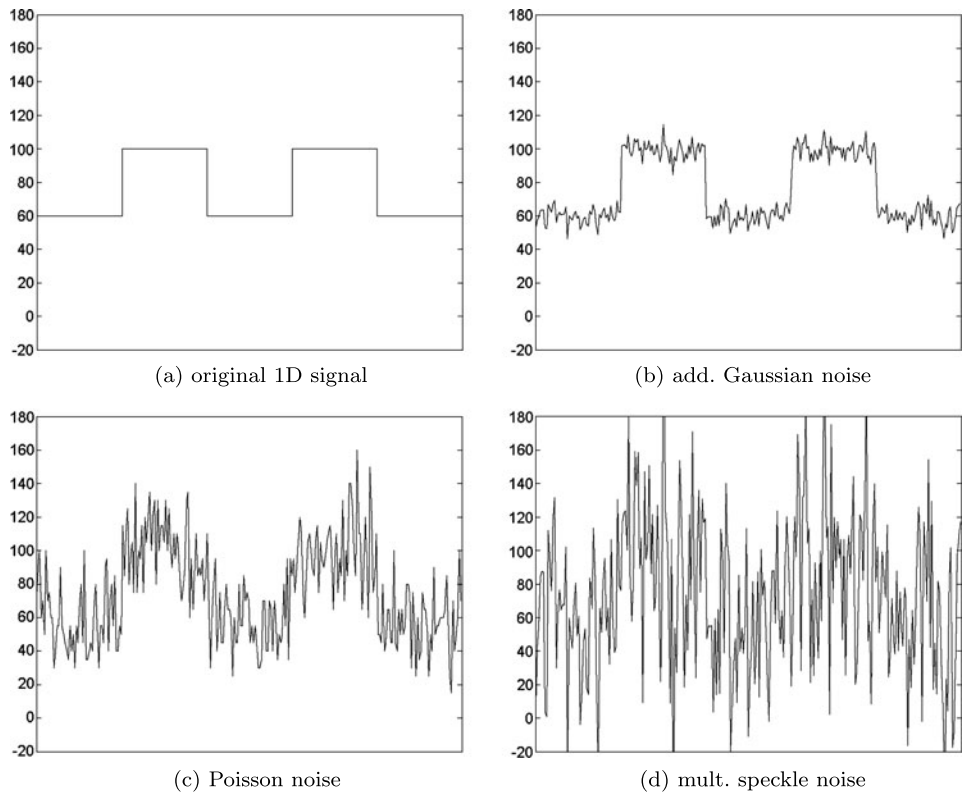
is signal-independent and has a globally identical distribution of noise. For this case the conditional probability $p_i(f(x) \mid u_i(x))$ in (11) is given by (cf. e.g., [11])

$$p_i(f(x) \mid u_i(x)) \propto e^{-\frac{1}{2\sigma^2}(u_i(x)-f(x))^2}.$$

Thus, this model leads to the following negative log-likelihood function in the energy functional E in (13),

$$-\log p_i(f(x) \mid u_i(x)) = \frac{1}{2\sigma^2}(u_i(x) - f(x))^2. \tag{14}$$

Consequently, the additive Gaussian noise model induces the commonly used L^2 data fidelity term, which is the canonical choice of fidelity in many segmentation formulations, e.g., in the Mumford-Shah or Chan-Vese model (see Sect. 4). Therefore, these segmentation methods are successful on a large class of images, since additive Gaussian noise is the most common form of noise. Finally, we mention that the factor σ^2 in (14) is neglected in the course of this work because it can be scaled by the regularization parameters α_i and β in the energy functional (13).

2.2.2 Poisson Noise

In contrast to the additive noise model described above, we are also interested in Poisson or the so-called 'photon counting noise'. In Fig. 1(c) the effect of Poisson noise on an unbiased signal can be observed. This type of noise is signal-

dependent and appears in a wide class of real-life applications, e.g., in positron emission tomography [78, 83], fluorescence microscopy [27, 41], CCD cameras [74], and astronomical images [46, 51]. For Poisson noise one indeed counts natural numbers as data, so also the image in the discrete modeling needs to be quantized. Thus, the conditional probability $p_i(f^N_{|Px} \mid u^N_{i\,|Px})$ in (11) with $Px \subset \Omega^N_i$ is modeled as (cf. e.g., [11])

$$p_i(f^N_{|Px} = k \mid u^N_{i\,|Px} = \lambda) = \frac{\lambda^k}{k!}e^{-\lambda},$$

and leads in the limit to the following negative log-likelihood function for the energy functional E in (13),

$$-\log p_i(f(x) \mid u_i(x)) = u_i(x) - f(x)\log u_i(x) + \text{const.} \tag{15}$$

Note that the fidelity can be rescaled in f and u_i in a straight-forward way, such that one can pass from quantized intensities on to real values in the reasonable limit of sufficiently high count rates. Appending additive terms independent of u_i, the corresponding data fidelity term in (15) becomes the so-called *Kullback-Leibler (KL) divergence* (also known as cross entropy or I-divergence) between two non-negative measures f and u_i. The most significant difference of the KL fidelity compared to the L^2 data term in (14) is the strong non-linearity in the KL functional, leading to complications in the computation of minimizers in (13).

2.2.3 Multiplicative Speckle Noise

The last noise model we want to investigate for the generalized segmentation framework (13) is signal-dependent noise of the form $f = u + \sqrt{u}\eta$, where η is a Gaussian-distributed random variable with mean 0 and variance σ^2. The appearance of this noise form is illustrated in Fig. 1(d), in which a spatial variation of noise variance yields different signal amplitudes. This effect occurs because the noise model is of multiplicative nature, i.e., the noise variance directly depends on the underlying signal intensity. This type of noise can be found, e.g., in diagnostic ultrasound imaging [43, 45, 52] and corresponds to an experimentally derived model of multiplicative speckle noise [77] in ultrasound images. Using this noise model, the conditional probability $p_i(f(x) \mid u_i(x))$ in (11) is modeled as

$$p_i(f(x) \mid u_i(x)) \propto (u_i(x))^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2} \frac{(u_i(x)-f(x))^2}{u_i(x)}},$$

and the negative log-likelihood function in the energy functional E in (13) is given by

$$-\log p_i(f(x) \mid u_i(x))$$
$$= \frac{1}{2\sigma^2} \frac{(u_i(x) - f(x))^2}{u_i(x)} + \frac{1}{2} \log u_i(x). \quad (16)$$

Although this noise model is somewhat similar to the additive Gaussian noise model introduced in Sect. 2.2.1, its impact on the given data f differs fundamentally from the influence of additive Gaussian noise due to the multiplicative adaption of the noise level η by the signal $\sqrt{u}$. Consequently, this aspect leads to a more complicated form of the data fidelity term and hence to more problems in the computation of minimizers in (13). As for additive Gaussian noise in Sect. 2.2.1, we can multiply the right-hand side of (16) by $2\sigma^2$ and incorporate this scaling factor in the regularization parameter α_i and β in (13).

2.3 Two-Phase Formulation of Region-Based Segmentation Framework

In the following we will restrict the proposed segmentation framework in Sect. 2.1 to a two-phase segmentation problem (i.e., we set $m = 2$ in (1) and (13)) for multiple reasons. First, we are mainly interested in medical imaging applications in which we want to segment objects in a complex background. Second, we extensively employ exact convex relaxation (see e.g., Lemma 1) in our analysis and numerical solution. In principle an extension to multiphase models can be performed with exactly the same difficulties as in the case of the standard Chan-Vese model (cf. [13, 24, 26, 37, 48, 49, 63, 64, 79]).

For a two-phase formulation of the region-based segmentation framework (13) we introduce the following notations.

First, we assume that we want to segment the image domain Ω into a background and a target subregion for the wanted partition $\mathcal{P}_2(\Omega)$ in (1), which we denote with Ω_b and Ω_t, respectively. Subsequently, we introduce an indicator function χ in order to represent both subregions such that

$$\chi(x) = \begin{cases} 1, & \text{if } x \in \Omega_b, \\ 0, & \text{else.} \end{cases} \quad (17)$$

The negative log-likelihood functions $-\log p_i(f \mid u_i)$ in (13) are defined as data fidelity functions using the notation

$$D_i(f, u_i) = -\log p_i(f \mid u_i) \quad \text{for } i \in \{b, t\}. \quad (18)$$

Finally, we use the well-known relation between the Hausdorff measure and the total variation of an indicator function (see e.g., [5, Sect. 3.3] or [35, Ex. 1.4]), which implies that $\mathcal{H}^{d-1}(\Gamma) = |\chi|_{BV(\Omega)}$. Here $\Gamma \subset \Omega$ is the edge set of the partition $\mathcal{P}_2(\Omega) = (\Omega_b, \Omega_t)$, χ is defined in (17), and $|\cdot|_{BV(\Omega)}$ denotes the total variation of a function in Ω. Then the energy functional E in (13) can be rewritten for the case of a two-phase segmentation problem as

$$E(u_b, u_t, \chi) = \int_\Omega \chi D_b(f, u_b) + (1 - \chi) D_t(f, u_t) dx$$
$$+ \alpha_b R_b(u_b) + \alpha_t R_t(u_t) + \beta |\chi|_{BV(\Omega)}. \quad (19)$$

3 Analysis of Region-Based Segmentation Framework

After introduction of the general region-based segmentation framework in Sect. 2.1 we provide a mathematical analysis of the variational problem induced by the energy functional E defined in (13) in the following. For the sake of simplicity, we restrict this analysis to the two-phase formulation proposed in (19) and thus consider a variational problem of the form

$$\min_{(u_b, u_t, \chi) \in \mathcal{D}(E)} E(u_b, u_t, \chi), \quad (20)$$

for which $\mathcal{D}(E)$ denotes the (effective) domain of the energy functional E.

3.1 Assumptions

In this section we introduce the necessary foundations for the analysis of the variational problem (20). In particular, we discuss the choice of function spaces in $\mathcal{D}(E)$ and state the required assumptions on the data fidelity terms D_i and the regularization functionals R_i in (19) with $i \in \{b, t\}$.

We start with the characterization of $\mathcal{D}(E)$ and deal with admissible sets of functions in the following (denoted by an additional (Ω)), which we assume all to be closed subsets of Banach spaces. First, we give a general formulation of

these spaces and then provide detailed conditions such that all integrals in (19) are well-defined.

Assumption 1 (Function sets) *Let $i \in \{b, t\}$, then we consider the following sets of functions*:

(i) *The domain $W_i(\Omega)$ of R_i in (19) is a subset of a Banach space $\widetilde{W}_i(\Omega)$, which is closed with respect to a topology $\tau_{\widetilde{W}_i}$ (which will later be the weak or strong norm topology depending on specific examples).*

(ii) *The domain $U_i(\Omega)$ of D_i in (19) is a subset of a Banach space $\widetilde{U}_i(\Omega)$ such that $\widetilde{W}_i(\Omega) \subset \widetilde{U}_i(\Omega)$ and the embedding $(\widetilde{W}_i(\Omega), \tau_{\widetilde{W}_i})$ into $\widetilde{U}_i(\Omega)$ with the strong norm topology is continuous.*

(iii) *The data function set is $V(\Omega) = V_b(\Omega) \cap V_t(\Omega)$, for which $V_i(\Omega)$ is a subset of a Banach space $\widetilde{V}_i(\Omega)$ such that $\widetilde{V}_i(\Omega)$ is continuously embedded in $\widetilde{U}_i(\Omega)$, both associated with the strong norm topologies.*

Thus, the effective domain $\mathcal{D}(E)$ of the energy functional E in (20) is given by

$$\mathcal{D}(E) = \{(u_b, u_t, \chi) : u_i \in W_i(\Omega),$$
$$\chi \in BV(\Omega; \{0, 1\}), E(u_b, u_t, \chi) < +\infty\}.$$

Note that Assumption 1(i) and (ii) are required since the purpose of the regularization functionals R_i is to restrict the admissible solution set of the minimization problem (20) to smooth functions. Additionally, the assumption on $V(\Omega)$ is necessary due to the occurrence of the data function f in both fidelity terms D_i in (19).

Remark 1 The properties of the function sets in Assumption 1 are chosen in this very general form in order to provide a unified framework for all cases of data fidelity and regularization terms. These abstract assumptions become more clear for specific models in the progress of this work.

To have a first impression of the function sets $W_i(\Omega)$, $\widetilde{W}_i(\Omega)$, $U_i(\Omega)$, $\widetilde{U}_i(\Omega)$, $V_i(\Omega)$, and $\widetilde{V}_i(\Omega)$, let us consider the choice of these sets for the simple example of additive Gaussian noise in the subregions Ω_i and a squared H^1-seminorm regularization functional (cf. Sect. 3.4.1 and 3.5.1). Then we have to choose the function sets as

$$\widetilde{W}_i(\Omega) = H^1(\Omega)$$

and

$$W_i(\Omega) = \left\{ u \in H^1(\Omega) : \left| \frac{1}{|\Omega|} \int_\Omega u \, dx \right| \le M_i \right\}$$

for a positive constant M_i big enough and

$$V_i(\Omega) = \widetilde{V}_i(\Omega) = U_i(\Omega) = \widetilde{U}_i(\Omega) = L^2(\Omega).$$

We continue with assumptions on the data fidelity terms D_i and the regularization functionals R_i in (19), which is necessary to ensure the existence of a regularized solution. For the sake of brevity, we introduce data fidelity functionals

$$\bar{D}_i : V_i(\Omega) \times U_i(\Omega) \times BV(\Omega; [0, 1]) \to \mathbb{R} \cup \{+\infty\}, \quad (21)$$

defined by

$$\bar{D}_i(f, u, v) = \begin{cases} \int_\Omega v D_i(f, u) \, dx, & \text{if } (v D_i(f, u)) \in L^1(\Omega), \\ +\infty, & \text{else.} \end{cases} \quad (22)$$

In the following we provide the required assumptions on the functionals $\bar{D}_i$. However, note that some of these can be transfered easily to assumptions on D_i in (22).

Assumption 2 (Energy functionals) *Let the function sets $W_i(\Omega)$, $U_i(\Omega)$, and $V_i(\Omega)$ satisfy Assumption 1. Then we assume the following*:

(i) *For any fixed $\varphi \in V_i(\Omega)$ and $\psi \in BV(\Omega; [0, 1])$, the function $u \mapsto \bar{D}_i(\varphi, u, \psi)$ is bounded from below.*

(ii) *For any fixed $\varphi \in V_i(\Omega)$ the function $u \mapsto \bar{D}_i(\varphi, u, \mathbb{1})$ is lower semi-continuous with respect to the topology $\tau_{\widetilde{W}_i}$, where $\mathbb{1}(x) = 1$ for all $x \in \Omega$.*

(iii) *The functional $R_i : W_i(\Omega) \to \mathbb{R}_{\geq 0}$ is convex and lower semi-continuous with respect to the topology $\tau_{\widetilde{W}_i}$.*

(iv) *For any fixed $\varphi \in V_i(\Omega)$ and $\psi \in BV(\Omega; [0, 1])$,*

$$\mathcal{D}(\bar{D}_i(\varphi, \cdot, \psi)) \cap \mathcal{D}(R_i) \neq \emptyset \quad (23)$$

holds, where $\mathcal{D}$ denotes the effective domain of a functional. In particular, this implies that the functionals R_i are proper.

(v) *For every $a > 0$, the sub-level sets $\mathcal{S}_{R_i}(a)$ of the functional R_i, defined by*

$$\mathcal{S}_{R_i}(a) := \{u \in W_i(\Omega) : R_i(u) \le a\}, \quad (24)$$

are sequentially precompact with respect to the topology $\tau_{\widetilde{W}_i}$.

3.2 Convex Relaxation

In this section we shortly anticipate the numerical realization of the minimization problem (20), for which we provide a theoretical basis in this section. Due to the simultaneous minimization with respect to u_b, u_t and χ, the problem (20) is hard to solve in general and we use an alternating minimization scheme to achieve our aim. This approach is commonly used for segmentation models in the literature (e.g., for the models of Ambrosio-Tortorelli [4], Chan-Vese [20],

or Mumford-Shah [55]) and leads to the following iterative minimization process,

$$(u_b^{k+1}, u_t^{k+1}) \in \underset{u_i \in W_i(\Omega)}{\arg\min} \; E(u_b, u_t, \chi^k), \tag{25a}$$

$$\chi^{k+1} \in \underset{\chi \in BV(\Omega;\{0,1\})}{\arg\min} \; E(u_b^{k+1}, u_t^{k+1}, \chi). \tag{25b}$$

However, even the minimization step (25b) is a difficult task since this problem is nonconvex due to the non-convexity of the function set $BV(\Omega; \{0, 1\})$. Considering the form of the energy functional E in (19), exact convex relaxation results for such problems have been proposed by Chan, Esedoglu and Nikolova in [22], which we recall in the following.

Lemma 1 (Exact convex relaxation) *Let $a \in \mathbb{R}$ and $g \in L^1(\Omega)$. Then there exists a minimizer of the constrained minimization problem*

$$\min_{\chi \in BV(\Omega;\{0,1\})} a + \int_{\Omega} g \chi \, dx + |\chi|_{BV(\Omega)}, \tag{26}$$

and every solution is also a minimizer of the relaxed problem

$$\min_{v \in BV(\Omega;[0,1])} a + \int_{\Omega} g v \, dx + |v|_{BV(\Omega)}, \tag{27}$$

consequently leading to the fact that the minimal functional values of (26) and (27) are equal. Moreover, if $\hat{v}$ solves (27), then for almost every $\mu \in (0, 1)$ the indicator function

$$\hat{\chi}(x) = \begin{cases} 1, & \text{if } \hat{v}(x) > \mu, \\ 0, & \text{else}, \end{cases}$$

solves (26) and thus also (27).

Since we use the alternating minimization strategy proposed in (25a)–(25b), there is an immediate implication of Lemma 1.

Theorem 1 *Let $W_i(\Omega)$, $U_i(\Omega)$, $V(\Omega)$, $\bar{D}_i$, and R_i satisfy Assumptions 1 and 2. Additionally, we assume that $f \in V(\Omega)$ and*

$$D_b(f, u_b) - D_t(f, u_t) \in L^1(\Omega), \quad \forall u_i \in W_i(\Omega). \tag{28}$$

Here, if $(\hat{u}_b, \hat{u}_t, \hat{\chi})$ is a minimizer of (20), then it is also a minimizer of the relaxed problem

$$\min_{(u_b, u_t, v) \in \mathcal{D}_{rel}(E)} E(u_b, u_t, v), \tag{29}$$

where $\mathcal{D}_{rel}(E)$ denotes the relaxed (effective) domain of the energy functional E and is given by

$$\mathcal{D}_{rel}(E) = \{(u_b, u_t, v) : u_i \in W_i(\Omega),$$
$$v \in BV(\Omega; [0, 1]), E(u_b, u_t, v) < +\infty\}.$$

Moreover, if $(\hat{u}_b, \hat{u}_t, \hat{v})$ solves (29), then for almost every $\mu \in (0, 1)$ and

$$\hat{\chi}(x) = \begin{cases} 1, & \text{if } \hat{v}(x) > \mu, \\ 0, & \text{else}, \end{cases}$$

the triple $(\hat{u}_b, \hat{u}_t, \hat{\chi})$ is a minimizer of (20).

Proof For fixed $u_b \in W_b(\Omega)$ and $u_t \in W_t(\Omega)$, set

$$a = \int_{\Omega} D_t(f, u_t) dx + \alpha_b R_b(u_b) + \alpha_t R_t(u_t)$$

and

$$g = D_b(f, u_b) - D_t(f, u_t) \overset{(28)}{\in} L^1(\Omega).$$

Thus, due to $\mathcal{D}(E) \subset \mathcal{D}_{rel}(E)$ and Lemma 1 we have

$$\min_{v \in BV(\Omega;[0,1])} E(u_b, u_t, v) = \min_{v \in BV(\Omega;\{0,1\})} E(u_b, u_t, v)$$

and we can conclude that

$$\min_{(u_b, u_t, v) \in \mathcal{D}_{rel}(E)} E(u_b, u_t, v)$$

$$= \min_{u_i \in W_i(\Omega)} \min_{v \in BV(\Omega;[0,1])} E(u_b, u_t, v)$$

$$= \min_{u_i \in W_i(\Omega)} \min_{v \in BV(\Omega;\{0,1\})} E(u_b, u_t, v)$$

$$= \min_{(u_b, u_t, v) \in \mathcal{D}(E)} E(u_b, u_t, v). \qquad \square$$

3.3 Existence of Minimizers

In this section we verify the existence of a minimizer of (20). For this reason, we first show the sequential lower semi-continuity of the energy functional E. With respect to the following lemma we note that the weak* topology represents the natural choice of a topology in $BV(\Omega)$.

Lemma 2 (Sequential lower semi-continuity) *Let $W_i(\Omega)$, $U_i(\Omega)$, $V(\Omega)$, $\bar{D}_i$, and R_i satisfy Assumptions 1 and 2, and let the data function $f \in V(\Omega)$. Moreover, let (u_i^n) be any $\tau_{\widetilde{W_i}}$-convergent sequence to some u_i in $W_i(\Omega)$ and $v^n \rightharpoonup^* v$ in $BV(\Omega)$ with $0 \le v^n \le 1$ almost everywhere. Then, for*

$$\underbrace{D_b(f, u_b^n) - D_t(f, u_t^n)}_{=:g^n} \overset{L^1(\Omega)}{\to} \underbrace{D_b(f, u_b) - D_t(f, u_t)}_{=:g}, \tag{30}$$

we have

$$E(u_b, u_t, v) \le \liminf_{n \to \infty} E(u_b^n, u_t^n, v^n).$$

Proof Due to the sequential lower semi-continuity of the functionals R_b, R_t, $\bar{D}_t(f,\cdot,\mathbb{1})$ and the total variation semi-norm $|\cdot|_{BV(\Omega)}$ (cf. e.g., [5, Prop. 3.6]), it is sufficient to show the following convergence

$$\int_\Omega v^n g^n dx \to \int_\Omega vg\,dx,$$

where g^n and g are defined as in (30). For this purpose we consider the estimate

$$\left| \int_\Omega (v^n g^n - vg)dx \right|$$
$$\leq \|v^n\|_{L^\infty(\Omega)} \|g^n - g\|_{L^1(\Omega)} + \left|\langle v^n - v, g\rangle\right|. \quad (31)$$

The first term on the right-hand side of (31) vanishes in the limit due to $\|v^n\|_{L^\infty(\Omega)} \leq 1$ and assumption (30). For the second term, the definition of the weak* topology on $BV(\Omega)$ implies $v^n \to v$ in $L^1(\Omega)$ (cf. e.g., [5, Def. 3.11]). Moreover, as $0 \leq v^n \leq 1$ a.e., the Banach-Alaoglu theorem [54, Thm. 2.6.18] delivers the existence of a subsequence (v^{n_j}) with $v^{n_j} \rightharpoonup^* \tilde{v}$ in $L^\infty(\Omega)$ and thus $v^{n_j} \rightharpoonup \tilde{v}$ in $L^1(\Omega)$. The uniqueness of the limit yields $\tilde{v} = v$, and we find that each subsequence of (v^n) has a subsequence converging to v in the weak*-sense, which implies $\left|\langle v^n - v, g\rangle\right| \to 0$. $\qquad\square$

Theorem 2 (Existence of a minimizer) *Let $W_i(\Omega)$, $U_i(\Omega)$, $V(\Omega)$, $\bar{D}_i$, and R_i satisfy Assumptions 1 and 2. Assume that α_i, $\beta > 0$ and $f \in V(\Omega)$ be fixed. Moreover, let (30) hold for any sequence (u_i^n) converging to u_i with respect to the topology $\tau_{\widetilde{W}_i}$. Then there exists a minimizer of (29) and consequently also of (20).*

Proof With respect to Assumption 2 and Lemma 2 we use the direct method of the calculus of variations (see e.g., [7, Sect. 2.1.2]). Due to the boundedness of $\bar{D}_i$ from below by a constant $C_i \in \mathbb{R}$ and due to α_i, $\beta > 0$ we have an inequality of the form

$$R_b(u_b^n) + R_t(u_t^n) + |v^n|_{BV(\Omega)}$$
$$\leq \frac{1}{\min\{\alpha_b, \alpha_t, \beta\}}(C - C_b - C_t) < \infty, \quad (32)$$

where (u_b^n, u_t^n, v^n) is a minimizing sequence of E in $\mathcal{D}_{rel}(E)$ with $E(u_b^n, u_t^n, v^n) \leq C$ and $C > (C_b + C_t)$. By using standard arguments, we obtain the existence of a minimizer of (29) and thus by Theorem 1 the existence of a minimizer of (20). $\qquad\square$

3.4 Discussion of Regularization Functionals

In this section we investigate different regularization functionals that allow to incorporate a-priori information of possible solutions in the proposed segmentation framework. We focus our attention on the Fisher information and the most frequently used squared H^1-seminorm regularization. The main objective of this section is to verify the properties of these functionals with respect to the requirements in Assumption 2.

3.4.1 The Squared H^1-Seminorm

In this section we consider the case of the squared H^1-seminorm regularization (sometimes also called Dirichlet functional),

$$R_i(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx, \quad i \in \{b, t\}. \quad (33)$$

Considering this smoothing functional, we choose the function space $\widetilde{W}_i(\Omega)$ as the Sobolev space $H^1(\Omega)$ and the subset $W_i(\Omega)$ as

$$W_i(\Omega) = \left\{ u \in H^1(\Omega) : \left| \frac{1}{|\Omega|}\int_\Omega u\,dx \right| \leq M_i \right\} \quad (34)$$

with a constant $M_i > 0$ big enough. Assumptions 2(iii) and (v) are then satisfied if $\tau_{\widetilde{W}_i}$ is the weak topology on $H^1(\Omega)$. More precisely, the following lemma gives an overview on properties of the functional in (33).

Lemma 3 (Properties of squared H^1-seminorm) *Let $W_i(\Omega)$ be the function set defined in (34) and R_i the squared H^1-seminorm (33) defined on $W_i(\Omega)$, then the following statements hold:*

(i) *R_i is convex.*
(ii) *R_i is lower semi-continuous with respect to the weak topology on $H^1(\Omega)$.*
(iii) *Let $a > 0$ and $\mathcal{S}_{R_i}(a)$ be a sub-level set of R_i defined in (24). Then $\mathcal{S}_{R_i}(a)$ is sequentially precompact with respect to the weak topology on $H^1(\Omega)$.*

Proof First, note that $W_i(\Omega) \subset H^1(\Omega)$ is convex and closed with respect to the strong norm topology on $H^1(\Omega)$. Thus, $W_i(\Omega)$ is also weakly closed (cf. e.g., [9, p. 29, Prop. 1.2.5]).

(i) It is obvious that R_i is convex on $H^1(\Omega)$ and thus on $W_i(\Omega)$, since $W_i(\Omega)$ is a convex subset.

(ii) It is also obvious that R_i defined on $H^1(\Omega)$ is lower semi-continuous with respect to the strong norm topology on $H^1(\Omega)$. By [28, p. 11, Cor. I.2.2], R_i defined on $H^1(\Omega)$ is thus also lower semi-continuous with respect to the weak topology and the assertion holds since $W_i(\Omega)$ is weakly closed.

(iii) First, we mention that the functional $\|\cdot\|^*_{p,H^1(\Omega)}$ given by

$$\|u\|^*_{p,H^1(\Omega)} := \left(\left| \frac{1}{|\Omega|}\int_\Omega u^p dx \right|^{\frac{2}{p}} + \int_\Omega |\nabla u|^2 dx \right)^{\frac{1}{2}} \quad (35)$$

defines an equivalent norm on $H^1(\Omega)$ for $p \in \{1, 2\}$ (cf. [56, Thm. 7.1]). Hence we obtain that $\|u\|_{1, H^1(\Omega)}^*$ is uniformly bounded for every $u \in \mathcal{S}_{R_i}(a)$ and the weak precompactness of sub-level sets follows directly from the Banach-Alaoglu theorem (cf. e.g., [54, Thm. 2.6.18]) since $H^1(\Omega)$ is reflexive. $\qquad\square$

Finally, we recall some basic results about the Sobolev space $H^1(\Omega)$, which is needed in the course of following sections.

Lemma 4 (Embedding of $H^1(\Omega)$) *Let $\Omega \subset \mathbb{R}^d$ be an open and bounded subset with a Lipschitz boundary, and $d \geq 1$. Then the following compact embedding of $H^1(\Omega)$ holds,*

$$H^1(\Omega) \stackrel{c}{\hookrightarrow} L^r(\Omega) \quad for \; 2 \leq r \begin{cases} \leq \infty, & if \; d = 1, \\ < \infty, & if \; d = 2, \\ < \frac{2d}{d-2}, & if \; d \geq 3. \end{cases} \quad (36)$$

Additionally, $H^1(\Omega)$ is continuously embedded in $L^r(\Omega)$ for $r = \frac{2d}{d-2}$ and $d \geq 3$.

Proof See [2, Thm. 6.2] and [3, Thm. 4.12]. $\qquad\square$

3.4.2 The Fisher Information

In the following we discuss the Fisher information as regularization functional, which can be written as

$$R_i(u) = \frac{1}{2} \int_\Omega \frac{|\nabla u|^2}{u} dx, \quad u \geq 0 \text{ a.e.}, \quad (37)$$

with $i \in \{b, t\}$. We use this form of Fisher information since the numerical motivation for this regularization energy gets more apparent in (37). For a more rigorous definition we refer to [34], in which its formulation based on a weak definition of the "logarithmic gradient" in suitable measure spaces. Under some regularity assumptions on u the energy in (37) can also be written as (cf. [34, 80])

$$R_i(u) = \frac{1}{2} \int_\Omega u |\nabla \log u|^2 dx = 2 \int_\Omega |\nabla \sqrt{u}|^2 dx. \quad (38)$$

The use of this regularization energy is motivated by the fact that the functional in (37) is one-homogeneous and thus seems to be more appropriate for density functions than the squared H^1-seminorm in (33). This is particularly significant in the context of problems with measured data corrupted by Poisson or speckle noise, since in such applications the desired functions typically represent densities or intensity information. Furthermore, the adaptive regularization property of the denominator u in (37) is additionally useful, since the background region of an image (with assumed low intensities) will be regularized stronger than the target subregion. Finally, we note that the Fisher information energy has already been used as regularization functional in density estimation problems [14, 58].

In the case of the Fisher information (37) as regularization energy, we use the result stated in [34, Lemma 2.2],

$$R_i(u) < +\infty \quad \text{if and only if}$$

$$u \geq 0 \text{ a.e.} \quad \text{and} \quad \sqrt{u} \in H^1(\Omega),$$

for which the conditions get obvious with regard to the second identity in (38). Thus we choose the function space $\widetilde{W}_i(\Omega)$ as $L^{\frac{r}{2}}(\Omega)$ with r given in (36), as well as

$$W_i(\Omega) = \Big\{ u \in L^{\frac{r}{2}}(\Omega) : u \geq 0 \text{ a.e.},$$

$$\sqrt{u} \in H^1(\Omega) \text{ and } \left| \frac{1}{|\Omega|} \int_\Omega u dx \right| \leq M_i \Big\} \quad (39)$$

with a positive constant M_i big enough. Then Assumptions 2(iii) and (v) are satisfied if $\tau_{\widetilde{W}_i}$ is the strong norm topology on $L^{\frac{r}{2}}(\Omega)$ as we can see in the following lemma.

Lemma 5 (Properties of Fisher information) *Let $W_i(\Omega)$ be the function set defined in (39) with r chosen as in (36) and R_i the Fisher information (37) defined on $W_i(\Omega)$. Then the following statements hold*:

(i) *R_i is convex.*
(ii) *R_i is lower semi-continuous with respect to the strong norm topology on $L^{\frac{r}{2}}(\Omega)$.*
(iii) *Let $a > 0$ and $\mathcal{S}_{R_i}(a)$ be a sub-level set of R_i defined in (24). Then $\mathcal{S}_{R_i}(a)$ is sequentially precompact with respect to the strong norm topology on $L^{\frac{r}{2}}(\Omega)$.*

Proof (i) It is easy to show that the mapping $(x, y) \mapsto \frac{x^2}{y}$ is convex for $y \geq 0$ defining $\frac{x^2}{y} = +\infty$ for $y = 0$. Thus R_i is convex since $W_i(\Omega)$ is a convex subset of $\{u \in L^{\frac{r}{2}}(\Omega) : u \geq 0 \text{ a.e.}\}$.

(ii) Let (u_n) be a sequence in

$$S(\Omega) := \{u \in L^{\frac{r}{2}}(\Omega) : u \geq 0 \text{ a.e.} : \sqrt{u} \in H^1(\Omega)\}$$

and $u \in S(\Omega)$ such that $u_n \to u$ in $L^{\frac{r}{2}}(\Omega)$, i.e., also $\sqrt{u_n} \to \sqrt{u}$ in $L^r(\Omega)$. Since r is given as in (36) and $\nabla : L^2(\Omega) \to H^{-1}(\Omega)$ is a continuous linear operator, we additionally have that $\nabla \sqrt{u_n} \to \nabla \sqrt{u}$ in $H^{-1}(\Omega)$, where $H^{-1}(\Omega)$ is the dual space of $H_0^1(\Omega)$. We can also conclude that the sequence $(\sqrt{u_n})$ is bounded in $H^1(\Omega)$ and thus there exists a weakly convergent subsequence in $H^1(\Omega)$ which converges against $\sqrt{u}$ due to the strong convergence of $(\sqrt{u_n})$ in $L^2(\Omega)$ and $(\nabla \sqrt{u_n})$ in $H^{-1}(\Omega)$. Thus we obtain that every subsequence of $(\sqrt{u_n})$ has a weakly convergent subsequence in $H^1(\Omega)$ with the same limit and consequently

$\sqrt{u_n} \rightharpoonup \sqrt{u}$ in $H^1(\Omega)$. Finally, since the squared H^1-seminorm (33) is lower semi-continuous with respect to the weak convergence in $H^1(\Omega)$ and due to relation (38), we obtain that the Fisher information is lower semi-continuous with respect to the strong norm topology on $L^{\frac{r}{2}}(\Omega)$.

(iii) Using the equivalent norm on $H^1(\Omega)$ in (35) with $p = 2$, we see that the set

$$\left\{ \sqrt{u} : u \in \mathcal{S}_{R_i}(a) \right\} \tag{40}$$

is bounded in $H^1(\Omega)$. Additionally, we note that the pointwise mapping $u \mapsto u^2$ is continuous from $L^r(\Omega)$ to $L^{\frac{r}{2}}(\Omega)$. Thus, the assertion is ensured since the set in (40) is compact in $L^r(\Omega)$ due to the compact embedding of $H^1(\Omega)$ in Lemma 4 and the convexity of this set. □

Remark 2 In view of the squared H^1-seminorm in (33), one might also consider an approximation of the Fisher information (37) of the form

$$R_i(u) = \frac{1}{2} \int_\Omega \frac{|\nabla u|^2}{w} dx, \quad w \in L^\infty(\Omega) \text{ and } w \geq 0 \text{ a.e.,} \tag{41}$$

where w is a given function and $i \in \{b, t\}$. Notice that the condition $w \in L^\infty(\Omega)$ (or more precisely the boundedness from above) is required to ensure the coercivity of the functional in (41). The main motivation to use this weighted version of the squared H^1-seminorm is the simpler form of (41) compared to the Fisher information formulation in (37), since the denominator in (41) is known. Furthermore, if the function w is chosen as an appropriate approximation of u such that the magnitude of w is close to the data function f, then the functional (41) is approximately one-homogeneous. Simultaneously, the analytical results from Lemma 3 also hold for the weighted version (41), due to the non-negativity of w.

3.5 Discussion of Data Fidelity Terms

We discuss in this section the properties of data fidelity terms that are used for the cases of additive Gaussian noise, Poisson noise, and multiplicative speckle noise. For this purpose we recall that a data fidelity functional $\bar{D}_i(f, u, v)$, $i \in \{b, t\}$, is induced by a negative log-likelihood function $D_i(f, u)$ (cf. (18) and (22)) by setting

$$\bar{D}_i(f, u, v) = \int_\Omega v D_i(f, u) dx \tag{42}$$

with

$$D_i(f, u) = -\log p_i(f \mid u).$$

In the case of the noise models mentioned above the negative log-likelihood function $D_i(f, u)$ is chosen as described in Sect. 2.2. In the following the objective is to verify the properties of $\bar{D}_i(f, u, v)$ regarding the requirements stated in Assumption 2. We also prove the conditions (28) and (30), which are required for the existence of a minimizer in Sect. 3.3 above. However, it is sufficient to verify only (30), due to the similarity of the respective statements. Finally, note that the following analysis strongly depends on the general choice of regularization functionals in the segmentation framework. Therefore, we focus in this work on the Fisher information and squared H^1-seminorm regularization proposed in Sect. 3.4.

3.5.1 Additive Gaussian Noise

We begin our investigation by discussion of the additive Gaussian noise model and consider the fidelity functional $\bar{D}_i(f, u, v)$ in (42) using the following negative log-likelihood function (cf. (14)),

$$D_i(f, u) = \frac{1}{2}(u - f)^2. \tag{43}$$

Due to this form, we choose the function sets in Assumption 1 as

$$V_i(\Omega) = \widetilde{V}_i(\Omega) = U_i(\Omega) = \widetilde{U}_i(\Omega) = L^2(\Omega)$$

and summarize the properties of this fidelity term in the following lemma.

Lemma 6 (Properties of additive Gaussian noise model) *Let $D_i(f, u)$ be defined as in (43). Moreover, we assume that $f \in L^2(\Omega)$ and $v \in BV(\Omega; [0, 1])$ are fixed. Then the following statements hold*:

(i) *$\bar{D}_i(f, \cdot, v)$ is nonnegative and convex on $L^2(\Omega)$.*

(ii) *$\bar{D}_i(f, \cdot, \mathbb{1})$ is lower semi-continuous with respect to the weak topology on $H^1(\Omega)$ and the strong norm topology on $L^2(\Omega)$.*

(iii) *The statement $D_i(f, u^n) \to D_i(f, u)$ in $L^1(\Omega)$ (cf. (30)) holds*

  - *if $u^n \rightharpoonup u$ in $H^1(\Omega)$, i.e., using H^1-seminorm regularization;*
  - *if $u^n \to u$ in $L^{\frac{r}{2}}(\Omega)$, r as in (36), and $d \leq 3$, i.e., using Fisher information regularization.*

Proof (i) Obviously both properties hold due to the non-negativity of v and $D_i(f, u)$, and the convexity of the mapping $u \mapsto u^2$.

(ii) It is clear that the functional $\bar{D}_i(f, \cdot, \mathbb{1})$ is lower semi-continuous with respect to the strong norm topology on $H^1(\Omega)$ and $L^2(\Omega)$. With [28, p. 11, Cor. I.2.2] it is then

also lower semi-continuous with respect to the weak topology on $H^1(\Omega)$ since the functional is convex.

(iii) Since $f \in L^2(\Omega)$, we have to show that

$$\|(u^n)^2 - u^2\|_{L^1(\Omega)} \to 0 \quad \text{and} \quad \|u^n - u\|_{L^2(\Omega)} \to 0.$$

However, due to the inequality

$$\|(u^n)^2 - u^2\|_{L^1(\Omega)} \le \|u^n + u\|_{L^2(\Omega)} \|u^n - u\|_{L^2(\Omega)}$$

and the fact that (u^n) is uniformly bounded in $L^2(\Omega)$ if the sequence is convergent, it is sufficient to show that

$$u^n \to u \quad \text{in } L^2(\Omega). \tag{44}$$

Hence, if we assume $u^n \rightharpoonup u$ in $H^1(\Omega)$, we directly obtain the required condition (44) due to the compact embedding of $H^1(\Omega)$ in $L^2(\Omega)$. Furthermore, assuming $u^n \to u$ in $L^{\frac{r}{2}}(\Omega)$ with r as given in (36), the condition (44) is fulfilled if $d \le 3$. □

3.5.2 Poisson Noise

In this section we consider the case of the Poisson noise model, for which the data fidelity functional $\bar{D}_i(f, u, v)$ in (42) is induced by the following negative log-likelihood function (cf. (15)),

$$D_i(f, u) = u - f \log u. \tag{45}$$

Disregarding additive terms independent of u, the corresponding fidelity functional $\bar{D}_i(f, u, \mathbb{1})$ in (42) is the so-called *Kullback-Leibler divergence* (also known as cross entropy or I-divergence) between two nonnegative measures f and u. Consequently, we set in Assumption 1 the function sets

$$\widetilde{U}_i(\Omega) = L^1(\Omega) \quad \text{and} \quad \widetilde{V}_i(\Omega) = L^\infty(\Omega),$$

as well as (see [65, Sects. 3.1 and 3.3])

$$U_i(\Omega) = \Big\{ u \in L^1(\Omega) : u \ge c > 0 \text{ a.e. and}$$

$$u \log u \in L^1(\Omega) \Big\} \tag{46}$$

and

$$V_i(\Omega) = \Big\{ f \in L^\infty(\Omega) : f \ge 0 \text{ a.e. and}$$

$$f \log f \in L^1(\Omega) \Big\}, \tag{47}$$

and summarize the properties of this fidelity term in the following lemma.

Lemma 7 (Properties of Poisson noise model) *Let $U_i(\Omega)$ and $V_i(\Omega)$ be defined as in (46) and (47), respectively, and $D_i(f, u)$ as in (45). We also assume that $v \in BV(\Omega; [0, 1])$ and $f \in V_i(\Omega)$ are fixed. Then the following statements hold*:

(i) $\bar{D}_i(f, \cdot, v)$ *is bounded from below on $U_i(\Omega)$.*
(ii) $\bar{D}_i(f, \cdot, \mathbb{1})$ *is lower semi-continuous with respect to the weak topology on $H^1(\Omega)$ and the strong norm topology on $L^1(\Omega)$.*
(iii) *The statement $D_i(f, u^n) \to D_i(f, u)$ in $L^1(\Omega)$ (cf. (30)) holds*

- *if $u^n \rightharpoonup u$ in $H^1(\Omega)$, i.e., using H^1-seminorm regularization;*
- *if $u^n \to u$ in $L^{\frac{r}{2}}(\Omega)$, r as in (36), i.e., using Fisher information regularization.*

Proof (i) As the function $u \mapsto D_i(f, u)$ attains its minimum at $u = f$, the assertion follows directly since v and $\|f \log f - f\|_{L^1(\Omega)}$ are bounded.

(ii) Due to the compact embedding of $H^1(\Omega)$ in $L^2(\Omega) \subset L^1(\Omega)$, it is sufficient to show the lower semi-continuity with respect to the strong norm topology on $L^1(\Omega)$. In the following we show that the functional $\bar{D}_i(f, \cdot, \mathbb{1})$ is even continuous with respect to the strong norm topology on $L^1(\Omega)$. For this purpose note that

$$|\bar{D}_i(f, u^n, \mathbb{1}) - \bar{D}_i(f, u, \mathbb{1})|$$

$$\overset{(42)}{\le} \|D_i(f, u^n) - D_i(f, u)\|_{L^1(\Omega)},$$

such that we need to prove $D_i(f, u^n) \to D_i(f, u)$ for $u^n \to u$ in $L^1(\Omega)$. Thus, using that $f \in L^\infty(\Omega)$, we have to show

$$\|u^n - u\|_{L^1(\Omega)} \to 0 \quad \text{and} \quad \|\log u^n - \log u\|_{L^1(\Omega)} \to 0.$$

However, since the log function is Lipschitz continuous on $[c, +\infty)$ with c given in (46), we obtain with the mean value theorem the following inequality,

$$\|\log u^n - \log u\|_{L^1(\Omega)} \le \frac{1}{c} \|u^n - u\|_{L^1(\Omega)},$$

and conclude that $D_i(f, u^n) \to D_i(f, u)$ if $u^n \to u$ in $L^1(\Omega)$.

(iii) To prove the assertion it suffices to show that

$$u^n \to u \quad \text{in } L^1(\Omega), \tag{48}$$

as we have seen in the proof of Lemma 7(ii). Hence, if we assume $u^n \rightharpoonup u$ in $H^1(\Omega)$, we directly obtain the required condition (48) due to the compact embedding of $H^1(\Omega)$ in $L^2(\Omega) \subset L^1(\Omega)$. Furthermore, supposing $u^n \to u$ in $L^{\frac{r}{2}}(\Omega)$ with r as given in (36), the condition (48) is fulfilled directly. □

3.5.3 Multiplicative Speckle Noise

Finally, we discuss the multiplicative speckle noise model presented in Sect. 2.2.3. Here the data fidelity functional $\bar{D}_i(f, u, v)$ in (42) is induced by the negative log-likelihood function given in (16),

$$D_i(f, u) = \frac{(u - f)^2}{u} + \sigma^2 \log u. \tag{49}$$

Considering the specific form of this function, we set the function sets in Assumption 1 as following,

$$V_i(\Omega) = \widetilde{V}_i(\Omega) = L^\infty(\Omega) \quad \text{and} \quad \widetilde{U}_i(\Omega) = L^1(\Omega),$$

as well as

$$U_i(\Omega) = \left\{ u \in L^1(\Omega) : u \geq c > 0 \text{ a.e.} \right\}, \tag{50}$$

and describe the properties of the corresponding fidelity functional in the lemma below.

Lemma 8 (Properties of multiplicative speckle noise model) *Let $U_i(\Omega)$ be as defined in (50) and $D_i(f, u)$ as in (49). Moreover, we assume that $f \in L^\infty(\Omega)$ and $v \in BV(\Omega; [0, 1])$ are fixed. Then the following statements hold:*

(i) *$\bar{D}(f, \cdot, v)$ is bounded from below on $U_i(\Omega)$.*
(ii) *$\bar{D}(f, \cdot, \mathbb{1})$ is lower semi-continuous with respect to the weak topology on $H^1(\Omega)$ and the strong norm topology on $L^1(\Omega)$.*
(iii) *The statement $D_i(f, u^n) \to D_i(f, u)$ in $L^1(\Omega)$ (cf. (30)) holds*

- *if $u^n \rightharpoonup u$ in $H^1(\Omega)$, i.e., using H^1-seminorm regularization;*
- *if $u^n \to u$ in $L^{\frac{r}{2}}(\Omega)$, r as in (36), i.e., using Fisher information regularization.*

Proof (i) Let $u \in U_i(\Omega)$, then $\bar{D}_i(f, u, v) \geq |\Omega| \sigma^2 \log c$ for all $c > 0$ in (50).

(ii) We proceed analogously to Lemma 7(ii) and thus need to prove $D_i(f, u^n) \to D_i(f, u)$ for $u^n \to u$ in $L^1(\Omega)$. In this context we use the identity

$$\frac{(u - f)^2}{u} = u - 2f + \frac{f^2}{u},$$

and exploit the results from Lemma 7(ii) with the observation that we also have to show

$$\left\| \frac{1}{u^n} - \frac{1}{u} \right\|_{L^1(\Omega)} \to 0.$$

We can use that the mapping $u \mapsto \frac{1}{u}$ is Lipschitz continuous on $[c, +\infty)$ with c as given in (50) and obtain with the mean

value theorem that

$$\left\| \frac{1}{u^n} - \frac{1}{u} \right\|_{L^1(\Omega)} \leq \frac{1}{c^2} \|u^n - u\|_{L^1(\Omega)}.$$

Finally, we obtain $D_i(f, u^n) \to D_i(f, u)$ if $u^n \to u$ in $L^1(\Omega)$ and the continuity of $\bar{D}(f, \cdot, \mathbb{1})$ with respect to the strong norm topology on $L^1(\Omega)$.

(iii) Using the results proven in Lemma 8(ii) we can argument analogously to the case of the Poisson noise model in Lemma 7(iii). □

We finally mention that the assumption $f \in L^\infty(\Omega)$ can be weakened in favor of the more natural condition $f \in L^2(\Omega)$ by an alternative proof taking into account the special structure of speckle noise data fidelity term and considering specific regularization terms. The coercivity of the functional is clear and it thus remains to consider the weak lower semi-continuity, which (by compact embeddings of $BV(\Omega)$, respectively $H^1(\Omega)$), can be reduced to the lower semi-continuity of functionals of the form

$$(v, u) \mapsto \int_\Omega v \frac{(u - f)^2}{u} dx$$

$$= \int_\Omega \left(vu - 2vf + v \frac{f^2}{u} \right) dx$$

with respect to the strong convergence of v and u in $L^p(\Omega)$ for an appropriate $p > 1$. Now using the a-priori boundedness of v it is straight-forward to pass to the limit in the first two terms. Thus it remains to verify the weak lower semi-continuity of

$$(v, u) \mapsto \int_\Omega v \frac{f^2}{u} dx.$$

We can now substitute variables and use that $w = \sqrt{v}$ still converges strongly in some $L^q(\Omega)$. Since the functional $\int_\Omega w^2 \frac{f^2}{u} dx$ is jointly convex with respect to w and u it is also lower semi-continuous in suitable reflexive Banach spaces $L^q(\Omega)$ and $L^p(\Omega)$, respectively.

We mention that in the above data term the restriction that u is bounded away from zero can also be removed using the joint convexity. However, in the overall functional the nonconvex term including $\log u$ cannot be shown to be lower semi-continuous without this restriction. This seems not to be just a technical issue, but a fundamental property of the multiplicative speckle noise model. Indeed, without a lower bound it is extremely favorable to have very small values of u if f is small due to the log term. Hence there is a certain bias for decreasing u, which might be removed if the logarithm term in the functional (49) is ignored. In our numerical tests both cases gave very similar results, so it might be possible to choose a convex speckle noise model

without the logarithmic term in practice, which is easier to handle numerically and allows a complete analysis without restrictive assumptions.

3.6 Boundedness of Function Mean Values

In previous sections we discussed the analytical properties of different regularization functionals and data fidelity terms, which were required to verify the existence of a minimizer of (20). In the course of this discussion the boundedness of function mean values in (34) and (39) was needed in order to prove the sequential precompactness of the sublevel sets of the squared H^1-seminorm and Fisher information functional. In the following we show that the existence of such bounds on mean values can be guaranteed directly if the energy functional E in (19) is bounded.

Lemma 9 (Boundedness of function mean values) *Let $V(\Omega)$ satisfy Assumption 1 and we choose the sets $U_i(\Omega)$ and $V_i(\Omega)$ as in Sect. 3.5. Moreover, let $f \in V(\Omega)$, $v \in BV(\Omega; [0, 1])$, and α_i, $\beta > 0$. Let $M_i = \infty$ for $W_i(\Omega)$ as given in (34) and (39). For a set of $(u_b, u_t, v) \in \mathcal{D}_{rel}(E)$ such that $E(u_b, u_t, v)$ is bounded, i.e., it exists a constant C with $E(u_b, u_t, v) \leq C$, there exist constants $\bar{C}_b(C)$, $\bar{C}_t(C) > 0$ independent of u_b, u_t, and v such that*

$$|c_b| \|v\|_{L^1(\Omega)} \leq \bar{C}_b(C) \quad and \quad |c_t| \|1 - v\|_{L^1(\Omega)} \leq \bar{C}_t(C), \tag{51}$$

where c_i are the mean values of u_i, $i \in \{b, t\}$.

Proof Analogously to the proof of Theorem 2, let $C > (C_b + C_t)$, for which $C_i \in \mathbb{R}$ are lower bounds of the data fidelity terms $\bar{D}_i$. We choose functions u_b and u_t such that $E(u_b, u_t, v) \leq C$ and

$$u_i \in H^1(\Omega),$$

   if $R_i =$ squared H^1-seminorm (33),

$$u_i \in \{u \in L^{\frac{r}{2}}(\Omega) : u \geq 0 \text{ a.e.}, \sqrt{u} \in H^1(\Omega)\},$$

   if $R_i =$ Fisher information (37),

where r is given as in (36) and $i \in \{b, t\}$. The aim is to show that there exist constants $\bar{C}_i(C) > 0$ independent of u_t, u_b, and v such that (51) is fulfilled. For the sake of completeness, we note that the constants $\bar{C}_i(C)$ also depend on the image domain $\Omega \subset \mathbb{R}^d$, the dimension d, the lower bounds C_i of the data fidelity terms $\bar{D}_i$, the data function f, and the regularization parameters α_i, β. However, for reasons of clarity, we refrain to indicate all dependencies and use only the bound C which is most crucial for the statement in Lemma 9.

To prove this assertion we proceed in two steps. At first we utilize the form of the data fidelity terms and show

that $\|v u_b\|_{L^1(\Omega)}$ and $\|(1 - v)u_t\|_{L^1(\Omega)}$ are bounded. Subsequently, these results and the form of the regularization functionals are used in order to show the boundedness presented in (51).

We start with the fidelity functional $\bar{D}_b(f, u_b, v)$ for the case of additive Gaussian noise discussed in Sect. 3.5.1. Due to the positivity of the used regularization functionals and the boundedness of $\bar{D}_t$ from below (cf. Assumption 2), we obtain by using the Cauchy-Schwarz inequality

$$\frac{1}{2} \left(\|v u_b\|_{L^1(\Omega)} - \|v f\|_{L^1(\Omega)} \right)^2$$
$$\leq \frac{1}{2} \|v (u_b - f)\|_{L^1(\Omega)}^2$$
$$= \frac{1}{2} \|\sqrt{v}\sqrt{v}(u_b - f)\|_{L^1(\Omega)}^2$$
$$\leq \|v\|_{L^1(\Omega)} \underbrace{\bar{D}_b(f, u_b, v)}_{\leq (C - C_t)},$$

which results in

$$\|v u_b\|_{L^1(\Omega)} \leq \sqrt{2|\Omega|(C - C_t)} + \sqrt{|\Omega|}\|f\|_{L^2(\Omega)}.$$

Thus $\|v u_b\|_{L^1(\Omega)}$ is bounded due to $f \in L^2(\Omega)$ in the case of additive Gaussian noise (cf. Sect. 3.5.1).

Next, we consider the Poisson noise fidelity functional $\bar{D}_b(f, u_b, v)$ investigated in Sect. 3.5.2. Using the properties of the Kullback-Leibler divergence presented, e.g., in [65, Lemma 3.3] (cf. comments to (45)), we obtain the estimate

$$\|v u_b - v f\|_{L^1(\Omega)}^2$$
$$\leq \left(\frac{2}{3} \|v f\|_{L^1(\Omega)} + \frac{4}{3} \|v u_b\|_{L^1(\Omega)} \right) \bar{D}_b(f, u_b, v)$$
$$\leq \left(2\|v f\|_{L^1(\Omega)} + \frac{4}{3} \|v u_b - v f\|_{L^1(\Omega)} \right)(C - C_t). \tag{52}$$

Adding $\frac{4(C - C_t)^2}{9}$ on both sides of (52) yields

$$\left(\|v u_b - v f\|_{L^1(\Omega)} - \frac{2(C - C_t)}{3} \right)^2$$
$$\leq 2\|v f\|_{L^1(\Omega)}(C - C_t) + \frac{4(C - C_t)^2}{9}$$

and consequently

$$\|v u_b\|_{L^1(\Omega)} \leq \sqrt{2\|f\|_{L^1(\Omega)}(C - C_t) + \frac{4(C - C_t)^2}{9}}$$
$$+ \frac{2(C - C_t)}{3} + \|f\|_{L^1(\Omega)}.$$

Thus $\|v u_b\|_{L^1(\Omega)}$ is bounded due to $f \in L^\infty(\Omega)$ in the case of Poisson noise (cf. Sect. 3.5.2).

We proved above that $\|vu_b\|_{L^1(\Omega)}$ is bounded in the case of additive Gaussian and Poisson noise fidelity term. Analogously we can show the boundedness of $\|(1-v)u_t\|_{L^1(\Omega)}$. In the following we prove (51), assuming that there exist constants $\tilde{C}_b(C), \tilde{C}_t(C) > 0$ such that

$$\|vu_b\|_{L^1(\Omega)} \leq \tilde{C}_b(C) \quad \text{and}$$
$$\|(1-v)u_t\|_{L^1(\Omega)} \leq \tilde{C}_t(C). \tag{53}$$

We start with the *squared H^1-seminorm* (33) as regularization functional R_i and split u_i into

$$u_i = c_i + w_i \tag{54}$$

with c_i being the mean value of u_i and $w_i := u_i - c_i$ satisfying $\int_\Omega w_i dx = 0$. For this purpose, we use the Poincaré-Wirtinger inequality (see e.g., [7, Sect. 2.5.1]), which gives us

$$\|w_i\|_{L^1(\Omega)}^2 \leq |\Omega| \|w_i\|_{L^2(\Omega)}^2$$
$$\leq C_1 \underbrace{\|\nabla w_i\|_{L^2(\Omega)}^2}_{=R_i(u_i)} \overset{(32)}{\leq} C_1 C_2, \tag{55}$$

where $C_1 > 0$ is a constant which depends on $\Omega \subset \mathbb{R}^d$ and d only, and $C_2 > 0$ is specified as in (32). Using the boundedness of $\|vu_b\|_{L^1(\Omega)}$, we obtain

$$\left(\|vc_b\|_{L^1(\Omega)} - \|vw_b\|_{L^1(\Omega)}\right)^2$$
$$\leq \|vu_b\|_{L^1(\Omega)}^2 \overset{(53)}{\leq} \left(\tilde{C}_b(C)\right)^2,$$

and consequently

$$|c_b| \|v\|_{L^1(\Omega)} \overset{(55)}{\leq} \tilde{C}_b(C) + \sqrt{C_1 C_2} \quad \left(=: \bar{C}_b(C)\right).$$

We can conclude analogously that there exists a constant $\bar{C}_t(C)$ such that $|c_t| \|1-v\|_{L^1(\Omega)} \leq \bar{C}_t(C)$.

Now we investigate the case of the *Fisher information* (37) as regularization functional R_i. Since we assume $u_i \geq 0$ a.e., we split $\sqrt{u_i}$ into

$$\sqrt{u_i} = c_i + w_i \tag{56}$$

with c_i being the mean value of $\sqrt{u_i}$ and $w_i := \sqrt{u_i} - c_i$ satisfying $\int_\Omega w_i dx = 0$. To show the boundedness of the mean value of u_i we observe

$$\left|\frac{1}{|\Omega|}\int_\Omega u_i dx\right| \leq \frac{1}{|\Omega|}\|u_i\|_{L^1(\Omega)}$$
$$\overset{(56)}{=} \frac{1}{|\Omega|}\|(c_i + w_i)^2\|_{L^1(\Omega)}$$
$$= \frac{1}{|\Omega|}\|c_i + w_i\|_{L^2(\Omega)}^2,$$

such that further estimates for $\|c_i\|_{L^2(\Omega)}$ and $\|w_i\|_{L^2(\Omega)}$ are required. As in (55) the Poincaré-Wirtinger inequality yields an estimate of $\|w_i\|_{L^2(\Omega)}$,

$$\|w_i\|_{L^2(\Omega)}^2 \leq C_3 \|\nabla w_i\|_{L^2(\Omega)}^2$$
$$\overset{(56)}{=} C_3 \underbrace{\int_\Omega |\nabla\sqrt{u_i}|^2 dx}_{\overset{(38)}{=}\frac{1}{2}R_i(u_i)}$$
$$\overset{(32)}{\leq} \frac{C_3 C_2}{2}, \tag{57}$$

where $C_3 > 0$ is a constant which depends only on $\Omega \subset \mathbb{R}^d$ and d, and $C_2 > 0$ is specified as in (32). In addition we have

$$\left(\|vc_b\|_{L^2(\Omega)} - \|vw_b\|_{L^2(\Omega)}\right)^2$$
$$\leq \|v\sqrt{u_b}\|_{L^2(\Omega)}^2 \overset{(56)}{=} \|vu_b\|_{L^1(\Omega)} \overset{(53)}{\leq} \left(\tilde{C}_b(C)\right)^2$$

and (57) yields the following estimate,

$$|c_b| \|v\|_{L^2(\Omega)} \leq \tilde{C}_b(C) + \sqrt{\frac{C_3 C_2}{2}}.$$

Thus, there exists a constant $\bar{C}_b(C)$ with $|c_b| \|v\|_{L^1(\Omega)} \leq \bar{C}_b(C)$. Analogously, we can conclude that there exists a constant $\bar{C}_t(C)$ such that $|c_t| \|1-v\|_{L^1(\Omega)} \leq \bar{C}_t(C)$. $\square$

Remark 3 With some modifications, the statement of Lemma 9 can also be proved in the case of the multiplicative speckle noise data fidelity term presented in Sect. 3.5.3. For this we note that $u_b \geq c > 0$ a.e. (50) and thus we can obtain the following estimate using the Cauchy-Schwarz inequality,

$$\|vu_b - vf\|_{L^1(\Omega)}^2 = \left\|\frac{\sqrt{vu_b}\sqrt{v}(u_b - f)}{\sqrt{u_b}}\right\|_{L^1(\Omega)}^2$$
$$\leq \|vu_b\|_{L^1(\Omega)} \int_\Omega v \frac{(u_b - f)^2}{u_b} dx.$$

Hence, for all $c > 0$ in (50) we have $\sigma^2 \log u_b \geq \sigma^2 \log c$ and

$$\|vu_b - vf\|_{L^1(\Omega)}^2$$
$$\leq \|vu_b\|_{L^1(\Omega)}\left(\bar{D}_b(f, u_b, v) - |\Omega|\sigma^2 \log c\right).$$

Now, we are able to use the same strategies as for the Poisson noise model in the proof of Lemma 9.

3.7 Existence of Minimizers for Proposed Noise and Regularization Models

In Theorem 2 we proved the existence of minimizers of (20) using the most general assumptions on data fidelity and

regularization functionals without further knowledge about the specific form of these terms (see Assumption 2). Subsequently, we verified the corresponding assumptions for three different noise models (additive Gaussian, Poisson, and multiplicative speckle noise) and two regularization functionals ((weighted) squared H^1-seminorm and Fisher information) in Sects. 3.4–3.6. For reasons of clarity we summarize the presented results for the proposed noise and regularization models again in the following.

Lemma 10 (Existence of a minimizer for proposed noise and regularization models) *Let $V(\Omega)$ satisfy Assumption 1 and we choose the sets $U_i(\Omega)$ and $V_i(\Omega)$ as in Sect. 3.5. Moreover, let $f \in V(\Omega)$ and $\alpha_i, \beta > 0$ be fixed. Let $M_i = \infty$ for $W_i(\Omega)$ as given in (34) and (39), then there exists a minimizer of (29) and consequently also of (20) corresponding to each investigated noise and regularization model.*

Proof Analogously to the proof of Theorem 2 let $C > (C_b + C_t)$, where $C_i \in \mathbb{R}$ are lower bounds of the data fidelity terms $\bar{D}_i$, which exist as proven in Lemmata 6–8(i). Moreover, let (u_b^n, u_t^n, v^n) be a minimizing sequence of E with $E(u_b^n, u_t^n, v^n) \leq C$ and

$$u_i^n \in H^1(\Omega),$$

if $R_i =$ squared H^1-seminorm (33),

$$u_i^n \in \{u \in L^{\frac{r}{2}}(\Omega) : u \geq 0 \text{ a.e.}, \sqrt{u} \in H^1(\Omega)\},$$

if $R_i =$ Fisher information (37),

for which r is given as in (36) and $i \in \{b, t\}$.

Since the functionals R_i are nonnegative the result in (32) and the constraint $\|v^n\|_{L^\infty(\Omega)} \leq 1$ imply the uniform boundedness of (v^n) in $BV(\Omega)$ and thus the existence of a weakly* convergent subsequence (cf. [1, Thm. 2.5] and [5, Prop. 3.13]), without restriction of generality v^n itself. Since the constraint $0 \leq v^n \leq 1$ prevails for $n \to \infty$ we see that the limit belongs to $BV(\Omega; [0, 1])$.

If $\|v^n\|_{L^1(\Omega)}$ and $\|1 - v^n\|_{L^1(\Omega)}$ do not tend to zero, then Lemma 9 implies that the mean values of (u_b^n) and (u_t^n) are uniformly bounded, and we can use the results proposed in Lemmata 3 and 5. Thus, we can satisfy the conditions required for the proof in Theorem 2 and obtain the existence of a minimizer of (29).

It remains to discuss the case $\|v^n\|_{L^1(\Omega)} \to 0$ as the case $\|1 - v^n\|_{L^1(\Omega)} \to 0$ is completely analogous. Due to $0 \leq v^n \leq 1$ we conclude that $\|1 - v^n\|_{L^1(\Omega)} \to 1 \neq 0$ such that by Lemma 9 the mean values of (u_t^n) are uniformly bounded. With the help of (32) we can use the results proposed in Lemmata 3(iii) and 5(iii), which ensure the existence of a convergent subsequence $(u_t^{n_j})$ that converges to some $\hat{u}_t \in W_t(\Omega)$. In addition, due to the nonnegativity of v^n and the uniqueness of the limit, v^n converges to zero also

in the weak* topology of $BV(\Omega)$. Thus, with the simple estimate

$$E(u_b^n, u_t^n, v^n) \geq \int_\Omega (1 - v^n) D_t(f, u_t^n) dx$$
$$+ \alpha_t R_t(u_t^n) + \beta |v^n|_{BV(\Omega)},$$

we can conclude that

$$\inf E(u_b, u_t, v) = \liminf_{j \to \infty} E(u_b^{n_j}, u_t^{n_j}, v^{n_j})$$
$$\geq \liminf_{j \to \infty} \left(\int_\Omega (1 - v^{n_j}) D_t(f, u_t^{n_j}) dx \right.$$
$$\left. + \alpha_t R_t(u_t^{n_j}) + \beta |v^{n_j}|_{BV(\Omega)} \right)$$
$$= \int_\Omega D_t(f, \hat{u}_t) dx + \alpha_t R_t(\hat{u}_t)$$
$$= E(\mathbb{1}, \hat{u}_t, 0),$$

showing that $(\mathbb{1}, \hat{u}_t, 0)$ is a minimizer of (29).

Finally, by Theorem 1 we conclude the existence of a minimizer of (20) corresponding to each noise and regularization model investigated in Sects. 3.4 and 3.5, namely the additive Gaussian, Poisson, and multiplicative speckle noise model in combination with the (weighted) squared H^1-seminorm and Fisher information regularization functional. □

4 Special Cases of the Segmentation Framework

In the following we investigate some special cases of the proposed segmentation framework in Sect. 2 which correspond to a region-based formulation of the popular Mumford-Shah model [84] and the Chan-Vese model [20].

4.1 Mumford-Shah Formulation

First, we study the Mumford-Shah model [55], which is well known within the segmentation community. This model can achieve edge-based segmentations of high quality for a large class of images and is formulated as a minimization problem of an energy functional of the following form,

$$E_{MS}(u, \Gamma) = \beta \mathcal{H}^{d-1}(\Gamma)$$
$$+ \frac{1}{2} \int_\Omega (u - f)^2 dx + \frac{\alpha}{2} \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx. \tag{58}$$

In this context $\Gamma \subset \Omega$ denotes the edge set, which is measured in the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$, and u is a smoothed approximation of the perturbed image

f on $\Omega \setminus \Gamma$. As already mentioned the original formulation of the Mumford-Shah model (58) is an edge-based segmentation method but can be turned into a region-based model if Γ is restricted to be the contour delineating the different subregions of Ω, e.g., in the case of two subregions Ω_1 and $\Omega_2 = \Omega \setminus \overline{\Omega}_1$ by the restriction $\Gamma = \partial \Omega_1 \setminus \partial \Omega$.

In [84] the author proposes an efficient region-based variant of the Mumford-Shah model, which is based on the modification of the boundary conditions on the edge set leading to different extensions of functions outside the subregions. The efficiency of this approach is induced by reformulating the Helmholtz-type optimality equations to the whole image domain Ω, which can be enforced by the following energy functional,

$$E_{MS*}(u_1, u_2, \Gamma)$$
$$= \beta \mathcal{H}^{d-1}(\Gamma) + \frac{1}{2}\int_{\Omega_1}(u_1 - f)^2 dx + \frac{1}{2}\int_{\Omega_2}(u_2 - f)^2 dx$$
$$+ \frac{\alpha}{2}\int_{\Omega}\left(|\nabla u_1|^2 + |\nabla u_2|^2\right)dx. \qquad (59)$$

Note that the case of two subregions Ω_1 and $\Omega_2 = \Omega \setminus \overline{\Omega}_1$ is used in (59) only for the sake of simplicity. Apparently, this functional is equivalent to our generalized segmentation framework in (13) choosing the negative log-likelihood functions $-\log p_i(f|u_i)$ as squared distances (14) and the regularizers R_i as squared H^1-seminorms (33). As already discussed in Sect. 2.2.1 this choice of likelihood functions corresponds to the assumption of the additive Gaussian noise model for the observed data f. Thus, the region-based version of the Mumford-Shah formulation in (59) represents a special case of the proposed generalized segmentation framework for an additive Gaussian noise model and H^1-seminorm regularization.

4.2 Chan-Vese Formulation

The traditional Chan-Vese segmentation model in [20] evolved as a simplification of the Mumford-Shah formulation presented in (58). It is based on the assumption that the intensities in the different subregions of Ω can be modeled as piecewise-constant functions. The Chan-Vese model can be formulated as minimization problem of an energy functional from the following form,

$$E_{CV}(c_1, c_2, \Gamma)$$
$$= \beta \mathcal{H}^{d-1}(\Gamma)$$
$$+ \frac{1}{2}\int_{\Omega_1}(c_1 - f)^2 dx + \frac{1}{2}\int_{\Omega_2}(c_2 - f)^2 dx. \qquad (60)$$

Here f is the perturbed image to be segmented and Γ again denotes the delineating contour of subregions Ω_1 and $\Omega_2 =$

$\Omega \setminus \overline{\Omega}_1$ of Ω, defined by $\Gamma = \partial \Omega_1 \setminus \partial \Omega$. The functions c_1 and c_2 are constant approximations of f in Ω_1 and Ω_2, respectively. For the Chan-Vese segmentation method the relationship to our generalized segmentation framework (13) gets obvious choosing the fidelity functions $-\log p_i(f|u_i)$ as squared distances like for the Mumford-Shah model in Sect. 4.1 and the regularizers R_i as

$$R_i(u_i) = \begin{cases} 0, & \text{if } |\nabla u_i| = 0, \\ \infty, & \text{else.} \end{cases} \qquad (61)$$

In this way we can enforce the solutions u_1 and u_2 to be constant. Analogous to the case of the Mumford-Shah formulation, we discussed in Sect. 2.2.1 that the choice of fidelity functions above corresponds to an additive Gaussian noise model assumed in the given data f. Hence, the Chan-Vese segmentation model is a special case of the proposed segmentation framework for the additive Gaussian noise model and regularizers R_i in (61), which enforce constant solutions of the segmentation problem.

4.3 Extension of Chan-Vese Formulation to Non-Gaussian Noise Models

In this section we discuss the natural extension of the Chan-Vese formulation presented in the section above to non-Gaussian noise models described in Sect. 2.2. To perform this extension it suffices to exchange the L^2 distance functions in (60) by general negative log-likelihood functions $-\log p_i(f|u_i)$ such that the functional in (60) gets the following form,

$$E_{CV*}(c_1, c_2, \Gamma)$$
$$= \beta \mathcal{H}^{d-1}(\Gamma)$$
$$+ \int_{\Omega_1} -\log p_1(f \mid c_1)dx + \int_{\Omega_2} -\log p_2(f \mid c_2)dx. \qquad (62)$$

As we can see this energy functional corresponds to the region-based segmentation framework (13) for the two-phase formulation ($m = 2$) using the regularization functionals R_i defined in (61) in order to enforce constant solutions c_1 and c_2. Actually, these optimal constants can be computed explicitly using the form of the negative log-likelihood functions. Thus, we obtain in the case of additive Gaussian (14) and Poisson (15) noise model the following constants,

$$c_i = \frac{\int_{\Omega_i} f \, dx}{|\Omega_i|},$$

and in the case of multiplicative speckle noise (16),

$$c_i = \frac{1}{2}\left(\sqrt{\sigma^4 + \frac{4\int_{\Omega_i} f^2 dx}{|\Omega_i|}} - \sigma^2\right).$$

Due to the simple form of these constants, the extension of the Chan-Vese segmentation method to non-Gaussian noise models in (62) is easy to implement and allows to be used in a wide range of applications in which piecewise constant approximations can be expected.

5 Numerical Realization

After the introduction and analysis of our region-based segmentation framework in earlier sections, we now discuss the details of its numerical implementation. For the sake of simplicity we restrict this discussion to the two-phase formulation proposed in (19) and thus deal with the realization of minimization problem (20). We use an alternating minimization strategy proposed in (25a)–(25b) in order to minimize the energy functional E in (19). Deviating from (25a)–(25b) we split the minimization step (25a) into separate subproblems regarding u_b^{k+1} and u_t^{k+1} and have to solve the following minimization problems,

$$u_b^{k+1} \in \underset{u_b \in W_b(\Omega)}{\arg\min} \left\{ \int_\Omega \chi^k D_b(f, u_b) dx + \alpha_b R_b(u_b) \right\}, \quad (63a)$$

$$u_t^{k+1} \in \underset{u_t \in W_t(\Omega)}{\arg\min} \left\{ \int_\Omega (1 - \chi^k) D_t(f, u_t) dx + \alpha_t R_t(u_t) \right\}, \quad (63b)$$

$$\chi^{k+1} \in \underset{\chi \in BV(\Omega; \{0,1\})}{\arg\min} \left\{ \beta |\chi|_{BV(\Omega)} \right. $$
$$\left. + \int_\Omega (D_b(f, u_b^{k+1}) - D_t(f, u_t^{k+1})) \chi \, dx \right\}. \quad (63c)$$

The first two minimization problems (63a) and (63b) can be interpreted as *denoising problems*, since their solutions u_i^{k+1} are meant to be approximations of the given noisy data f inside the subregions specified by χ^k and $(1 - \chi^k)$ with appropriate extensions outside of these, and which are characterized by the choice of regularization functionals R_i. The numerical realization of these two subproblems are discussed in Sects. 5.2 and 5.3 for the cases of additive Gaussian noise, multiplicative speckle noise, and Poisson noise. The third minimization problem (63c) is derived from (25b) by neglecting additive terms independent of χ. This subproblem is related to the actual partition of Ω into subregions given by χ and $(1 - \chi)$ and depends on the previously computed solutions u_i^{k+1}. The numerical solution of the minimization step (63c) is discussed in Sect. 5.4.

5.1 Numerical Realization of Denoising Steps: Preliminary Remarks

First, we discuss the numerical realization of denoising problems (63a) and (63b) for each of the noise models described in Sect. 2.2, i.e., additive Gaussian noise, Poisson

noise, and multiplicative speckle noise. To simplify this discussion, we see that both problems are of the form

$$u_i^{k+1} \in \underset{u_i \in W_i(\Omega)}{\arg\min} \left\{ \int_\Omega \tilde{\chi}_i^k D_i(f, u_i) dx + \alpha_i R_i(u_i) \right\} \quad (64)$$

with $i \in \{b, t\}$ and an indicator function $\tilde{\chi}_i^k$ such that $\tilde{\chi}_b^k = \chi^k$ in the case of (63a) and $\tilde{\chi}_t^k = (1 - \chi^k)$ in the case of (63b). In sections below we distinguish between the choice of the regularization functional R_i as the squared H^1-seminorm (33) (respectively its weighted version (41)) and the Fisher information (37) in order to propose a numerical strategy solving (64).

Finally, we mention the following aspects of the numerical realization below.

- In Assumption 2(iii) we assumed the regularization functionals R_i to be convex. Thus, we use the concept of subdifferentials from convex analysis (see e.g., [28, Sect. 5.1]) to compute the first order optimality condition of the minimization problem (64) and denote the subdifferential with ∂.
- With respect to the analytical results in Sect. 3 the optimization problem (64) implies additional complications in the computation of minimizers using the Poisson and speckle noise model, or for the Fisher information regularization. As we can see in (39) and (46) the admissible function spaces only allow non-negative solutions in these cases, which needs to be enforced in the minimization problem (64). However, for reasons of clarity, we refrain to note the non-negativity constraint on u_i in (64) explicitly but include this condition in the function set $W_i(\Omega)$ whenever it is required.

5.2 Numerical Realization of Denoising Steps: (Weighted) H^1-Seminorm Regularization

We start with a discussion of the problem (64) using the squared H^1-seminorm regularization (33) and its weighted variant (41). Since the weighted version is more general than the basic H^1-seminorm, we focus on the weighted H^1-seminorm regularization functional and denote it with $R_{H_w^1}$. In order to propose an appropriate numerical strategy minimizing the problem (64) with $R_i = R_{H_w^1}$ for each of the noise models described in Sect. 2.2, we proceed in two steps. First, we show that for each fixed $k \in \mathbb{N}$ the problem (64) can be equivalently realized by a sequence of convex variational problems of the form

$$\left(u_i^{k+1}\right)^{n+1} \in \underset{u_i}{\arg\min} \left\{ \frac{1}{2} \int_\Omega \tilde{\chi}_i^k \frac{(u_i - q^n)^2}{h^n} dx \right. $$
$$\left. + \alpha_i R_{H_w^1}(u_i) \right\}, \quad (65)$$

Table 1 Overview for the setting of functions q^n and h^n in (65) with respect to the different physical noise models proposed in Sect. 2.2

Noise model	q^n	h^n
Additive Gaussian noise	f	$\mathbb{1}$
Poisson noise	f	$(u_i^{k+1})^n$
Multiplicative speckle noise	$\frac{f^2}{(u_i^{k+1})^n} - \sigma^2$	$(u_i^{k+1})^n$

with an appropriate setting of the noise function q^n and the weighting function h^n with respect to the present noise model. The choice of these functions is summarized in Table 1. The advantage of this strategy is that it allows to propose a uniform numerical framework for the realization of the denoising steps (64) without depending on the actual form of the fidelity functions $D_i(f, u_i)$. In the second part we discuss the numerical realization of the variational problem (65).

5.2.1 Additive Gaussian Noise

The first case that we want to investigate is the minimization problem (64) using the negative log-likelihood function $D_i(f, u_i)$ for additive Gaussian noise,

$$D_i(f, u_i) \overset{(18)}{=} -\log p_i(f \mid u_i) \overset{(14)}{=} \frac{1}{2}(u_i - f)^2.$$

Then, the problem (64) is apparently a special case of (65) for $h^n \equiv 1$ and $q^n = f$, and hence no inner iterations are needed for this case.

5.2.2 Poisson Noise

We concentrate now on the problem of denoising images perturbed by Poisson noise. For this purpose we insert the following fidelity term $D_i(f, u_i)$ into (64),

$$D_i(f, u_i) \overset{(18)}{=} -\log p_i(f \mid u_i) \overset{(15)}{=} u_i - f \log u_i.$$

Note that the Poisson noise model needs an additional non-negativity constraint on the solution u_i in (64) (cf. (46)). Hence, we use the Karush-Kuhn-Tucker (KKT) optimality conditions, which formally provide the existence of a Lagrange multiplier $\lambda \geq 0$, such that the stationary points of the functional in (64) need to fulfill the equations,

$$0 = \tilde{\chi}_i^k \left(\mathbb{1} - \frac{f}{u_i^{k+1}} \right) + \alpha_i p_i^{k+1} - \lambda, \tag{66a}$$

$$0 = \lambda u_i^{k+1}, \tag{66b}$$

where $p_i^{k+1} \in \partial R_{H_w^1}(u_i^{k+1})$ and $\mathbb{1}(x) = 1$ on Ω. Multiplying (66a) by u_i^{k+1} the Lagrange multiplier λ can be eliminated

by (66b) leading to a fixed point equation of the form

$$0 = \tilde{\chi}_i^k(u_i^{k+1} - f) + \alpha_i u_i^{k+1} p_i^{k+1}.$$

Here, we use a semi-implicit iteration approach proposed in [70] (see [69, Sect. 4.5] for details) and obtain the following iteration scheme to compute u_i^{k+1},

$$\tilde{\chi}_i^k(u_i^{k+1})^{n+1} = \tilde{\chi}_i^k f - \alpha_i(u_i^{k+1})^n(p_i^{k+1})^{n+1} \tag{67}$$

with

$$(p_i^{k+1})^{n+1} \in \partial R_{H_w^1}((u_i^{k+1})^{n+1}).$$

Considering the form of (67), we can see that each step of this iteration sequence can be realized by solving the following convex variational problem,

$$(u_i^{k+1})^{n+1} \in \underset{u_i \geq 0}{\arg\min} \left\{ \frac{1}{2} \int_\Omega \tilde{\chi}_i^k \frac{(u_i - f)^2}{(u_i^{k+1})^n} dx \right.$$
$$\left. + \alpha_i R_{H_w^1}(u_i) \right\}. \tag{68}$$

Inspecting the first order optimality condition of (68) confirms the equivalence of this minimization problem with the iteration step in (67) due to convexity. Hence, the original problem (64) can be realized in the Poisson case by a sequence of minimization problems of the form (68), which is a special case of (65) with $q^n = f$ and $h^n = (u_i^{k+1})^n$. We note that we incorporate an additional non-negativity constraint in (68), and thus automatically ensure the complementarity condition (66b). Finally, using the convergence result in [69], we terminate the sequence (68) if the relative distance between two consecutive iteration solutions falls below a specified threshold, i.e., if

$$\frac{\|(u_i^{k+1})^{n+1} - (u_i^{k+1})^n\|_{L^2(\Omega)}}{\|(u_i^{k+1})^{n+1}\|_{L^2(\Omega)}} < \epsilon. \tag{69}$$

Note that we can initialize the function $(u_i^{k+1})^0$ in (68) as $(u_i^{k+1})^0 \equiv 1$. However, in our experiments we observed that a combination of smoothed results of the last step of our outer minimization strategy (63a)–(63c), i.e., $(u_i^{k+1})^0 = u^k = \chi^k u_b^k + (1 - \chi^k) u_t^k$, resulted in a better performance.

5.2.3 Multiplicative Speckle Noise

The last case that we want to discuss is dedicated to images biased by multiplicative speckle noise proposed in Sect. 2.2.3 and thus we investigate the following negative log-likelihood function $D_i(f, u_i)$ in (64),

$$D_i(f, u_i) \overset{(16)}{=} \frac{(u_i - f)^2}{u_i} + \sigma^2 \log u_i.$$

Considering the non-negativity constraint on the solution u_i in (64) (cf. Sect. 3.5.3), we can proceed analogously to the case of the Poisson noise fidelity term in Sect. 5.2.2. Thus, we obtain that each step of the resulting semi-implicit iteration sequence can be realized by solving a convex variational problem of the form (see [69, Sect. 5.3] for details),

$$\left(u_i^{k+1}\right)^{n+1} \in \operatorname*{arg\,min}_{u_i \geq 0} \left\{ \alpha_i R_{H_w^1}(u_i) \right.$$
$$\left. + \frac{1}{2} \int_\Omega \tilde{\chi}_i^k \frac{(u_i - (\frac{f^2}{(u_i^{k+1})^n} - \sigma^2))^2}{(u_i^{k+1})^n} dx \right\}. \quad (70)$$

We can see that (70) corresponds to (65) by setting $q^n = f^2/(u_i^{k+1})^n - \sigma^2$ and $h^n = (u_i^{k+1})^n$. Finally, we terminate the sequence (70) as in (69) and initialize $(u_i^{k+1})^0$ analogously to Sect. 5.2.2 as $(u_i^{k+1})^0 = u^k = \chi^k u_b^k + (1 - \chi^k) u_t^k$.

5.2.4 Numerical Realization of Variational Problem (65)

In Sects. 5.2.1–5.2.3 we have seen that the proposed numerical schemes lead to solving a sequence of variational problems introduced in its most general form in (65). In this section we discuss the numerical realization of variational problems occurring in the sequence (65). For the sake of simplicity, we neglect in the following the indices k, n and i, and hence deal with the following minimization problem using the definition of the regularization functional $R_{H_w^1}$ in (41),

$$\min_{u \geq 0} \frac{1}{2} \int_\Omega \tilde{\chi} \frac{(u - q)^2}{h} dx + \frac{\alpha}{2} \int_\Omega \frac{|\nabla u|^2}{w} dx. \quad (71)$$

Due to the presented results in (68) and (70) regarding the Poisson and speckle noise model, respectively, we consider in (71) the more interesting case of an additional non-negativity constraint. The case of the additive Gaussian noise model (i.e., where the non-negativity condition is unnecessary) is indicated in Remark 5 below.

In order to solve the problem (71) efficiently, we propose to use an augmented Lagrangian minimization approach and refer, e.g., to [30, 36, 42] for an introduction to this topic. The major motivation for this approach is that we want to obtain a unified method which can handle the weights h and w in (71) in a simple way and simultaneously incorporate the non-negativity condition on the solution without great effort. A scheme of that form has been used by the authors for inverse problems with Poisson noise and total variation regularization [69] demonstrating good performance and variability. In principle it would be possible to use alternative techniques, e.g., Newton's method or a conjugate gradient method, particularly for the case of additive Gaussian noise and the standard H^1-seminorm regularization (i.e., using $h = w \equiv 1$ in (71)). However, one has to take special

care for the non-negativity constraint. Furthermore, if the weights h in the data fidelity term and w in the regularization functional are not identity functions, further problems can arise, such as choosing a proper preconditioner matrix for the conjugate gradient method for instance. Finally, another motivation for using the augmented Lagrangian approach is that the most expensive part of this iteration sequence is the computation of a Helmholtz-type optimality equation (see (74) below), which can be solved efficiently and exactly using the discrete cosine transform. This has the advantage that it provides the opportunity to accelerate this sub-step directly using fast methods, e.g., by using GPU-based approaches.

Due to the weights $\tilde{\chi}$, h and w occurring in (71), we introduce auxiliary functions $\tilde{u}$ and v to simplify handling of these weights. Hence, we consider the following equivalent constrained optimization problem,

$$\min_{u, \tilde{u}, v} \frac{1}{2} \int_\Omega \tilde{\chi} \frac{(\tilde{u} - q)^2}{h} dx + \frac{\alpha}{2} \int_\Omega \frac{|v|^2}{w} dx + i_{\geq 0}(\tilde{u})$$
$$\text{s.t.} \quad \tilde{u} = u \quad \text{and} \quad v = \nabla u, \quad (72)$$

where $i_{\geq 0}$ is an indicator functional given by $i_{\geq 0}(u) = 0$ if $u \geq 0$ a.e. and $+\infty$ else. Following the idea of augmented Lagrangian methods and using the standard Uzawa algorithm (without preconditioning) [29], we obtain an alternating minimization scheme given by (cf. [69, Sect. 6.3.4] for a detailed description in the case of total variation regularization)

$$u^{l+1} \in \operatorname*{arg\,min}_u \left\{ \left\langle \lambda_{\tilde{u}}^l, \tilde{u}^l - u \right\rangle + \frac{\mu_{\tilde{u}}}{2} \|\tilde{u}^l - u\|_{L^2(\Omega)}^2 \right.$$
$$\left. + \left\langle \lambda_v^l, v^l - \nabla u \right\rangle + \frac{\mu_v}{2} \|v^l - \nabla u\|_{L^2(\Omega)}^2 \right\}, \quad (73a)$$

$$\tilde{u}^{l+1} \in \operatorname*{arg\,min}_{\tilde{u}} \left\{ \frac{1}{2} \int_\Omega \tilde{\chi} \frac{(\tilde{u} - q)^2}{h} dx + i_{\geq 0}(\tilde{u}) \right.$$
$$\left. + \left\langle \lambda_{\tilde{u}}^l, \tilde{u} - u^{l+1} \right\rangle + \frac{\mu_{\tilde{u}}}{2} \|\tilde{u} - u^{l+1}\|_{L^2(\Omega)}^2 \right\}, \quad (73b)$$

$$v^{l+1} \in \operatorname*{arg\,min}_v \left\{ \frac{\alpha}{2} \int_\Omega \frac{|v|^2}{w} dx \right.$$
$$\left. + \left\langle \lambda_v^l, v - \nabla u^{l+1} \right\rangle + \frac{\mu_v}{2} \|v - \nabla u^{l+1}\|_{L^2(\Omega)}^2 \right\}, \quad (73c)$$

$$\lambda_{\tilde{u}}^{l+1} = \lambda_{\tilde{u}}^l + \mu_{\tilde{u}} \left(\tilde{u}^{l+1} - u^{l+1} \right), \quad (73d)$$

$$\lambda_v^{l+1} = \lambda_v^l + \mu_v \left(v^{l+1} - \nabla u^{l+1} \right), \quad (73e)$$

where $\lambda_{\tilde{u}}$, λ_v are Lagrange multipliers and $\mu_{\tilde{u}}$, μ_v are positive relaxation parameters. The efficiency of this strategy strongly depends on how fast one can solve each of the sub-problems in (73a)–(73e). First, the problem (73a) is differ-

entiable with the following Helmholtz-type optimality equation assuming Neumann boundary conditions,

$$(\mu_{\tilde{u}} I - \mu_v \Delta) u^{l+1}$$
$$= \underbrace{\lambda_{\tilde{u}}^l + \mu_{\tilde{u}} \tilde{u}^l - \text{div}\left(\lambda_v^l + \mu_v v^l\right)}_{=:z^l}, \tag{74}$$

where I is the identity operator and Δ denotes the Laplace operator. In the discrete setting using finite difference discretization on a rectangular domain Ω, this equation can be solved efficiently by the discrete cosine transform (DCT-II), since $-\Delta$ is diagonalizable in the DCT-transformed space,

$$u^{l+1} = DCT^{-1}\left(\frac{DCT\left(z^l\right)}{\mu_{\tilde{u}} + \mu_v \hat{k}}\right), \tag{75}$$

where z^l is defined in (74), $\hat{k}$ represents the negative Laplace operator in the discrete cosine space (see [69, Sect. 6.3.4]), and DCT^{-1} denotes the inverse discrete cosine transform. Moreover, the solution of the minimization problem (73b) can be computed by an explicit formula of the form

$$\tilde{u}^{l+1} = \max\left\{\frac{\tilde{\chi} q + h\left(\mu_{\tilde{u}} u^{l+1} - \lambda_{\tilde{u}}^l\right)}{\tilde{\chi} + \mu_{\tilde{u}} h}, 0\right\}, \tag{76}$$

where the maximum operation has to be interpreted pointwise (i.e., for each $x \in \Omega$). Finally, the sub-problem (73c) can be computed by the following explicit formula,

$$v^{l+1} = \frac{w\left(\mu_v \nabla u^{l+1} - \lambda_v^l\right)}{\alpha + \mu_v w}. \tag{77}$$

Remark 4 From a practical point-of-view, we note that the augmented Lagrangian approach presented above can be simplified for the original squared H^1-seminorm regularization (33) in (71), i.e., considering $w \equiv 1$ in (71). In this special case the auxiliary function $v = \nabla u$ in (72) is not necessarily required and hence the iteration steps (73e) and (77) disappear. In addition, the formula in (75) can be simplified and the resulting numerical scheme has only to perform the following iteration sequence,

$$u^{l+1} = DCT^{-1}\left(\frac{DCT\left(\lambda_{\tilde{u}}^l + \mu_{\tilde{u}} \tilde{u}^l\right)}{\mu_{\tilde{u}} + \alpha \hat{k}}\right),$$

followed by steps in (76) and (73d). This increases the algorithm performance significantly in the special case of $w \equiv 1$ compared to the more general approach presented in (73a)–(73e) and (77).

Remark 5 Finally, we want to discuss the case where the non-negativity condition in (71) is unnecessary, i.e. considering the additive Gaussian noise model. The only difference is that the indicator functional $i_{\geq 0}$ in (72) and (73b) is

no longer required and hence we only have to modify (76) to

$$\tilde{u}^{l+1} = \frac{\tilde{\chi} q + h(\mu_{\tilde{u}} u^{l+1} - \lambda_{\tilde{u}}^l)}{\tilde{\chi} + \mu_{\tilde{u}} h}.$$

5.3 Numerical Realization of Denoising Steps: Fisher Information Regularization

In Sect. 5.2 we considered the numerical realization of the denoising problem (64) using the (weighted) squared H^1-seminorm regularization R_i. In this section we discuss the case of the Fisher information regularization defined in (37), i.e., we consider

$$u_i^{k+1} \in \arg\min_{u_i \geq 0}\left\{\int_\Omega \tilde{\chi}_i^k D_i(f, u_i)dx + \frac{\alpha_i}{2}\int_\Omega \frac{|\nabla u_i|^2}{u_i}dx\right\}. \tag{78}$$

Note that in the case of the Fisher information regularization a non-negativity constraint on the solution is always required due to the admissible domain of the functional in (39).

Before we discuss the implementation of (78), we explain why the differentiation between the Fisher information and squared H^1-seminorm regularization is required to propose a numerical strategy minimizing (64). It is apparent that for the Fisher information regularization we can proceed analogously to Sect. 5.2 and propose an iteration scheme solving the following minimization problem (cf. (71)),

$$\min_{u \geq 0} \frac{1}{2}\int_\Omega \tilde{\chi}\frac{(u-q)^2}{h}dx + \frac{\alpha}{2}\int_\Omega \frac{|\nabla u|^2}{u}dx. \tag{79}$$

The complexity of this minimization functional, compared to the one in (71), arises from the unknown denominator in the regularization energy in (79). Hence, the augmented Lagrangian approach used in Sect. 5.2.4 leads for (79) to several realization problems, which require further detailed investigations and hence go beyond the scope of this work. Thus, we decided to omit these additional discussions from this paper.

We rather propose to use another approach to solve (64) with the Fisher information (37) as regularization functional R_i in the following. There we derive the following iteration scheme for each fixed $k \in \mathbb{N}$,

$$\left(cI - \alpha_i \Delta\right)(u_i^{k+1})^{n+1}$$
$$= \underbrace{(c - r^n)(u_i^{k+1})^n + s^n - \frac{\alpha_i}{2}\frac{|\nabla(u_i^{k+1})^n|^2}{(u_i^{k+1})^n}}_{=:\tilde{z}^n}, \tag{80}$$

where $c > 0$ is a constant, while I and Δ denote the identity and Laplace operator, respectively. The choice of functions r^n and s^n in (80) is discussed below for each of the noise

Table 2 Overview for the setting of functions r^n and s^n in (80) with respect to the different physical noise models proposed in Sect. 2.2

Noise model	r^n	s^n
Additive Gaussian noise		
if $f \geq 0$ everywhere	$\tilde{\chi}_i^k (u_i^{k+1})^n$	$\tilde{\chi}_i^k (u_i^{k+1})^n f$
else	$\tilde{\chi}_i^k ((u_i^{k+1})^n - f)$	0
Poisson noise	$\tilde{\chi}_i^k$	$\tilde{\chi}_i^k f$
Multiplicative speckle noise	$\tilde{\chi}_i^k (1 + \frac{\sigma^2}{(u_i^{k+1})^n})$	$\tilde{\chi}_i^k \frac{f^2}{(u_i^{k+1})^n}$

models described in Sect. 2.2 and summarized in Table 2. Now, each iteration step in (80) requires the solution of a Helmholtz-type equation, which can be realized efficiently in the discrete setting using the discrete cosine transform,

$$(u_i^{k+1})^{n+1} = DCT^{-1} \left(\frac{DCT(\tilde{z}^n)}{c + \alpha_i \hat{k}} \right),$$

using the notation of (75) with $\tilde{z}^n$ as defined as in (80). Finally, we have to ensure the non-negativity of $(u_i^{k+1})^n$ for each $n \in \mathbb{N}$ since the solution of this iteration sequence should be non-negative in the case of Fisher information regularization (see (39)). This can be guaranteed if the right-hand side of (80) is non-negative due to the maximum principle for elliptic partial differential equations. To ensure this property we have to choose the functions s^n and $(u_i^{k+1})^0$ non-negative and

$$c \geq \underset{\Omega}{\text{ess sup}} \left(r^n + \frac{\alpha_i}{2} \frac{|\nabla (u_i^{k+1})^n|^2}{((u_i^{k+1})^n)^2} \right) \tag{81}$$

for each $n \in \mathbb{N}$, whereas the latter condition is a simple consequence of the negative terms occurring in (80). The termination of (80) is realized as in (69). With respect to the convergence of (80), which is depending on c, we refer to [73] and the references therein, where the convergence of similar numerical schemes is discussed. It is obvious that we cannot guarantee a sufficiently fast convergence of the iteration sequence (80) due to potentially high values of c in (81). Thus, further development of alternative algorithms will be required in the future.

In the following we discuss the choice of functions r^n and s^n in (80) with respect to additive Gaussian, Poisson, and multiplicative speckle noise.

5.3.1 Additive Gaussian Noise

We start with the investigation of (78) using the negative log-likelihood function $D_i(f, u_i)$ corresponding to the additive Gaussian noise presented in (14), i.e., we have $D_i(f, u_i) = \frac{1}{2}(u_i - f)^2$. Due to the non-negativity constraint on the solution u_i in (78) we use KKT conditions, which formally provide the existence of a Lagrange multiplier $\lambda \geq 0$, such

that the stationary points of the functional in (78) need to fulfill the equations,

$$0 = \tilde{\chi}_i^k (u_i^{k+1} - f) - \alpha_i \frac{\Delta u_i^{k+1}}{u_i^{k+1}} + \frac{\alpha_i}{2} \frac{|\nabla u_i^{k+1}|^2}{(u_i^{k+1})^2} - \lambda, \tag{82a}$$

$$0 = \lambda u_i^{k+1}, \tag{82b}$$

for which we assume Neumann boundary conditions. Multiplying (82a) by u_i^{k+1} the Lagrange multiplier λ can be eliminated by (82b) and the subsequent addition of the term $c u_i^{k+1} - c u_i^{k+1}$ with $c > 0$ leads to a fixed point equation of the form

$$0 = c u_i^{k+1} - c u_i^{k+1} + \tilde{\chi}_i^k u_i^{k+1} (u_i^{k+1} - f)$$
$$- \alpha_i \Delta u_i^{k+1} + \frac{\alpha_i}{2} \frac{|\nabla u_i^{k+1}|^2}{u_i^{k+1}}. \tag{83}$$

Using a semi-implicit approach, we now obtain the iteration sequence (80) by setting

- $r^n = \tilde{\chi}_i^k (u_i^{k+1})^n$ and $s^n = \tilde{\chi}_i^k (u_i^{k+1})^n f$, if f is non-negative everywhere;
- $r^n = \tilde{\chi}_i^k ((u_i^{k+1})^n - f)$ and $s^n = 0$ else.

The case differentiation presented here is motivated by practical reasons. It is obvious that we can always choose c in (81) large enough such that the right-hand side of (80) is non-negative. On the other hand, it should be as small as possible to yield a reasonable convergence behavior of the iteration sequence. Thus, in the case of additive Gaussian noise, where the data function f can also be negative, using the functions above is more convenient. Finally, note that the resulting fixed point of (83) is a minimizer of (78) if the corresponding derivative has the correct sign, a condition which has to be checked in the numerical realization.

5.3.2 Poisson Noise

Here we consider the problem (78) in the case of Poisson noise, where the negative log-likelihood function $D_i(f, u_i)$ is given in (15), i.e., $D_i(f, u_i) = u_i - f \log u_i$. In this case we can proceed analogously to the additive Gaussian noise discussed in Sect. 5.3.1 and hence obtain the following fixed point equation using KKT and Neumann boundary conditions,

$$c u_i^{k+1} - c u_i^{k+1} + \tilde{\chi}_i^k (u_i^{k+1} - f)$$
$$- \alpha_i \Delta u_i^{k+1} + \frac{\alpha_i}{2} \frac{|\nabla u_i^{k+1}|^2}{u_i^{k+1}} = 0.$$

Then a semi-implicit approach yields the iteration sequence (80) with $r^n = \tilde{\chi}_i^k$ and $s^n = \tilde{\chi}_i^k f$. Note that the non-negativity of s^n results from Sect. 3.5.2 since $f \in V_i(\Omega)$ is non-negative.

5.3.3 Multiplicative Speckle Noise

Finally, we discuss (78) for the case of the speckle noise likelihood function introduced in (49), i.e., we consider $D_i(f, u_i) = \frac{(u_i - f)^2}{u_i} + \sigma^2 \log u_i$. In this case the KKT conditions give a fixed point equation of the form

$$cu_i^{k+1} - cu_i^{k+1} + \tilde{\chi}_i^k \left(u_i^{k+1} - \frac{f^2}{u_i^{k+1}} + \sigma^2 \right)$$

$$- \alpha_i \Delta u_i^{k+1} + \frac{\alpha_i}{2} \frac{|\nabla u_i^{k+1}|^2}{u_i^{k+1}} = 0,$$

with Neumann boundary conditions. Consequently, a semi-implicit approach results in the iteration sequence (80) with

$$r^n = \tilde{\chi}_i^k \left(1 + \frac{\sigma^2}{(u_i^{k+1})^n} \right) \quad \text{and} \quad s^n = \tilde{\chi}_i^k \frac{f^2}{(u_i^{k+1})^n}.$$

5.4 Numerical Realization of the Segmentation Step

After discussing the numerical realization of the denoising problems (63a) and (63b) we now concentrate on the segmentation problem (63c) in the proposed alternating minimization strategy. In this context we are interested in computing the current segmentation indicated by χ^{k+1} using the updated denoised images u_b^{k+1} and u_t^{k+1}.

The standard approaches to solve geometric problems are active contour models such as snakes initially proposed by Kass, Witkin and Terzopoulos in [44] or level set methods introduced by Osher and Sethian in [60]. Although these models have attracted strong attention in the past, there are several drawbacks leading to complications in the computation of segmentation results. For example, the explicit curve representation of snake models do not allow changes in curve topology or the level set methods require an expensive re-initialization of the level set function during the evolution process (cf. e.g., [50, 59]). However, the main drawback of these methods is the non-convexity of the associated energy functionals and consequently the existence of local minima leading to unsatisfactory results with wrong scales of details.

Recently, several globally convex segmentation models have been proposed in [12, 22, 39] to overcome the fundamental issue of existence of local minima. The main idea of all these approaches is based on the unification of image segmentation and image denoising tasks into a global minimization framework. In our work we follow the idea found in [16], where a relation between the well known Rudin-Osher-Fatemi (ROF) model [66] and the minimal surface problem is presented. In order to do so we recall this relation in the following lemma and note that the ROF model always admits a unique solution, since the associated energy functional is strictly convex.

Lemma 11 *Let $\beta > 0$ be a fixed parameter, $g \in L^2(\Omega)$, and $\hat{u}$ the unique solution of the ROF minimization problem*

$$\min_{u \in BV(\Omega)} \frac{1}{2} \int_\Omega (u - g)^2 dx + \beta |u|_{BV(\Omega)}. \tag{84}$$

Then, for almost every $t \in \mathbb{R}$, the indicator function

$$\hat{\chi}(x) = \begin{cases} 1, & \text{if } \hat{u}(x) > t, \\ 0, & \text{else,} \end{cases} \tag{85}$$

is a solution of the minimal surface problem

$$\min_{\chi \in BV(\Omega; \{0,1\})} \int_\Omega \chi(x)(t - g(x)) dx + \beta |\chi|_{BV(\Omega)}. \tag{86}$$

In particular, for all t but a countable set, the solution of the problem (86) is even unique.

Using Lemma 11 we are capable to translate the segmentation problem (63c) to a standard ROF denoising problem. We can observe that the problem (63c) corresponds to the minimal surface problem (86) by setting

$$t = 0 \quad \text{and} \quad g = D_t(f, u_t^{k+1}) - D_b(f, u_b^{k+1}). \tag{87}$$

Therefore, the solution χ^{k+1} of the segmentation step (63c) can now be computed by simple thresholding of the form (85) with $t = 0$, where $\hat{u}$ is the solution of the ROF problem (84) with g specified in (87).

In conclusion, the ROF model (84) is a well understood and intensively studied variational problem in mathematical image processing. Hence, a variety of numerical schemes have been already proposed in literature in order to solve this problem. Exemplarily, we refer to the projected gradient descent algorithm of Chambolle in [17], a nonlinear primal-dual method of Chan, Golub and Mulet in [21], the split Bregman algorithm of Goldstein and Osher in [38], and some first-order algorithms in [8, 18].

6 Experimental Results on Synthetic Data

In this section we validate the proposed variational segmentation framework on synthetic data and focus on the choice of an appropriate data fidelity term and the selection of regularization functionals. To fully understand the influence of these variables we propose different experimental settings in the following, which are constructed to highlight observed characteristics of our segmentation framework. In order to segment the images discussed below, we use the alternating minimization scheme given in (63a)–(63c), for which 25 iteration steps are sufficient in most cases. To terminate the inner iteration loops discussed in Sects. 5.2 and 5.3 we choose $\epsilon = 10^{-6}$ in (69).
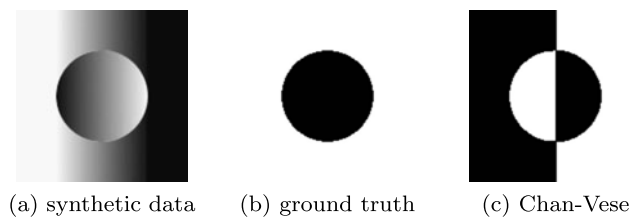
(a) synthetic data   (b) ground truth   (c) Chan-Vese

Fig. 2 Comparison of Chan-Vese segmentation result with ground truth on synthetic data

6.1 Impact of Regularization Term

First, we investigate the impact of the selected regularization term (cf. Sect. 3.4) on the results of segmentation and especially on the estimated approximations to the original object data. For this purpose we use an image of size 150×150 pixels scaled to $[0, 1]$ with a simple object structure illustrated in Fig. 2(a). In this image we put inhomogeneities in Ω_t and Ω_b covering the full range of intensities, such that both regions have the same mean value. The challenge of this data is the occurence of the same grayscale values in both regions with strong intensity changes at the border of the object structure. Apparently, this situation is quite untypical in real-life applications. However, it illustrates the limits of the proposed segmentation framework.

Due to this constructed properties it is obvious that a regularization functional enforcing piecewise constant functions is not feasible in this situation, as can be seen in Fig. 2(c) for the popular Chan-Vese algorithm from Sect. 4.2. This motivates the investigation of the squared H^1-seminorm and the Fisher information regularization discussed in Sect. 3.4. For this purpose we concentrate on data which are perturbed by Poisson noise (cf. Sect. 2.2.2) as illustrated in Fig. 3(b). In order to guarantee identical preconditions we keep $\beta = 0.4$ fix for both experiments.

For the squared H^1-seminorm the regularization parameters are chosen as $\alpha_b = 7$ and $\alpha_t = 540$ in order to achieve a satisfying segmentation result illustrated in Fig. 3(c). The strong difference in the absolute values of α_b and α_t is caused by the extraordinary form of the given data. However, as a consequence of this choice we obtain an oversmoothing of the target region and on the other hand a low regularization in the background region as presented in Fig. 3(e). The same observations can also be made for the Fisher information regularization, where the parameters are chosen as $\alpha_b = 20$ and $\alpha_t = 30$. The corresponding segmentation and approximation results are illustrated in Figs. 3(d) and 3(f), respectively. However, in order to demonstrate the difference of approximations in Figs. 3(e) and 3(f) we plot the intensity values of these results in Fig. 4 along the illustrated profile lines. As we can see the Fisher information result approximates the target object significantly better than the squared H^1-seminorm result.
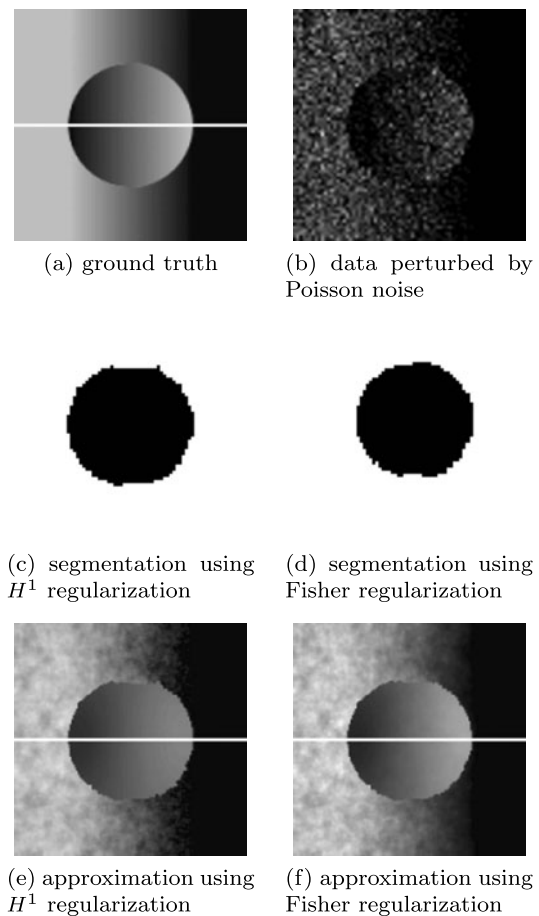

(a) ground truth   (b) data perturbed by Poisson noise

(c) segmentation using H^1 regularization   (d) segmentation using Fisher regularization

(e) approximation using H^1 regularization   (f) approximation using Fisher regularization

Fig. 3 Comparison of segmentation and approximation results for synthetic data biased by Poisson noise using the squared H^1-seminorm and Fisher information regularization. The grayscale values in (**a**), (**e**) and (**f**) are scaled to a uniform interval and the *horizontal white lines* represent the position of the profile *lines* plotted in Fig. 4
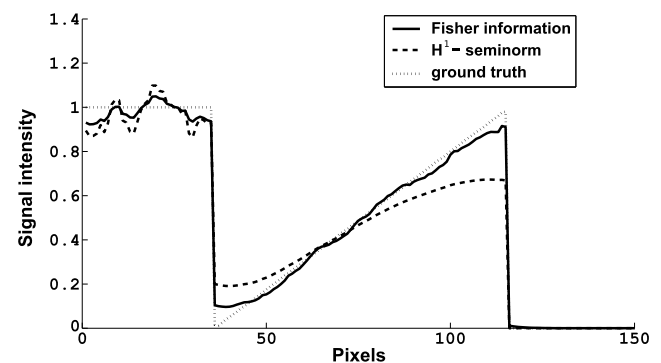

Fig. 4 Profile *lines* of the squared H^1-seminorm and Fisher information approximations from Fig. 3

During our experiments described above we also observed a possible convergence of our algorithm against the local minima presented in Fig. 2(c). For the Fisher information regularization we illustrate such segmentation result in Fig. 5(c), where the regularization parameters are chosen
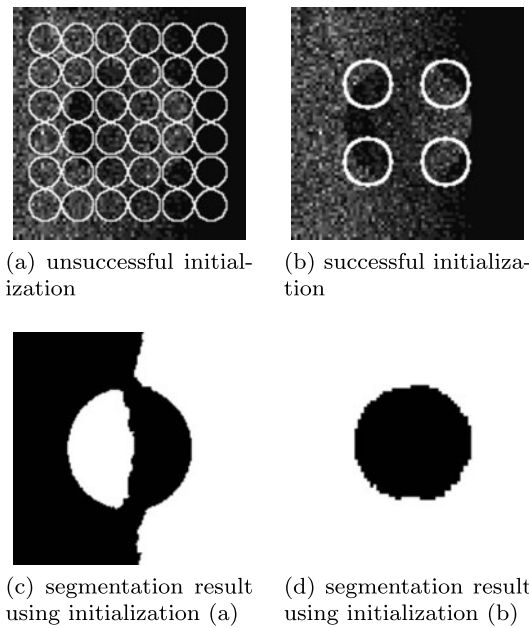
(a) unsuccessful initial-
ization

(b) successful initializa-
tion

(c) segmentation result
using initialization (a)

(d) segmentation result
using initialization (b)

Fig. 5 Comparison of segmentation results for synthetic data biased by Poisson noise using the Fisher information regularization and different initializations

(a) synthetic data

(b) ground truth

Fig. 6 Synthetic data with anatomical structures of different size

as described above and the initialization given in Fig. 5(a) are used. This unsatisfactory result can be avoided as illustrated in Fig. 5(d) using another initialization as presented in Fig. 5(b). The reason for this behavior is the non-convexity of the whole energy functional given in (19) such that a convergence against a global minimum cannot be guaranteed, although each subproblem in (63a)–(63c) has a global solution. Hence, we state that the convergence of the presented method against an optimal solution depends not only on the specified regularization parameters but also on a suitable initialization.

6.2 Impact of the Data Fidelity Term

To evaluate the importance of a correct noise model in automated image segmentation we investigate images perturbed by physical noise forms described in Sect. 2.2, i.e., additive Gaussian noise, Poisson noise, and multiplicative speckle noise. For the sake of simplicity we perform this experiment on piecewise constant images, since this case is known from the popular Chan-Vese formulation [20] and we can neglect the regularization parameters α_b and α_t. For the segmentation we use the generalized Chan-Vese two-phase segmentation model proposed in Sect. 4.3, i.e., we only separate regions-of-interest from the background.

We choose the objects to be segmented with respect to typical segmentation tasks from biomedical imaging. Often, only one major anatomical region-of-interest has to be segmented, e.g., the left ventricle of the heart in echocardiographic examinations [57, 76] or [^{18}F]FDG uptake studies
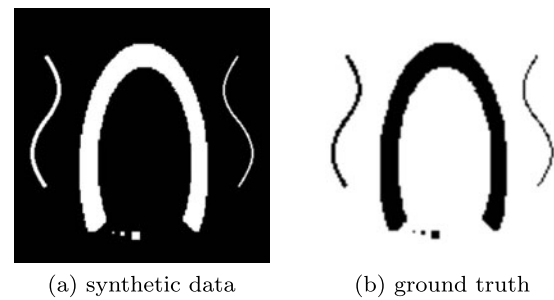
in positron emission tomography [62]. Furthermore, it is desirable to preserve as many image details as possible during the process of segmentation. Especially in tumor imaging [75] small lesions in a size of only few pixels can easily be overseen, due to a loss of details by too intense regularization. This leads to severe consequences if not taken into account and hence it is important to preserve details of small image regions.

Figure 6(a) shows our experimental data without perturbation of noise and the ground truth segmentation in Fig. 6(b). In the image center we place an approximate shape of the left ventricle of the human heart as it would be imaged in an apical four-chamber view in echocardiography. Below this major structure we put three small squares with sizes of 1, 2, and 4 pixels to simulate minor structures, such as small lesions, which we want to preserve during the process of automated image segmentation. On the left and right side of our heart phantom we set two curved lines with a diameter of 1 and 2 pixels to simulate vessel-like structures, which play an important role in perfusion studies of different organs [33, 82], e.g., liver veins or coronary arteries of the heart.

To validate the impact of the data fidelity term we bias the image in Fig. 6(a) with synthetic noise and try to find the optimal value of the regularization parameter β. This optimization was done with respect to the following two criterions,

- Segmentation of the main structure without noise artifacts.
- Preservation of small anatomical structures without loss of details.

Naturally, it is hard to fulfill both criterions simultaneously, since there is a trade-off between noise-free segmentation results and a detailed partition of the image. For our synthetic images we look for the highest possible value of β, which preserves as many small structures as possible, and on the other hand for the lowest possible value of β that ensures a complete segmentation of the main structure without noise-induced artifacts. These two limiting cases are the foundation for our observations in the following experiment.
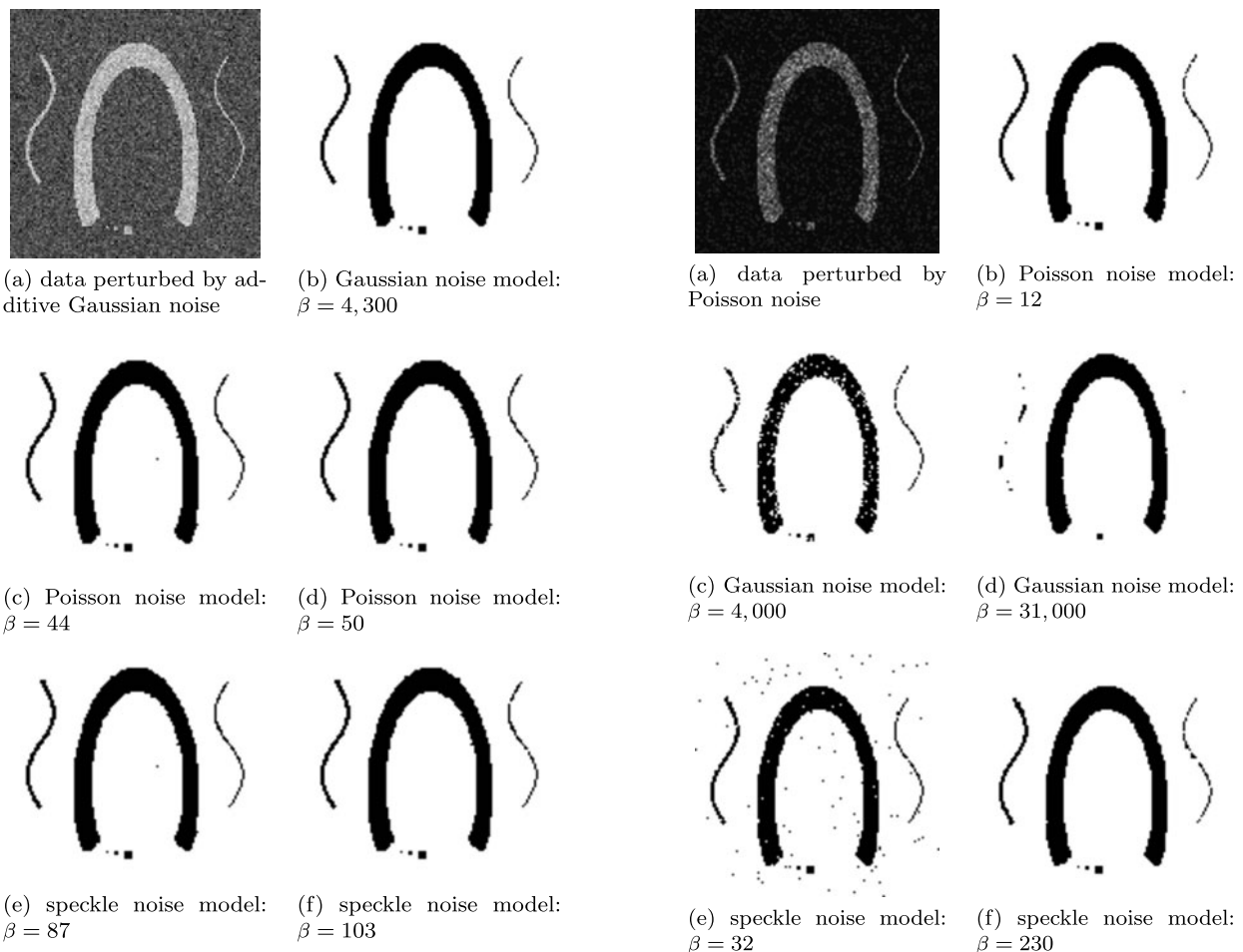
(a) data perturbed by additive Gaussian noise

(b) Gaussian noise model: $\beta = 4,300$

(c) Poisson noise model: $\beta = 44$

(d) Poisson noise model: $\beta = 50$

(e) speckle noise model: $\beta = 87$

(f) speckle noise model: $\beta = 103$

Fig. 7 Comparison of different data fidelity models on synthetic data of simulated anatomical structures biased by additive Gaussian noise

(a) data perturbed by Poisson noise

(b) Poisson noise model: $\beta = 12$

(c) Gaussian noise model: $\beta = 4,000$

(d) Gaussian noise model: $\beta = 31,000$

(e) speckle noise model: $\beta = 32$

(f) speckle noise model: $\beta = 230$

Fig. 8 Comparison of different data fidelity models on synthetic data of simulated anatomical structures biased by Poisson noise

First, we perturb the data with additive Gaussian noise as illustrated in Fig. 7(a). In Fig. 7(b) the data fidelity term for additive Gaussian noise (cf. (14)) produces a satisfying segmentation result, which fulfills both criterions discussed above. For this case we observe very similar results for the Poisson and speckle data fidelity term as illustrated in Figs. 7(c), 7(d) and 7(e), 7(f), respectively.

In the next experiment we perturb the synthetic data with Poisson noise as presented in Fig. 8(a). For this image we state that the Poisson data fidelity term in (15) is an appropriate choice as can be seen in Fig. 8(b). In Fig. 8(c) and 8(d) we test the additive Gaussian noise fidelity term and choose the regularization parameter β according to the criterions discussed above. In order to preserve all small structures in Fig. 8(c) we have to accept a significant amount of noise artifacts within the main structure. On the other hand, we lose almost all small structures in Fig. 8(d) by choosing β high enough to guarantee a segmentation of the center object without pertubation of noise. In this case the trade-off between smooth segmentations and the loss of details becomes obvious. Figure 8(e) and 8(f) show the results in
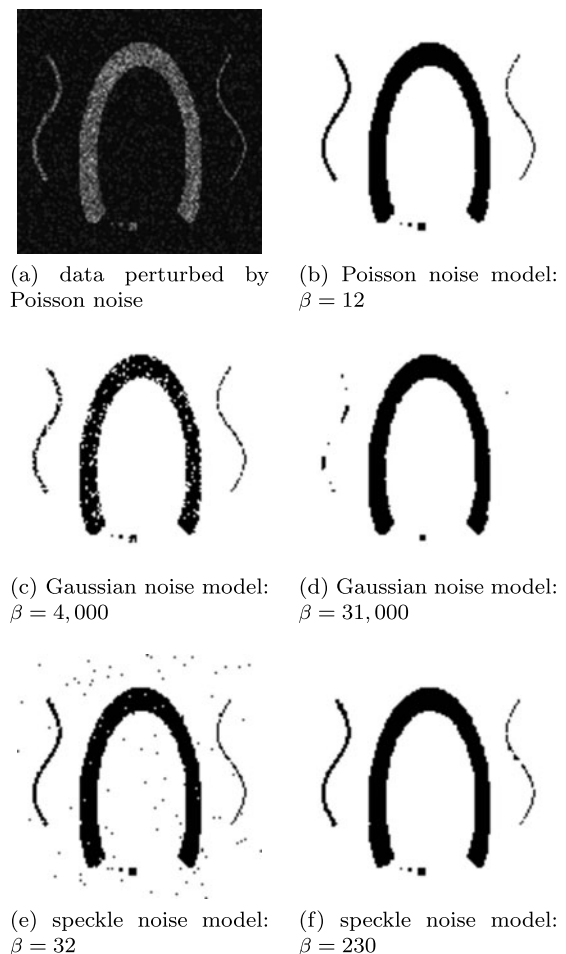
the case of the speckle data fidelity term. Since we do not observe any satisfying results for different values of β, we show two representative segmentations for this model. Compared to the results in the case of additive Gaussian noise, the choice of the appropriate noise model has a significantly higher impact on the segmentation results in the presence of Poisson noise.

Finally, we investigate the case of data biased by multiplicative speckle noise as shown in Fig. 9(a). The segmentation result for the corresponding speckle data fidelity term (cf. (16)) is presented in Fig. 9(b). Again, we observe that we are able to satisfy both segmentation criterions when choosing the correct noise model. In Fig. 9(c) and 9(d) the segmentation results for the data fidelity term of additive Gaussian noise are presented. As for the results in Fig. 8(c) and 8(d), the Gaussian model is inappropriate in the presence of multiplicative speckle noise. For the Poisson fidelity term we observe in Fig. 9(e) and 9(f) similar effects compared to the Gaussian case, but with a better trade-off between smooth segmentation results and the preservation of small image structures.
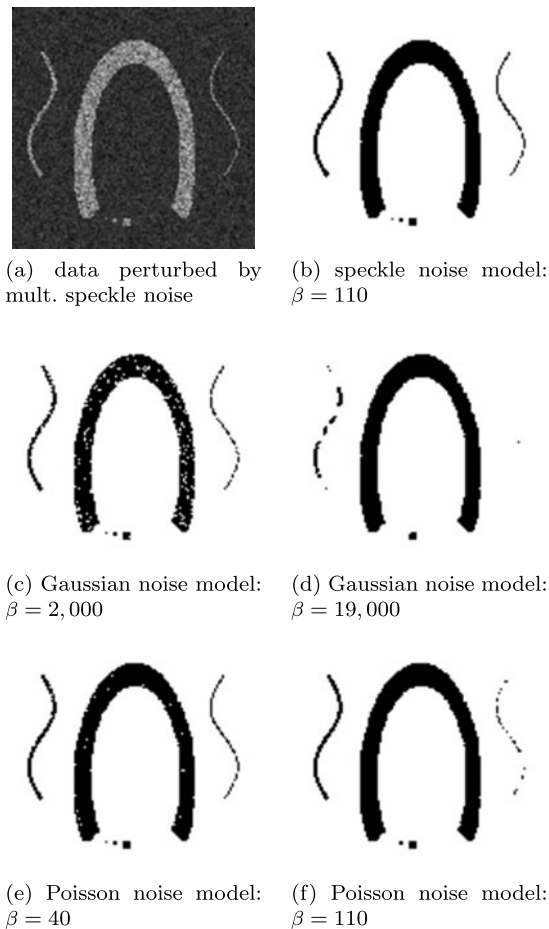
(a) data perturbed by mult. speckle noise

(b) speckle noise model: $\beta = 110$

(c) Gaussian noise model: $\beta = 2,000$

(d) Gaussian noise model: $\beta = 19,000$

(e) Poisson noise model: $\beta = 40$

(f) Poisson noise model: $\beta = 110$

Fig. 9 Comparison of different data fidelity models on synthetic data of simulated anatomical structures biased by multiplicative speckle noise

In conclusion, we emphasize that the incorporation of physical noise modeling for given data has a significant impact on segmentation results and leads to improved accuracy in applications dealing with non-Gaussian noise.

7 Experimental Results on Real Data

In this section we use real data from biomedical imaging and investigate the performance of the proposed variational segmentation framework in real segmentation tasks. Therefore, we present images from two imaging modalities dealing with the non-Gaussian noise models discussed in Sect. 2.2.

7.1 Positron Emission Tomography

To test the variational segmentation framework on data which are heavily perturbed by noise we take images from positron emission tomography (PET). PET belongs to the field of molecular imaging in which a specific radioactive tracer is injected into blood circulation and its binding to certain molecules is studied [83]. One possible tracer is, e.g.,

(a) EM reconstruction

(b) region-of-interest

(c) Fisher approximation

(d) segmentation contour using Fisher regularization (Poisson model)
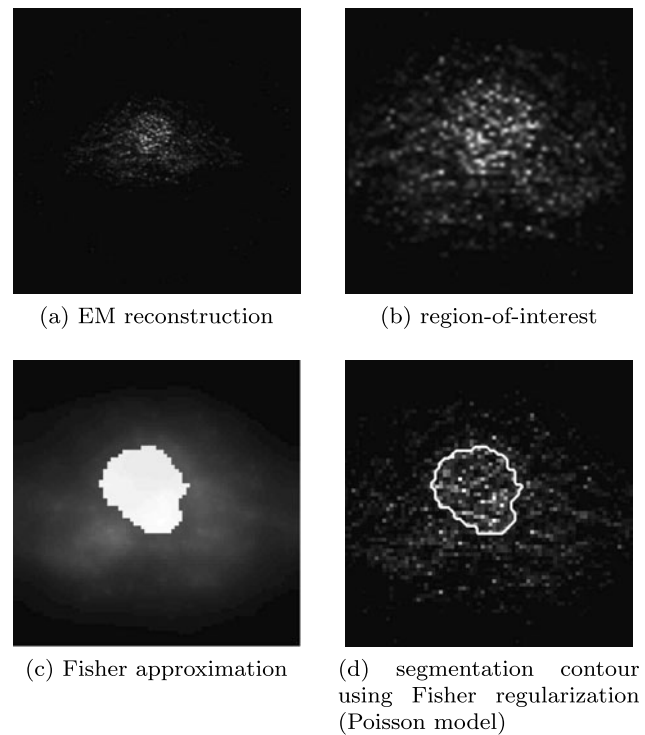
Fig. 10 Segmentation result for reconstruction of cardiac $H_2[^{15}O]$ PET data

$H_2[^{15}O]$ (radioactive-labeled water), which can be used for quantification of the myocardial blood flow [71]. However, this quantification needs a segmentation of myocardial tissue [10, 71], which is extremely difficult to realize due to a very low signal-to-noise (SNR) ratio of $H_2[^{15}O]$ data.

Figure 10(a) shows a slice of a reconstructed $H_2[^{15}O]$ study using the expectation maximization (EM) algorithm [72] with signal intensities in the range of up to 2,000,000 counts. The data were captured in the moment, when the tracer flooded into the left ventricle of the human heart. This very short interval leads to a high level of Poisson noise in the data and thus in the corresponding EM reconstructions, causing a challenging task for most segmentation algorithms. Based on our observations in the last section we choose the Poisson data fidelity term (cf. (15)) and use Fisher information as regularization functional. For our experiment we focus on the region-of-interest shown in Fig. 10(b), which corresponds to the area around the myocardial tissue. As regularization parameters we determine $\alpha_t = 140$, $\alpha_b = 20$, and $\beta = 380,000$ (low SNR).

In Fig. 10(c) and 10(d) we show the approximation and segmentation result in the region-of-interest, respectively. The segmented region corresponds to the myocardium including the left and right ventricle. As can be seen, the proposed segmentation framework can segment the myocardium in data with very low SNR. However, we were not able to separate left and right ventricle as it would be

needed for myocardial perfusion quantification. The reason for this is the used two-phase formulation in (19), which is not suitable to segment different uptake levels and additional background activity.

7.2 Medical Ultrasound Imaging

The last case we want to investigate is medical ultrasound data. As proposed in [45, 52] the physical noise occurring in this imaging modality can be modeled by multiplicative speckle noise as presented in Sect. 2.2.3 and hence the speckle data fidelity term (cf. (16)) is an appropriate choice in this context. In the following we concentrate on the task of blood vessel segmentation, which can be used in perfusion studies of different organs, e.g., the myocardium [33, 82].

In Fig. 11(a) we illustrate a section of an ultrasound B-mode image containing parts of a human liver. The darker regions within these data represent the blood vessels to be segmented. In Figs. 11(b)–11(d) we present three segmentation results for $\beta \in \{1000, 2000, 5000\}$ using the classical Chan-Vese model (additive Gaussian noise) from Sect. 4.2. As can be seen we were not able to segment the small structures of the blood vessels accurately without over-segmenting the data. We compare the best segmentation result for the classical CV approach with $\beta = 2000$ against the corresponding extension to the speckle noise model presented in Sect. 4.3. As can be seen in Fig. 11(e) by exchanging only the data fidelity term we can observe a significant improvement in the segmentation accuracy and suppression of noise artifacts. Although, the result in Fig. 11(e) appears to segment the blood vessels within the image correctly, we can identify some small structures that have been left out. In order to include these microvessels we exchanged the regularization term to the Fisher information for target and background region and used the regularization parameters $\alpha_t = 15$, $\alpha_b = 25$, and $\beta = 25$. We could not achieve any meaningful segmentation result using the squared H^1-seminorm regularization due to oversmoothing of small structures.

As illustrated in Fig. 11(f) the proposed segmentation framework is able to segment even smallest structures within the given data. Especially in low contrast regions, e.g., the lower-left corner of the image, we observe satisfying segmentation results. Figure 11(g) shows the corresponding approximation of the data for the Fisher information regularization and illustrates a non-constant approximation of the target region while preserving edges at the blood vessels within the image.
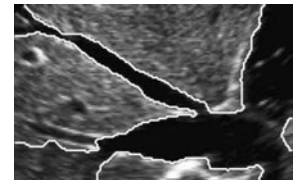
8 Conclusion

In this work we proposed a variational segmentation framework for different physical noise models and analyzed the
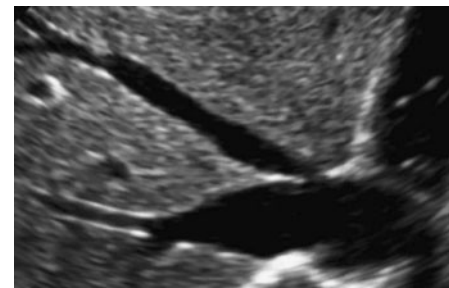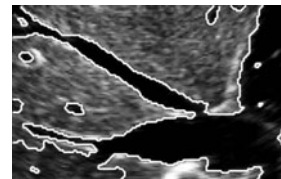

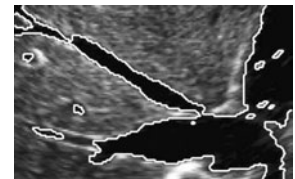(a) US data of human liver
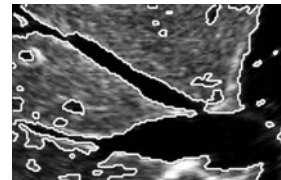
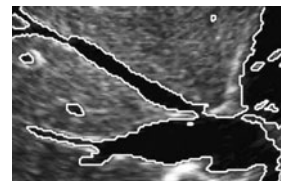(b) segmentation contour using classical CV ($\beta = 1000$)

(c) segmentation contour using classical CV ($\beta = 5000$)

(d) segmentation contour using classical CV ($\beta = 2000$)

(e) segmentation contour using extension of CV (speckle model)

(f) segmentation contour using Fisher regularization (speckle model)

(g) Fisher approximation (speckle model)

Fig. 11 Segmentation results for ultrasound (US) data from an examination of a human liver comparing classical and extended Chan-Vese (CV) method for speckle noise, and Fisher information for speckle noise

corresponding mathematical theory. In particular, we investigated a selected variety of regularization terms and noise models, i.e., additive Gaussian noise, Poisson noise, and multiplicative speckle noise, for automated image segmentation. Experimental results on synthetic and real data show the necessity to adapt an algorithm to present conditions, e.g., incorporating a-priori knowledge about the solution.

We plan to translate the theoretical fundament provided in this work to more real world applications and to validate the presented variational segmentation framework on a variety of realistic data. Furthermore, we plan to explore the incorporation of more physical noise models, e.g., Rayleigh- or Gamma-perturbed data.

References

1. Acar, R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. Inverse Probl. **10**, 1217–1229 (1994)

2. Adams, R.A.: Sobolev Spaces. Pure and Applied Mathematics, vol. 65. Academic Press, New York (1975)

3. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics, vol. 140. Elsevier, Amsterdam (2003)

4. Ambrosio, L., Tortorelli, V.M.: Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. Commun. Pure Appl. Math. **43**, 999–1036 (1990)

5. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. Oxford University Press, Oxford (2000)

6. Aubert, G., Aujol, J.-F.: A variational approach to removing multiplicative noise. SIAM J. Appl. Math. **68**, 925–946 (2008)

7. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, 2nd edn. Applied Mathematical Sciences, vol. 147. Springer, Berlin (2006)

8. Aujol, J.-F.: Some first-order algorithms for total variation based image restoration. J. Math. Imaging Vis. **34**, 307–327 (2009)

9. Barbu, V., Precupanu, T.: Convexity and Optimization in Banach Spaces. Sijthoff & Noordhoff, Rockville (1978)

10. Benning, M., Kösters, T., Wübbeling, F., Schäfers, K., Burger, M.: A nonlinear variational method for improved quantification of myocardial blood flow using dynamic $H_2^{15}O$ PET. In: Nuclear Science Symposium Conference Record, pp. 4472–4477 (2008)

11. Bertero, M., Lanteri, H., Zanni, L.: Iterative image reconstruction: a point of view. In: Censor, Y., Jiang, M., Louis, A. (eds.) Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT). Publications of the Scuola Normale, CRM Series, vol. 7, pp. 37–63 (2008)

12. Bresson, X., Esedoglu, S., Vandergheynst, P., Thiran, J.-P., Osher, S.: Fast global minimization of the active contour/snake model. J. Math. Imaging Vis. **28**, 151–167 (2007)

13. Brox, T., Weickert, J.: Level set segmentation with multiple regions. IEEE Trans. Image Process. **15**, 3213–3218 (2006)

14. Burger, M., Franek, M., Schönlieb, C.-B.: Regularized regression and density estimation based on optimal transport. Appl. Math. Res. Express **2012** (2012), 45 pp.

15. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vis. **22**, 61–79 (1997)

16. Caselles, V., Chambolle, A., Novaga, M.: The discontinuity set of solutions of the TV denoising problem and some extensions. Multiscale Model. Simul. **6**, 879–894 (2007)

17. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**, 89–97 (2004)

18. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**, 120–145 (2011)

19. Chan, T.F., Shen, J.: Image Processing and Analysis: Variational, PDE, Wavelet and Stochastic Methods. SIAM, Philadelphia (2005)

20. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**, 266–277 (2001)

21. Chan, T.F., Golub, G.H., Mulet, P.: A nonlinear primal-dual method for total variation-based image restoration. SIAM J. Sci. Comput. **20**, 1964–1977 (1999)

22. Chan, T.F., Esedoglu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM J. Appl. Math. **66**, 1632–1648 (2006)

23. Chesnaud, C., Réfrégier, P., Boulet, V.: Statistical region snake-based segmentation adapted to different physical noise models. IEEE Trans. Pattern Anal. Mach. Intell. **21**, 1145–1157 (1999)

24. Chung, G., Vese, L.A.: Energy minimization based segmentation and denoising using a multilayer level set approach. In: Proceedings of the 5th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition. LNCS, vol. 3757, pp. 439–455. Springer, Berlin (2005)

25. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. Int. J. Comput. Vis. **72**, 195–215 (2007)

26. Cremers, D., Pock, T., Kolev, K., Chambolle, A.: Convex relaxation techniques for segmentation, stereo, and multiview reconstruction. In: Markov Random Fields for Vision and Image Processing. MIT Press, New York (2011)

27. Dey, N., Blanc-Féraud, L., Zimmer, C., Roux, P., Kam, Z., Olivio-Marin, J.-C., Zerubia, J.: 3D microscopy deconvolution using Richardson-Lucy algorithm with total variation regularization. Tech. Rep. 5272, Institut National de Recherche en Informatique et en Automatique (2004)

28. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. Studies in Mathematics and Its Applications, vol. 1. North-Holland, Amsterdam (1976)

29. Elman, H.C., Golub, G.H.: Inexact and preconditioned Uzawa algorithms for saddle point problems. SIAM J. Numer. Anal. **31**, 1645–1661 (1994)

30. Fortin, M., Glowinski, R.: Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems. Studies in Mathematics and Its Applications, vol. 15. Elsevier, Amsterdam (1983)

31. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. J. Appl. Stat. **20**, 25–62 (1993)

32. Geman, S., McClure, D.E.: Bayesian image analysis: an application to single photon emission tomography. In: Statistical Computation Section, American Statistical Association, pp. 12–18 (1985)

33. Ghanem, A., et al.: Triggered replenishment imaging reduces variability of quantitative myocardial contrast echocardiography and allows assessment of myocardial blood flow reserve. Echocardiography **24**, 149–158 (2007)

34. Gianazza, U., Savaré, G., Toscani, G.: The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. Arch. Ration. Mech. Anal. **194**, 133–220 (2009)

35. Giusti, E.: Minimal Surfaces and Functions of Bounded Variation. Monographs in Mathematics, vol. 80. Birkhäuser, Basel (1984)

36. Glowinski, R., Le Tallec, P.: Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. Studies in Applied Mathematics, vol. 9. SIAM, Philadelphia (1989)

37. Goldluecke, B., Cremers, D.: Convex relaxation for multilabel problems with product label spaces. In: Proceedings of the 11th European Conference on Computer Vision. LNCS, vol. 6315, pp. 225–238. Springer, Berlin (2010)

38. Goldstein, T., Osher, S.: The split Bregman method for L^1-regularized problems. SIAM J. Imaging Sci. **2**, 323–343 (2009)

39. Goldstein, T., Bresson, X., Osher, S.: Geometric applications of the split Bregman method: segmentation and surface reconstruction. J. Sci. Comput. **45**, 272–293 (2010)

40. Helin, T., Lassas, M.: Hierarchical models in statistical inverse problems and the Mumford-Shah functional. Inverse Probl. **27**, 015008 (2011). 32 pp.

41. Hell, S.W.: Toward fluorescence nanoscopy. Nat. Biotechnol. **21**, 1347–1355 (2003)

42. Ito, K., Kunisch, K.: Lagrange Multiplier Approach to Variational Problems and Applications. Advances in Design and Control, vol. 15. SIAM, Philadelphia (2008)

43. Jin, Z., Yang, X.: A variational model to remove the multiplicative noise in ultrasound images. J. Math. Imaging Vis. **39**, 62–74 (2011)

44. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **1**, 321–331 (1988)

45. Krissian, K., Kikinis, R., Westin, C.-F., Vosburgh, K.: Speckle-constrained filtering of ultrasound images. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 547–552 (2005)

46. Lantéri, H., Theys, C.: Restoration of astrophysical images—the case of Poisson data with additive Gaussian noise. EURASIP J. Appl. Signal Process. **15**, 2500–2513 (2005)

47. Le, T., Chartrand, R., Asaki, T.J.: A variational approach to reconstructing images corrupted by Poisson noise. J. Math. Imaging Vis. **27**, 257–263 (2007)

48. Lellmann, J., Becker, F., Schnörr, C.: Convex optimization for multi-class image labeling with a novel family of total variation based regularizers. In: Proceedings of the IEEE 12th International Conference on Computer Vision, pp. 646–653 (2009)

49. Lellmann, J., Lenzen, F., Schnörr, C.: Optimality bounds for a variational relaxation of the image partitioning problem. In: Proceedings of the 8th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition. LNCS, vol. 6819, pp. 132–146. Springer, Berlin (2011)

50. Li, C., Xu, C., Gui, C., Fox, M.D.: Level set evolution without re-initialization: a new variational formulation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 430–436 (2005)

51. Llacer, J., Núñez, J.: Iterative maximum likelihood and Bayesian algorithms for image reconstruction in astronomy. In: White, R.L., Allen, R.J. (eds.) The Restoration of Hubble Space Telescope Images, pp. 62–69. The Space Telescope Science Institute, Baltimore (1990)

52. Loupas, T., McDicken, W.N., Allan, P.L.: An adaptive weighted median filter for speckle suppression in medical ultrasonic images. IEEE Trans. Circuits Syst. **36**, 129–135 (1989)

53. Martin, P., Réfrégier, P., Goudail, F., Guérault, F.: Influence of the noise model on level set active contour segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 799–803 (2004)

54. Megginson, R.E.: An Introduction to Banach Space Theory. Graduate Texts in Mathematics, vol. 183. Springer, Berlin (1998)

55. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. **42**, 577–685 (1989)

56. Nečas, J.: Les Méthodes Directes en Théorie des Équations Elliptiques. Academia/Masson et Cie, Prague/Paris (1967)

57. Noble, J.A., Boukerroui, D.: Ultrasound image segmentation: a survey. IEEE Trans. Med. Imaging **25**, 987–1010 (2006)

58. Obereder, A., Scherzer, O., Kovac, A.: Bivariate density estimation using BV regularisation. Comput. Stat. Data Anal. **51**, 5622–5634 (2007)

59. Osher, S., Fedkiw, R.: Level Set Methods and Dynamic Implicit Surfaces. Applied Mathematical Sciences, vol. 153. Springer, Berlin (2003)

60. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. J. Comput. Phys. **79**, 12–49 (1988)

61. Paragois, N., Deriche, R.: Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision. J. Vis. Commun. Image Represent. **13**, 249–268 (2002)

62. Pirich, C., Schwaiger, M.: The clinical role of positron emission tomography in management of the cardiac patient. Port. J. Cardiol. **19**(Suppl 1), 89–100 (2000)

63. Pock, T., Chambolle, A., Cremers, D., Bischof, H.: A convex relaxation approach for computing minimal partitions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 810–817 (2009)

64. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: Global solutions of variational models with convex regularization. SIAM J. Imaging Sci. **3**, 1122–1145 (2010)

65. Resmerita, E., Anderssen, R.S.: Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems. Math. Methods Appl. Sci. **30**, 1527–1544 (2007)

66. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)

67. Rudin, L., Lions, P.-L., Osher, S.: Multiplicative denoising and deblurring: theory and algorithms. In: Geometric Level Set Methods in Imaging, Vision, and Graphics, pp. 103–119. Springer, Berlin (2003)

68. Sarti, A., Corsi, C., Mazzini, E., Lamberti, C.: Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **52**, 947–960 (2005)

69. Sawatzky, A.: (Nonlocal) Total Variation in Medical Imaging. PhD thesis, University of Münster (2011). CAM Report 11-47, UCLA

70. Sawatzky, A., Brune, C., Müller, J., Burger, M.: Total variation processing of images with Poisson statistics. In: Jiang, X., Petkov, N. (eds.) Computer Analysis of Images and Patterns. Lecture Notes in Computer Science, vol. 5702, pp. 533–540 (2009)

71. Schäfers, K.P., et al.: Absolute quantification of myocardial blood flow with $H_2^{15}O$ and 3-dimensional PET: an experimental validation. J. Nucl. Med. **43**, 1031–1040 (2002)

72. Shepp, L.A., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imaging **1**, 113–122 (1982)

73. Smereka, P.: Semi-implicit level set methods for curvature and surface diffusion motion. J. Sci. Comput. **19**, 439–456 (2003)

74. Snyder, D.L., Hammoud, A.M., White, R.L.: Image recovery from data acquired with a charge-coupled-device camera. J. Opt. Soc. Am. A, Opt. Image Sci. Vis. **10**, 1014–1023 (1993)

75. Soret, M., Bacharach, S.L., Buvat, I.: Partial-volume effect in PET tumor imaging. J. Nucl. Med. **48**, 932–945 (2007)

76. Stypmann, J., et al.: Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L. Proc. Natl. Acad. Sci. USA **99**, 6234–6239 (2002)

77. Tur, M., Chin, K.C., Goodman, J.W.: When is speckle noise multiplicative? Appl. Opt. **21**, 1157–1159 (1982)

78. Vardi, Y., Shepp, L.A., Kaufman, L.: A statistical model for positron emission tomography. J. Am. Stat. Assoc. **80**, 8–20 (1985)

79. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the Mumford and Shah model. Int. J. Comput. Vis. **50**, 271–293 (2002)

80. Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, New York (2003)

81. Vovk, U., Pernuš, F., Likar, B.: A review of methods for correction of intensity inhomogeneity in MRI. IEEE Trans. Med. Imaging **26**, 405–421 (2007)

82. Wellnhofer, E., et al.: Angiographic assessment of cardiac allograft vasculopathy: results of a consensus conference of the task force for thoracic organ transplantation of the German cardiac society. Transpl. Int. **23**, 1094–1104 (2010)

83. Wernick, M.N., Aarsvold, J.N. (eds.): Emission Tomography: The Fundamentals of PET and SPECT. Elsevier, Amsterdam (2004)

84. Wirtz, D.: SEGMEDIX: Development and application of a medical image segmentation framework. Master's thesis, University of Münster (2009) www.agh.ians.uni-stuttgart.de/uploads/media/DA_Wirtz.pdf

85. Xiao, G., Brady, M., Noble, J.A., Zhang, Y.: Segmentation of ultrasound B-mode images with intensity inhomogeneity correction. IEEE Trans. Med. Imaging **21**, 48–57 (2002)

Alex Sawatzky studied Mathematics at the University of Münster and received his PhD in 2011. Since 2012 he is a Postdoctoral Research Scholar at the Department of Biomedical Engineering at Washington University in St. Louis. His research interests are mathematical image processing and inverse problems in biomedical imaging.

Daniel Tenbrinck graduated computer science in 2009 at the University of Münster, Germany. He is currently a PhD student in the Computer Vision and Pattern Recognition Group at the Department of Mathematics and Computer Science in Münster, Germany. His research interests include computer vision, medical imaging, and variational methods.

Xiaoyi Jiang studied Computer Science at Peking University and received his PhD and Venia Docendi (Habilitation) degree from University of Bern, Switzerland. He was an associate professor at Technical University of Berlin. Since 2002 he is a full professor of Computer Science at University of Münster, Germany. Currently, he is Editor-in-Chief of Int. Journal of Pattern Recognition and Artificial Intelligence and also serves on the editorial board of several other journals including Pattern Recognition, IEEE TSMC—Part B, and Chinese Science Bulletin. He is Senior Member of IEEE and Fellow of IAPR.

Martin Burger studied Mathematics at Johannes Kepler University Linz and University of Milano. In 2000 he received his PhD and 2005 the Venia Docendi. He was assistant professor in Linz and UCLA and now is full professor in applied mathematics at the Westfälische Wilhelms-University in Münster since 2006. His research interests are mathematical image processing, inverse problems, and mathematical modelling in the life sciences. He is editorial board member of Inverse Problems, Communications in Mathematical Sciences, and of the Springer Lecture Notes in Mathematical Modelling in the Life Sciences. In 2009 he received the Calderon prize for distinguished contributions to the field of inverse problems. Since 2012 he is founding and executive board member of the cluster of excellence "Cells in Motion" in Münster.