

Generalized TV- ℓ_p Structured Priors for Bayesian T_1 Mapping

Disi Lin ¹, Martin Berggren ¹, Tommy Löfstedt ¹

¹ Department of Computing Science, Umeå University, Sweden

Abstract

We propose an extended family of structured spatial priors that incorporates the total variation (TV) function with ℓ_p norms. The prior is proven to be proper and incorporated into a Bayesian regression framework to enable uncertainty quantification in T_1 mapping, with posterior inference performed using the No-U-Turn Sampler (NUTS). This TV- ℓ_p construction is proven to constitute a well-defined family of prior distributions, and it naturally enforces spatial consistency and smooth variations in the estimated parameter maps. The method was evaluated in comparison to maximum-likelihood estimation and several Bayesian alternative priors based on the uniform, Gamma, and bounded TV priors. The evaluation includes experiments on synthetic brain and cardiac T_1 mapping datasets, as well as a real in-vivo breast T_1 mapping dataset. The results show that the TV- ℓ_p prior yields more concentrated posterior densities, indicating reduced uncertainty. It also consistently achieves lower variance and smaller (negative) bias, leading to more reliable estimates. Overall, embedding a TV-based structured penalty along with ℓ_p norms in a prior in a Bayesian model improves spatial coherence in T_1 maps and enhances uncertainty quantification, offering a robust approach for T_1 mapping with uncertainties. Our code is available at <https://github.com/Disi2022/Proper-TV-Structured-Priors>.

Keywords

Bayesian Inference, T_1 Mapping, Uncertainty Quantification, Structured Prior, Total Variation, ℓ_p Norms

Article informations

<https://doi.org/https://doi.org/10.59275/j.me1ba.2026-g41g>

©2026 Lin, Berggren and Löfstedt. License: CC-BY 4.0

Volume 2026, Received: 2026-01, Published 2026-06

Corresponding author: disi.lin@umu.se

Special issue: Uncertainty for Safe Utilization of Machine Learning in Medical Imaging (UNSURE) 2025

Guest editors: Mobarak Hoque, Raghav Mehta, Cheng Ouyang, Chen Qin, Marianne Rakic, Sandy Wells



1. Introduction

Magnetic Resonance Imaging (MRI) is a widely used medical imaging modality to visualize a subject's anatomical structures and physiological processes. For example, cardiac MRI (Rajiah et al., 2023) provides detailed images of the cardiovascular system while breast MRI is used to examine internal breast structures (Wekking et al., 2023). Functional MRI (fMRI) can be used to measure brain activity by detecting changes in cerebral blood flow and oxygenation (DeYoe et al., 1994). MRI techniques play crucial roles in disease diagnosis and treatment. Compared to other imaging modalities, such as ultrasound and computed tomography (CT), MRI provides superior soft tissue contrast and eliminates the risks of exposure to ionizing radiation.

Despite its advantages, conventional MRI primarily provides qualitative images based on contrast differences, which limits its sensitivity to subtle or early pathological changes (Lavrova et al., 2021). Additionally, MR image

intensity values are not standardized and can vary substantially depending on factors such as manufacturer, sequence type, and acquisition parameters (Carré et al., 2020). Quantitative MRI (qMRI) is a variation of MRI that produces measurements in physical units rather than relying on contrast weighting (Granziera et al., 2021). Techniques using qMRI have been developed to quantify biophysical parameters such as relaxation times (T_1 , T_2 , and T_2^*), proton density, diffusion, etc. (Jara et al., 2022). Obtaining parametric maps typically involves two steps: 1) Transform raw k-space data to the image domain, resulting in contrast-weighted images; and 2) fit the parametrically linked contrast-weighted images to a biophysical or MRI signal model, resulting in quantitative parametric maps (Shafieizargar et al., 2023). Overall, qMRI provides a more objective and reproducible alternative to conventional MRI (Hellström, 2025).

In particular, T_1 mapping is a qMRI technique that measures the longitudinal or spin-lattice relaxation time. T_1 mapping has, e.g., been used to uncover pathological

processes in the brain (Gräfe et al., 2021) and myocardial tissue characterization (Becker et al., 2017). Multiple T_1 mapping pulse sequences have been proposed (Tirkes et al., 2019), with inversion recovery (IR) widely regarded as the gold standard due to its accuracy. IR records the recovery of longitudinal magnetization following an inversion pulse (Pykett et al., 1983). However, it requires a long acquisition time, which has prompted the development of various accelerated alternatives. One such alternative is saturation recovery, which replaces the inversion pulse with a saturation pulse. Another prominent strategy is the Look-Locker (LL) method, which samples the recovery curve after inversion (Look and Locker, 1970). Building upon the LL framework, several refined variants have been proposed to improve efficiency and accuracy, including the Modified Look-Locker IR (MOLLI) sequence (Messroghli et al., 2004) and the Modulated Repetition Time Look-Locker (MORTLL) method (Gai and Butman, 2009). Differing in magnetization preparation, the Variable Flip Angle (VFA) technique (Christensen et al., 1974) estimates T_1 from steady-state signals across multiple flip angles instead of recovery curves following inversion or saturation.

However, these T_1 mapping techniques are typically limited by long repetition times and introduce systematic errors in the T_1 estimation (Preibisch and Deichmann, 2009). Undersampling the k-space, to sample below the Nyquist rate, has become the mainstream approach for accelerating MRI acquisition. However, undersampling results in underdetermined equation systems, leading to artifacts when using linear reconstruction methods and lower signal-to-noise ratio in acquired data, which leads to poor reconstructed image quality (Virtue and Lustig, 2017). Regularization, e.g., based on low-rank constraint and/or total variation (TV), has been used to solve such ill-posed inverse problems. Zhang et al. (2015) proposed a local low-rank constraint for VFA T_1 mapping, as well as multi-echo spin-echo T_2 mapping. Pandey et al. (2021) proposed a joint TV-based reconstruction for accelerated qMRI reconstruction. Le et al. (2022) used recurrent neural networks with a cyclic, model-based loss for cardiac MOLLI T_1 mapping. Wang et al. (2023) proposed a model-based reconstruction method using a joint ℓ_1 -Wavelet spatial regularization and temporal TV regularization for free-breathing myocardial T_1 mapping.

The methods mentioned above are all frequentist approaches that maximize a likelihood (usually assuming Gaussian errors, corresponding to least-squares problems). However, using (regularized) maximum likelihood estimation on each pixel independently typically leads to noisy reconstructions and does not account for uncertainty in the estimation. Bayesian inference has proven effective in modeling uncertainties by incorporating prior knowledge and observed data within a probabilistic framework (Gelman et al., 1995). Bayesian models provide full posterior distri-

butions over model parameters and predictions, allowing not only point estimates but also quantification of credibility and uncertainty. For instance, Beirinckx et al. (2022) used a Bayesian approach with a TV prior for T_1 and T_2 mapping with super-resolution. Löfstedt et al. (2020) developed a Bayesian method based on a bounded TV prior for T_1 mapping that both reduced and estimated uncertainty in parametric maps. Huang et al. (2025) proposed a Bayesian formulation that models the wavelet coefficients of VFA-multi-echo images as Laplace distributed to achieve model-based parametric map reconstructions.

In Bayesian inference, the prior distribution encodes our beliefs about the unknown parameters before observing any data. Common priors include noninformative priors, such as the uniform or Jeffreys priors, weak priors, and informative priors. Selecting an appropriate prior distribution is crucial since an inappropriate informative prior can disproportionately affect the resulting posterior inferences (Jones et al., 2022). The Laplace prior is a commonly used prior that promotes sparsity but uses a single scale parameter to control both mass near zero and tail heaviness, often reducing flexibility and causing over-shrinkage of large coefficients. To address this issue, alternative priors have been introduced that decouple sparsity enforcement from tail behavior. Examples include the Horseshoe (Carvalho et al., 2010) or spike-and-slab priors (Ishwaran and Rao, 2005), both of which allow strong shrinkage of small coefficients while preserving large coefficients. Armagan et al. (2013) proposed a generalized double Pareto prior which retains the sharp peak at zero characteristic of the Laplace distribution while exhibiting heavy, Student's t-like tails. Tang et al. (2019) adapted this prior for rank penalization under the Bayesian framework enabling adaptive learning of the rank.

While these priors effectively encode sparsity, they do not explicitly capture spatial structure. In images, neighboring pixels/voxels are typically strongly correlated, reflecting the inherent spatial continuity of physical tissues or structures. Incorporating such spatially structured information into the prior can therefore substantially improve inference quality. TV regularization encourages smooth solutions by promoting sparsity in the spatial gradients of the coefficients, effectively favoring homogeneous parameter maps and sharp edges, and suppressing small, noisy fluctuations (Michel et al., 2011; Hadj-Selem et al., 2018). Incorporating TV directly into the Bayesian framework poses challenges: the standard TV penalty does not trivially correspond to a proper prior distribution, as it lacks a normalizable probability density over the parameter space (see Lemma 1). To address this, Löfstedt et al. (2020) used a bounded version of the TV prior, restricting the set of acceptable parameter values. While that approach results in a proper prior, it still suffers from key limitations. It requires the exclusion of certain parameter values from the prior's support, limit-

ing the flexibility and interpretability. To overcome these issues, our previous work (Lin et al., 2025) combined the anisotropic TV function and the ℓ_1 norm, forming a proper prior, *i.e.*, one that defines a valid probability distribution—without excluding any parameter values. And experiments on synthetic brain T_1 mapping data showed that the method provides smoother results and narrower credible intervals.

In this paper, we rigorously show that a prior constructed solely from the TV functional cannot be normalized and therefore fails to define a proper probability distribution over the parameter space. We subsequently generalize the prior construction from previous work (Lin et al., 2025) to accommodate both anisotropic and isotropic TV combined with any ℓ_p norm with $p \geq 1$, and prove that the resulting prior is proper. We applied this proposed TV- ℓ_p prior to Bayesian T_1 mapping.

The proposed method was evaluated by comparing it with maximum-likelihood estimation and several Bayesian alternatives based on uniform, Gamma, and bounded TV priors. The evaluation includes experiments on synthetic brain and cardiac T_1 mapping datasets (acquired with VFA and MOLLI, respectively), as well as a real in-vivo breast T_1 mapping dataset (acquired with VFA). The results show that across all datasets, the TV- ℓ_p prior consistently produces more concentrated posterior densities, indicating a substantial reduction in uncertainty compared to alternative methods. Moreover, it achieves lower variance and a smaller negative bias in the estimates, resulting in more robust and reliable T_1 estimates. These results highlight the ability of the TV- ℓ_p prior to preserve spatial structure while providing uncertainty quantification, making it particularly effective for both synthetic simulations and real in-vivo measurements.

2. Background

2.1 Signal Models

T_1 mapping is performed by fitting the measured data to a corresponding signal model. The proposed prior is evaluated using data acquired with the VFA and MOLLI sequences. The following subsection describes the signal models employed for these two acquisition schemes.

2.1.1 VFA T_1 Mapping

VFA T_1 mapping quantifies T_1 values from repeated scans at different excitation flip angles. Specifically, VFA acquires two or more gradient-echo (GRE) data sets with different excitation angles, α_i for $i = 1, 2, \dots, I$ (Baudrexel et al., 2018). The signal intensity, S , is a function of the longitudinal relaxation time, T_1 , repetition time, TR, flip angle, α_i , and the equilibrium longitudinal magnetization, M_0 , which is determined by the proton density and other factors (Deoni

et al., 2003). The signal intensity is defined as

$$S_{\theta}(\alpha_i) = M_0 \frac{1 - e^{-\frac{TR}{T_1}}}{1 - \cos(\alpha_i) e^{-\frac{TR}{T_1}}} \sin \alpha_i, \quad (1)$$

where $\theta = (T_1, M_0)$ denotes the parameters to be estimated, which vary for tissues with different properties. In practice, however, residual unspoiled transverse magnetization may yield discrepancies between measured signal intensities and the theoretical values modeled in Equation (1) (Baudrexel et al., 2018). To account for this variability, the measurement at the n -th voxel is modeled as

$$y_n^{(i)} = S_{\theta_n}^{(i)} + \epsilon_n^{(i)}, \quad (2)$$

where the additive noise terms, $\epsilon_n^{(i)}$, are assumed to be independent and follow a zero-mean Normal distribution with variance σ^2 , *i.e.*, $\epsilon_n^{(i)} \sim \mathcal{N}(0, \sigma^2)$. Here, θ_n denotes the pair (T_1, M_0) at voxel n , and the index i identifies the acquisition associated with the i -th flip angle.

2.1.2 MOLLI T_1 Mapping

In contrast to VFA, the MOLLI method estimates T_1 using IR data. The MOLLI signal is modeled as

$$S_{\theta}(t_i) = A - B \exp(-t_i/T_1^*), \quad (3)$$

where $\theta = (T_1^*, A, B)$ denotes the parameters to be estimated, which vary for tissues with different properties. As with VFA, the actual measurements are affected by systematic error (Messroghli et al., 2004). Therefore, the observed signal at voxel n and inversion time t_i is modeled following the same formulation as in Equation (2). The true longitudinal relaxation time, T_1 , is derived from the estimated parameters as

$$T_1 = T_1^* \left(\frac{B}{A} - 1 \right). \quad (4)$$

2.2 MLE Reconstruction

Typically, the parameters in θ are estimated by fitting the model in Equation (2) using (regularized) maximum log-likelihood estimation (MLE). Assuming the measurements, $y_n^{(i)} \sim \mathcal{N}(S_{\theta_n}^{(i)}, \sigma^2)$, are independent leads to a voxel-wise optimization problem that minimizes the least squares error, formulated as,

$$\hat{\theta}_{\text{ML}} = \arg \max_{\theta} \sum_{i=1}^I \sum_{n=1}^N \log \mathcal{N}(y_n^{(i)} | S_{\theta_n}^{(i)}, \sigma^2) \\ = \arg \min_{\theta} \left[\frac{1}{2\sigma^2} \sum_{i=1}^I \sum_{n=1}^N (y_n^{(i)} - S_{\theta_n}^{(i)})^2 \right], \quad (5)$$

$$(6)$$

where N denotes the total number of voxels.

2.3 Bayesian Model

A Bayesian model combines a likelihood, such as that in Equation (5), with a prior distribution, that accounts for available prior information or any other prior beliefs. Specifically, a Bayesian model assigns a posterior distribution over the parameters, constructed using a likelihood and a prior, as

$$p(\boldsymbol{\theta} | \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\mathbf{y} | \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}} \propto p(\mathbf{y} | \boldsymbol{\theta})p(\boldsymbol{\theta}), \quad (7)$$

where $p(\boldsymbol{\theta} | \mathbf{y})$, $p(\mathbf{y} | \boldsymbol{\theta})$, and $p(\boldsymbol{\theta})$ are the posterior distribution, the likelihood, and the prior distribution, respectively. The proportionality is in the parameters, $\boldsymbol{\theta}$. Since it is impossible or at least computationally intractable to compute the integral in the denominator, an efficient means is to draw samples from the posterior using Markov chain Monte Carlo (MCMC) methods (Hoffman and Gelman, 2014).

2.4 Total Variation

The TV of a smooth intensity function ξ defined on a domain Ω is

$$\text{TV}(\xi) = \int_{\Omega} \|\nabla \xi\|_2 d\Omega.$$

For a two-dimensional (2D) image, where the intensity is represented by pixel values in an n -by- n matrix, \mathbf{x} , the gradient $\nabla \xi$ can be approximated with a finite difference scheme on the pixel values, for instance, the forward differences

$$\nabla_{i,j} \mathbf{x} = (x_{i+1,j} - x_{i,j}, x_{i,j+1} - x_{i,j})^T. \quad (8)$$

There are two common versions of the TV of \mathbf{x} , an isotropic and an anisotropic version. Using the scheme in Equation (8), the isotropic TV is defined as (Rudin et al., 1992)

$$\begin{aligned} \text{TV}_{iso}(\mathbf{x}) &= \sum_{i,j} \|\nabla_{i,j} \mathbf{x}\|_2 \\ &= \sum_{i,j} \sqrt{|x_{i+1,j} - x_{i,j}|^2 + |x_{i,j+1} - x_{i,j}|^2}, \end{aligned} \quad (9)$$

and the anisotropic TV as (Esedoğlu and Osher, 2004)

$$\begin{aligned} \text{TV}_{aniso}(\mathbf{x}) &= \sum_{i,j} \|\nabla_{i,j} \mathbf{x}\|_1 \\ &= \sum_{i,j} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}|. \end{aligned} \quad (10)$$

Although here the argument, \mathbf{x} , is assumed to be a 2D structure (e.g., a 2D image), the definitions generalize trivially to d -dimensional structures.

Let $\bar{\mathbf{x}} \in \mathbb{R}^N$ be the vectorized (flattened) version of \mathbf{x} . The isotropic TV can then be expressed as

$$\text{TV}_{iso}(\mathbf{x}) = \sum_{i,j} \|\mathbf{D}_{i,j} \bar{\mathbf{x}}\|_2,$$

and the anisotropic TV as

$$\text{TV}_{aniso}(\mathbf{x}) = \sum_{i,j} \|\mathbf{D}_{i,j} \bar{\mathbf{x}}\|_1 = \|\mathbf{D} \bar{\mathbf{x}}\|_1,$$

where $\mathbf{D}_{i,j} \in \{-1, 0, 1\}^{d \times N}$ denotes the elements of a finite-difference matrix (e.g., implementing the forward difference scheme in Equation (8)) that computes discrete derivatives along d dimensions from $\bar{\mathbf{x}}$ taking into account the spatial indices of \mathbf{x} . The operator \mathbf{D} denotes the global linear difference operator obtained by stacking all local operators $\mathbf{D}_{i,j}$, thereby computing all finite differences of $\bar{\mathbf{x}}$ simultaneously.

3. Proposed Bayesian Model

Choosing the prior that best represents available prior information is vital for Bayesian inference. In this work, we propose to construct a spatially structured prior based on the TV function, which promotes smoothness by assigning high probabilities to small spatial changes in the parameter maps. The proposed formulation defines a proper prior density with full support over the parameter domain, ensuring that all possible values retain nonzero prior probability. We further formulate T_1 mapping within a Bayesian framework and employ the proposed prior to encode smooth spatial structure, thereby enabling uncertainty quantification and improving estimation performance.

3.1 Structured Priors Based on Total Variation

In images, neighboring pixels are often strongly correlated, and in qMRI images, neighboring pixels often represent the same underlying tissues. Hence, the parametric maps are likely to be locally spatially smooth. In this section, we first naively construct a prior based on TV but then show that the resulting distribution is not integrable and hence does not induce a proper prior. To address this, we propose a novel family of priors that contains an additional energy term that ensures they are integrable.

A prior distribution based on the TV function can naively be defined using the Boltzmann distribution (Gibbs distribution) as

$$p_{\text{TV}}(\mathbf{x}) = \frac{1}{Z} e^{-\lambda \text{TV}(\mathbf{x})}, \quad (11)$$

where $\lambda > 0$, the Z is a normalizing constant, and TV represent TV_{iso} or TV_{aniso} . This distribution results in high probabilities for sparse spatial changes in the model parameters. However, the normalizing constant,

$$Z = \int_{\mathbb{R}^N} e^{-\lambda \text{TV}(\mathbf{x})} d\bar{\mathbf{x}}, \quad (12)$$

diverges due to invariance in the direction of $\mathbb{1}$, i.e., for constant images. We have the following result.

Lemma 1 Let $\lambda \in \mathbb{R}_+$, the integral

$$Z = \int_{\mathbb{R}^N} e^{-\lambda \text{TV}(\bar{\mathbf{x}})} d\bar{\mathbf{x}}$$

diverges for both isotropic and anisotropic TV.

Proof Let $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N]$ be an orthonormal basis in \mathbb{R}^N , where $\mathbf{v}_1 = \frac{1}{\sqrt{N}}\mathbb{1}$, where $\mathbb{1} = (1, 1, \dots, 1)^T \in \mathbb{R}^N$ is the constant vector (with all elements being 1). Then $\bar{\mathbf{x}} = \mathbf{V}\mathbf{k}$, for some $\mathbf{k} \in \mathbb{R}^N$.

For isotropic TV, $D_{i,j}\mathbf{v}_1 = \mathbf{0}$ for any i and j , thus,

$$\begin{aligned} D_{i,j}\bar{\mathbf{x}} &= D_{i,j}(\mathbf{V}\mathbf{k}) = D_{i,j} \left(k_1\mathbf{v}_1 + \sum_{n=2}^N k_n\mathbf{v}_n \right) \\ &= D_{i,j} \sum_{n=2}^N k_n\mathbf{v}_n, \end{aligned}$$

which depends only on $\tilde{\mathbf{k}} = (k_2, \dots, k_N)^T$. The integral for isotropic TV can thus be rewritten as

$$\begin{aligned} Z &= \int_{\mathbb{R}^N} e^{-\lambda \sum_{i,j} \|D_{i,j}\bar{\mathbf{x}}\|_2} d\bar{\mathbf{x}} \\ &= \int_{\mathbb{R}^N} e^{-\lambda \sum_{i,j} \|D_{i,j} \sum_{n=2}^N k_n\mathbf{v}_n\|_2} \underbrace{|\det \mathbf{V}|}_{=1} d\mathbf{k} \quad (13) \\ &= \int_{-\infty}^{\infty} \left(\int_{\mathbb{R}^{N-1}} e^{-\lambda \sum_{i,j} \|D_{i,j} \sum_{n=2}^N k_n\mathbf{v}_n\|_2} d\tilde{\mathbf{k}} \right) dk_1. \end{aligned}$$

Let $C := \int_{\mathbb{R}^{N-1}} e^{-\lambda \sum_{i,j} \|D_{i,j} \sum_{n=2}^N k_n\mathbf{v}_n\|_2} d\tilde{\mathbf{k}}$. The integrand is strictly positive and bounded for all $\tilde{\mathbf{k}}$, and therefore $C > 0$. Moreover, since $\text{null}(\text{TV}) = \text{span}\{\mathbf{v}_1\}$, this guarantees exponential decay of the integrand at infinity and thus $C < \infty$. Therefore,

$$Z = \int_{-\infty}^{\infty} C dk_1 = \lim_{R \rightarrow \infty} \int_{-R}^R C dk_1 = \lim_{R \rightarrow \infty} 2RC = \infty.$$

For anisotropic TV, $D\mathbf{v}_1 = \mathbf{0}$, and thus,

$$D\bar{\mathbf{x}} = D(\mathbf{V}\mathbf{k}) = D \left(k_1\mathbf{v}_1 + \sum_{n=2}^N k_n\mathbf{v}_n \right) = \sum_{n=2}^N k_n D\mathbf{v}_n.$$

The integral for anisotropic TV can thus be rewritten as

$$\begin{aligned} Z &= \int_{\mathbb{R}^N} e^{-\lambda \|D\bar{\mathbf{x}}\|_1} d\bar{\mathbf{x}} \\ &= \int_{\mathbb{R}^N} e^{-\lambda \left\| \sum_{n=2}^N k_n D\mathbf{v}_n \right\|_1} \underbrace{|\det \mathbf{V}|}_{=1} d\mathbf{k} \quad (14) \end{aligned}$$

$$= \int_{-\infty}^{\infty} \left(\int_{\mathbb{R}^{N-1}} e^{-\lambda \left\| \sum_{n=2}^N k_n D\mathbf{v}_n \right\|_1} d\tilde{\mathbf{k}} \right) dk_1, \quad (15)$$

Let $C' := \int_{\mathbb{R}^{N-1}} e^{-\lambda \left\| \sum_{n=2}^N k_n D\mathbf{v}_n \right\|_1} d\tilde{\mathbf{k}}$. Again, $0 < C' < \infty$, and thus

$$Z = \int_{-\infty}^{\infty} C' dk_1 = \lim_{R \rightarrow \infty} \int_{-R}^R C' dk_1 = \lim_{R \rightarrow \infty} 2RC' = \infty.$$

Hence, the normalizing constants diverge for both $\text{TV} = \text{TV}_{iso}$ and $\text{TV} = \text{TV}_{aniso}$. \blacksquare

In Lemma 1, the normalization constant is shown to diverge when the integration is taken over the entire space \mathbb{R}^N , without imposing any constraints on the values of $\bar{\mathbf{x}}$. However, in the practical quantitative MRI application of T_1 mapping considered in this work, the unknown parameter, $\bar{\mathbf{x}}$, represents T_1 relaxation times, and they are physically meaningful only for nonnegative values. It is therefore natural to ask whether restricting the domain to the non-negative orthant, \mathbb{R}_+^N , suffices to render the corresponding normalization constant finite.

The following Lemma 2 shows that this is not the case when the integration domain is restricted to \mathbb{R}_+^N , because the normalization constant associated with the TV prior remains divergent. This result highlights that the divergence is not caused by allowing negative values, but rather is a direct consequence of the translation invariance of the TV semi-norm with respect to constant shifts, namely that $\text{TV}(\bar{\mathbf{x}}) = \text{TV}(\bar{\mathbf{x}} + c\mathbb{1})$, $\forall c \in \mathbb{R}$.

Lemma 2 Let $\bar{\mathbf{x}} \in \mathbb{R}_+^N$ and $\lambda \in \mathbb{R}_+$, the integral

$$Z = \int_{\mathbb{R}_+^N} e^{-\lambda \text{TV}(\bar{\mathbf{x}})} d\bar{\mathbf{x}}$$

diverges for both isotropic and anisotropic TV.

Proof Fix $r > 0$ and define the compact set

$$B_r := \{\mathbf{u} \in \mathbb{R}^N : \|\mathbf{u}\|_{\infty} \leq r\}.$$

Since TV is continuous and B_r is compact, there exists

$$M_r := \max_{\mathbf{u} \in B_r} \text{TV}(\mathbf{u}) < \infty.$$

Now define

$$S_r^s := \{\bar{\mathbf{x}} = t\mathbb{1} + \mathbf{u} : t \in [r, s], \mathbf{u} \in B_r\}.$$

For any $\bar{\mathbf{x}} \in S_r^s$, each element thus satisfies $\bar{x}_k = t + u_k \geq r - r = 0$ and hence $S_r^s \subset \mathbb{R}_+^N$. From Equation (9) and Equation (10) it follows immediately that,

$$\text{TV}(\bar{\mathbf{x}}) = \text{TV}(t\mathbb{1} + \mathbf{u}) = \text{TV}(\mathbf{u}) \leq M_r, \quad \forall \bar{\mathbf{x}} \in S_r^s,$$

and therefore

$$e^{-\lambda \text{TV}(\bar{\mathbf{x}})} \geq e^{-\lambda M_r}, \quad \forall \bar{\mathbf{x}} \in S_r^s.$$

Consequently, for any $s > r$

$$\begin{aligned} Z &= \int_{\mathbb{R}_+^N} e^{-\lambda \text{TV}(\bar{\mathbf{x}})} d\bar{\mathbf{x}} \geq \int_{S_r^s} e^{-\lambda \text{TV}(\bar{\mathbf{x}})} d\bar{\mathbf{x}} \\ &\geq e^{-\lambda M_r} \int_{S_r^s} d\bar{\mathbf{x}} = e^{-\lambda M_r} \text{Vol}(S_r^s), \end{aligned}$$

where $\text{Vol}(S_r^s)$ denotes the volume of the set S_r^s . Taking the limit as $s \rightarrow \infty$ yields

$$Z \geq \lim_{s \rightarrow \infty} e^{-\lambda M_r} \text{Vol}(S_r^s) = \infty.$$

Löfstedt et al. (2020) proposed to incorporate a uniform prior alongside the TV prior, such that the TV was only defined within the domain of the uniform prior, to force a proper prior on the model parameter and to ensure convergence during the sampling process. This constraint bounds the model parameters to be in a certain range, making the corresponding integral finite. Our previous work (Lin et al., 2025) proposed an ℓ_1 -based modification of the anisotropic TV prior to ensure that it defines a proper probability distribution. Building on that formulation, the present work generalizes the prior by incorporating an ℓ_p norm with $p \geq 1$ into the isotropic or anisotropic TV function. We propose the following modified TV function.

Lemma 3 Let TV_μ^p be a modified TV function, such that

$$\text{TV}_\mu^p(\mathbf{x}) = \text{TV}(\mathbf{x}) + \mu \|\bar{\mathbf{x}}\|_p, \quad (16)$$

where $\text{TV} = \text{TV}_{iso}$ or $\text{TV} = \text{TV}_{aniso}$, $\mu \in \mathbb{R}_+$, and $\|\cdot\|_p$ is an ℓ_p norm with $p \geq 1$. Then the integral

$$Z = \int_{\mathbb{R}^N} e^{-\lambda \text{TV}_\mu^p(\mathbf{x})} d\mathbf{x} < \infty,$$

for all $\lambda \in \mathbb{R}_+$.

Proof Note that $\lambda \|\mathbf{D}\bar{\mathbf{x}}\|_1 \geq 0$ for any $\bar{\mathbf{x}} \in \mathbb{R}^N$, so we can write

$$\begin{aligned} \int_{\mathbb{R}^N} e^{-\lambda \text{TV}_\mu^p(\mathbf{x})} d\bar{\mathbf{x}} &= \int_{\mathbb{R}^N} e^{-\lambda \|\mathbf{D}\bar{\mathbf{x}}\|_1 - \lambda \mu \|\bar{\mathbf{x}}\|_p} d\bar{\mathbf{x}} \\ &\leq \int_{\mathbb{R}^N} e^{-\lambda \mu \|\bar{\mathbf{x}}\|_p} d\bar{\mathbf{x}} \\ &\leq \int_{\mathbb{R}^N} e^{-\lambda \mu \|\bar{\mathbf{x}}\|_1 / N^{1-\frac{1}{p}}} d\bar{\mathbf{x}} \\ &= \prod_{n=1}^N \int_{-\infty}^{\infty} e^{-\lambda \mu |x_n| / N^{1-\frac{1}{p}}} dx_n \\ &= \left(\frac{2N^{1-\frac{1}{p}}}{\lambda \mu} \right)^N < \infty. \end{aligned}$$

The second inequality follows by noting that when $p = 1$ then $N^{1-\frac{1}{p}} = 1$, and that when $p > 1$ Hölder's inequality gives $\|\mathbf{x}\|_1 \leq N^{1-\frac{1}{p}} \|\mathbf{x}\|_p$. ■

A geometric illustration of Lemmas 1 and 3 is provided in Figure 1. In Figure 1(a), along the subspace spanned by

$\mathbb{1}$, we have $\|\mathbf{D}\bar{\mathbf{x}}\|_1 = 0$, and thus the integrand $e^{-\lambda \|\mathbf{D}\bar{\mathbf{x}}\|_1}$ remains constant. As a result, integration along this unbounded direction leads to divergence of the integral. In contrast, Figures 1(b) and 1(c) illustrate the integrand $e^{-\text{TV}_\mu^p(\mathbf{x})}$, where the inclusion of the additional $\|\bar{\mathbf{x}}\|_p$ term induces exponential decay in all directions of \mathbb{R}^n . This ensures that the integral converges and is finite.

We see now that a finite normalization constant leads to a proper prior.

Theorem 4 The modified prior,

$$p_{\text{TV}_\mu^p}(\mathbf{x}) = \frac{1}{Z} e^{-\lambda \text{TV}_\mu^p(\mathbf{x})},$$

is a proper prior.

Proof Follows from Lemma 3, which says we can normalize it to unit integral. ■

The generalized TV_μ^p prior broadens the class of valid TV-based spatial priors and enhances modeling flexibility, allowing better adaptation to varying image characteristics and regularization requirements.

3.2 Posterior Based on Total Variation

We propose a Bayesian framework for T_1 mapping, where the likelihood is based on the signal models in Equations (1) or (3), and spatial prior information is incorporated through a generalized TV_μ^p prior. A half-normal prior with scale parameter $\sigma_\epsilon > 0$ is placed on σ to enforce positivity and encode prior knowledge about the noise level. The posterior is then given by,

$$\begin{aligned} p_{\text{TV}_\mu^p}(\boldsymbol{\theta}, \sigma \mid \mathbf{y}) &\propto p(\mathbf{y} \mid \boldsymbol{\theta}, \sigma) p_{\text{TV}_\mu^p}(\boldsymbol{\theta}) p(\sigma) \\ &= \left(\sqrt{2\pi\sigma^2} \right)^{-NI} \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^I \sum_{n=1}^N (y_n^{(i)} - S_{\boldsymbol{\theta}_n}^{(i)})^2 \right) \\ &\quad \cdot \prod_{\boldsymbol{\theta}} \frac{1}{Z_{\boldsymbol{\theta}}} e^{-\lambda \text{TV}_\mu^p(\boldsymbol{\theta})} \cdot \frac{\sqrt{2}}{\sigma_\epsilon \sqrt{\pi}} \exp \left(-\frac{\sigma^2}{2\sigma_\epsilon^2} \right), \end{aligned} \quad (17)$$

where for VFA acquisitions, $\boldsymbol{\theta} = (T_1, M_0)$, and for MOLLI acquisitions, $\boldsymbol{\theta} = (T_1^*, A, B)$. Since the overall normalizing constant is intractable, we sample from the unnormalized posterior using MCMC to quantify the uncertainty in the estimated parameters.

We further extend the Bayesian framework for T_1 mapping to a hierarchical model by assigning hyper-priors to the hyper-parameters λ and μ , which control the strength and shape of the generalized TV_μ^p prior. The hierarchical model is defined as,

$$\begin{aligned} p_{\text{TV}_\mu^p}(\boldsymbol{\theta}, \sigma, \lambda, \mu \mid \mathbf{y}) \\ \propto p(\mathbf{y} \mid \boldsymbol{\theta}, \sigma, \lambda, \mu) p_{\text{TV}_\mu^p}(\boldsymbol{\theta} \mid \mu, \lambda) p(\sigma) p(\lambda) p(\mu). \end{aligned} \quad (18)$$

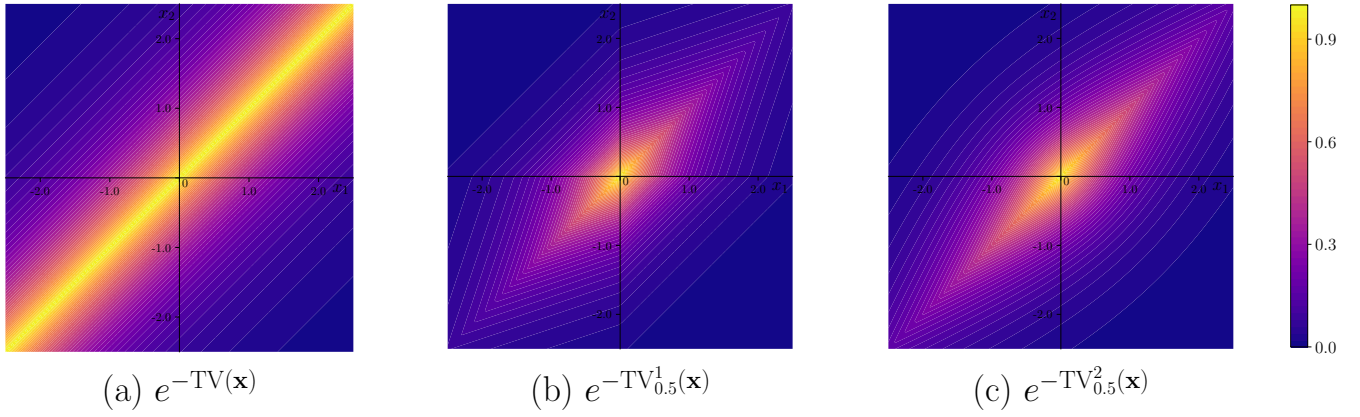


Figure 1: 2D heatmaps for the unnormalized priors (a): $e^{-TV(\mathbf{x})}$; (b): $e^{-TV_{0.5}^1(\mathbf{x})}$ and (c): $e^{-TV_{0.5}^2(\mathbf{x})}$.

4. Experiments

In the experiments, we evaluated the anisotropic TV_{μ}^p prior with $p \in \{1, 2\}$ for VFA and MOLLI T_1 mapping within the Bayesian models on both synthetic and real datasets. The performance was compared to that of MLE and alternative Bayesian models using uniform, Gamma, and bounded TV priors. The hyperparameters λ and μ in the TV_{μ}^p prior were selected either by assigning hyperpriors, referred to as the Hierarchical TV_{μ}^p model, or through hyperparameter tuning, denoted simply as the TV_{μ}^p model.

For the Hierarchical TV_{μ}^p model, the λ and μ are constrained to be non-negative and were therefore assigned exponential priors. The rate parameters of these exponential priors were determined using Bayesian optimization. For the TV_{μ}^p model, λ and μ were treated as fixed hyperparameters and were tuned using Bayesian optimization. During the Bayesian optimization, the optimal combination of hyperparameters was selected as the one that maximized the expected log point-wise predictive density (elpd; Vehtari et al., 2017), estimated using the widely applicable information criterion (WAIC; Watanabe and Opper, 2010), defined as,

WAIC \approx

$$\frac{1}{NI} \sum_{i=1}^I \sum_{n=1}^N \log \left(\frac{1}{S} \sum_{s=1}^S p(y_n^{(i)} | \theta_n^s, \sigma^s) \right) - \frac{1}{NI} \sum_{i=1}^I \sum_{n=1}^N \frac{1}{S-1} \sum_{s=1}^S \left(\log p(y_n^{(i)} | \theta_n^s, \sigma^s) - \mu_n^{(i)} \right)^2, \quad (19)$$

where $\mu_n^{(i)} = \frac{1}{S} \sum_{s=1}^S \log p(y_n^{(i)} | \theta_n^s, \sigma)$, and θ_n^s and σ^s with $s = 1, 2, \dots, S$ denote samples drawn from the posterior distribution. The parameter $p \in \{1, 2\}$ in the TV_{μ}^p model was also selected as the value that maximizes the elpd.

4.1 Baselines

MLE formulates the estimation as a voxel-wise optimization problem that minimizes the sum of squared errors, expressed as in Equation (5). The MLE solution was found using the L-BFGS-B algorithm (Liu and Nocedal, 1989). The baseline priors used in the experiments were Uniform(0, 50), Gamma(3, 1) and the bounded TV prior. Specifically, the Uniform(0, 50) prior represent weak prior knowledge. It enforces non-negativity and restricts the parameter to a physically meaningful range that covers all expected parameter values across tissues and acquisition settings. The Gamma(3, 1) prior was used to enforce non-negativity while providing moderate concentration around physiologically realistic T_1 values, offering more informative guidance than the Uniform(0, 50) prior. The bounded TV prior was defined as (Löfstedt et al., 2020)

$$p_{bTV}(\mathbf{x}) = p_{TV}(\mathbf{x}) p_{\text{uni}}(\mathbf{x} | L, H), \quad \text{with } L = 0, H = 50,$$

where the uniform prior, $p_{\text{uni}}(\mathbf{x} | L, H)$, imposes hard bounds on the parameter range. The hyperparameter λ was included in the bounded TV prior as well, and was selected through Bayesian optimization.

4.2 Data

The experimental evaluation was performed using datasets comprising both synthetic and real T_1 mapping data. The synthetic brain T_1 mapping data were simulated using Equations 1 and 2 with $TR = 6.8$ ms, $TE = 2.1$ ms, and flip angles $\alpha \in \{2^\circ, 4^\circ, 11^\circ, 13^\circ, 15^\circ\}$. The ground truth values of T_1 and M_0 were generated from the BrainWeb phantom¹ using Hero Imaging². The parameter maps generated had a matrix size of 256×256 and a voxel size of $0.98 \times 0.98 \times 2.00$ mm³. The added noise was complex circular Gaussian noise.

1. <https://brainweb.bic.mni.mcgill.ca/brainweb>
 2. <https://www.heroimaging.com>

For the synthetic cardiac T_1 mapping data, the ground-truth T_1 maps and inversion time were obtained from a publicly available dataset (Le et al., 2022). Specifically, the inversion times were

$$TI \in [128, 208, 881, 951, 1628, 1711, 2381, 3131] \text{ ms.}$$

Synthetic measurements were generated using the MOLLI signal model described in Equations (3) and (4), with additive Gaussian noise applied according to Equation (2).

We also used a real breast MRI dataset from the publicly available *QIN-BREAST-02* collection³ (Yankeelov et al., 2019). The data were acquired using a radio frequency-spoiled 3D GRE VFA technique with ten flip angles ranging from 2° to 20° in 2° increments. The acquisition matrix was $192 \times 192 \times 20$, covering the full breast in the sagittal plane over a 22 cm^2 field of view, with a slice thickness of 5 mm. From this dataset, we selected one subject (QIN-BREAST-02-0011) diagnosed with *left-sided invasive ductal carcinoma*. The subject underwent neoadjuvant chemotherapy consisting of doxorubicin and cyclophosphamide, followed by paclitaxel.

4.3 Implementation details

All Bayesian models were developed using PyMC 5.15.0⁴. Samples were drawn from the posterior distributions using the No-U-Turn Sampler algorithm (NUTS; Hoffman and Gelman, 2014). Four independent Markov chains were run, each generating 2,000 samples after discarding 2,000 burn-in iterations. Convergence was evaluated by inspecting trace plots and calculating the \hat{R} statistic, which assesses the consistency between and within chains to confirm reliable mixing and sampling (Gelman and Rubin, 1992).

The results from the different methods were evaluated by analyzing the estimated posterior distributions, as well as the bias and variance of the posterior samples. Specifically, the probability density functions (PDFs) of T_1 values were estimated using kernel density estimation with Gaussian kernels (Silverman, 2018), based on posterior samples for each method. PDFs in representative regions were then visualized to examine their shape and spread, thereby providing insight into the estimation uncertainties. The uncertainty was further quantified by computing the credible interval spanning two standard deviations around the posterior mean. Bias and variance were also used for quantitative comparisons and were estimated as

$$\text{Bias} = \frac{1}{S} \sum_{s=1}^S (\theta^s - \theta), \text{ and} \quad (20)$$

$$\text{Variance} = \frac{1}{S-1} \sum_{s=1}^S (\theta^s - \theta_{\text{mean}})^2, \quad (21)$$

where θ denotes the ground truth, θ^s is the s -th posterior sample, and $\theta_{\text{mean}} = \frac{1}{S} \sum_{s=1}^S \theta^s$ is the empirical posterior mean.

5. Results and Discussion

This section presents the T_1 mapping results obtained on the synthetic Brain and Cardiac MRI datasets, as well as the real Breast MRI dataset, to compare all baseline and proposed methods. The synthetic experiments provide controlled settings for quantitative evaluation against known ground truth, allowing us to investigate the influence of different priors. In contrast, the real-data experiments assess the practical applicability of the methods under realistic imaging conditions, where unknown complex noise are presented.

5.1 Brain T_1 Mapping

For the Brain T_1 Mapping, the TV_μ^1 model had the highest epld, so $p = 1$ was selected. Figure 2 shows the PDFs of the estimated T_1 values within the regions of interest (ROIs). The corresponding summary statistics are provided in Table 1. The bias and variance for each Bayesian method are visualized as spatial maps in Figure 3, alongside the mean T_1 maps, and as overall comparison metrics across all voxels in Figure 4.

Figure 2 illustrates that, across the three selected ROIs, the TV_μ^1 model produced the most concentrated posterior PDFs, followed by the Bounded TV and the Hierarchical TV_μ^1 models. In contrast, the Gamma and Uniform models result in progressively wider posterior distributions. This behaviour reflects a notable reduction in posterior uncertainty in the T_1 estimates achieved by the proposed method. Table 1 shows that the standard deviations corresponding to the TV_μ^1 prior are smaller than all others across different tissues. The voxelwise MLE is particularly unstable in CSF, yielding unrealistically large T_1 estimates (e.g., 32.01s in Table 1). This can be caused by the fact that, in CSF, the VFA signal is nearly insensitive to variations in T_1 , resulting in a flat likelihood surface. Consequently, the inverse problem becomes weakly identifiable, allowing the MLE to drift toward excessively large values with minimal impact on the data fidelity term.

Figure 3 shows that the average T_1 maps obtained using the TV_μ^1 method appear smoother than those obtained with other priors. This observation is consistent with the bias maps. Additionally, the variance maps indicate a lower variance when using the TV_μ^1 method.

Figure 4 shows that the samples from the model using the TV_μ^1 prior have a smaller negative bias compared to other methods, especially when the T_1 values are high. Also, the variance from the proposed method is clearly much lower than for the other methods.

3. <https://www.cancerimagingarchive.net/collection/qin-breast-02>

4. <https://www.pymc.io>

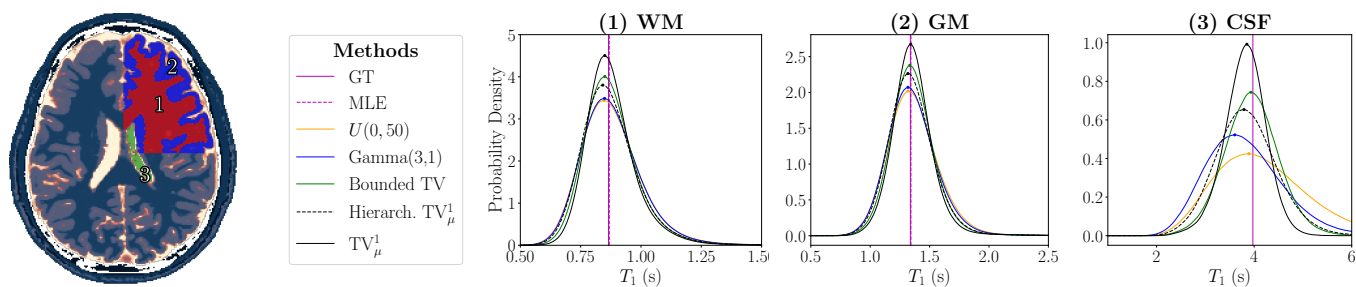


Figure 2: PDFs of estimated T_1 values within three selected ROIs. The ROIs correspond to (1) white matter (WM), (2) gray matter (GM), and (3) cerebrospinal fluid (CSF), highlighted in red, blue, and green, respectively, in the anatomical reference image. For each ROI, the PDFs from different Bayesian methods are shown, along with the corresponding ground truth (GT) and MLE T_1 values for comparison.

Tissue	GT	MLE	U(0, 50)	Gamma(3, 1)	Bounded TV	Hierarch. TV_μ^1	TV_μ^1
WM	0.87	0.87	0.85, 0.88 ± 0.26	0.85, 0.88 ± 0.26	0.85, 0.88 ± 0.23	0.84, 0.87 ± 0.24	0.85, $0.88 \pm \mathbf{0.21}$
GM	1.34	1.34	1.32, 1.37 ± 0.44	1.32, 1.36 ± 0.43	1.33, 1.36 ± 0.38	1.32, 1.35 ± 0.40	1.34, $1.36 \pm \mathbf{0.35}$
CSF	3.97	32.01	3.89, 4.21 ± 2.07	3.60, 3.81 ± 1.59	3.93, 3.94 ± 1.12	3.79, 3.84 ± 1.26	3.85, $3.81 \pm \mathbf{0.86}$

Table 1: Statistics of the estimated T_1 values in (1) WM, (2) GM, and (3) CSF across the different methods. For GT and MLE, the table reports the mean T_1 values. For each Bayesian model, as indicated in the top column, the table reports the posterior mode (*i.e.*, the value with the highest probability density), the posterior mean, and the associated uncertainty expressed as two standard deviations around the mean, in the format “mode, mean $\pm 2\sigma$ ”. All T_1 values are in seconds (s).

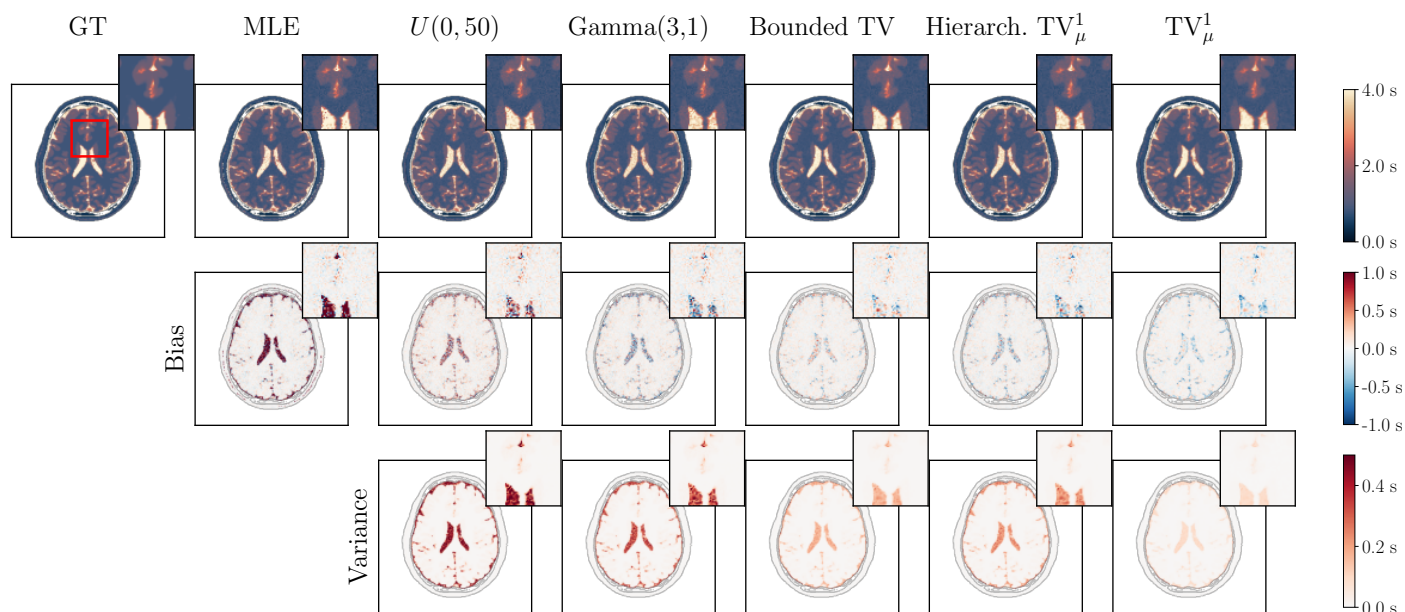


Figure 3: Comparison of mean T_1 , bias and variance maps from different methods. The first row displays the GT T_1 map, the MLE point estimate, and the mean T_1 maps obtained from each Bayesian model, as labeled at the top of the figure. The second row presents the corresponding bias maps, while the third row shows the variance maps. Each subplot includes a zoomed-in view of a region (marked by the red square) in the top-right corner for detailed comparison.

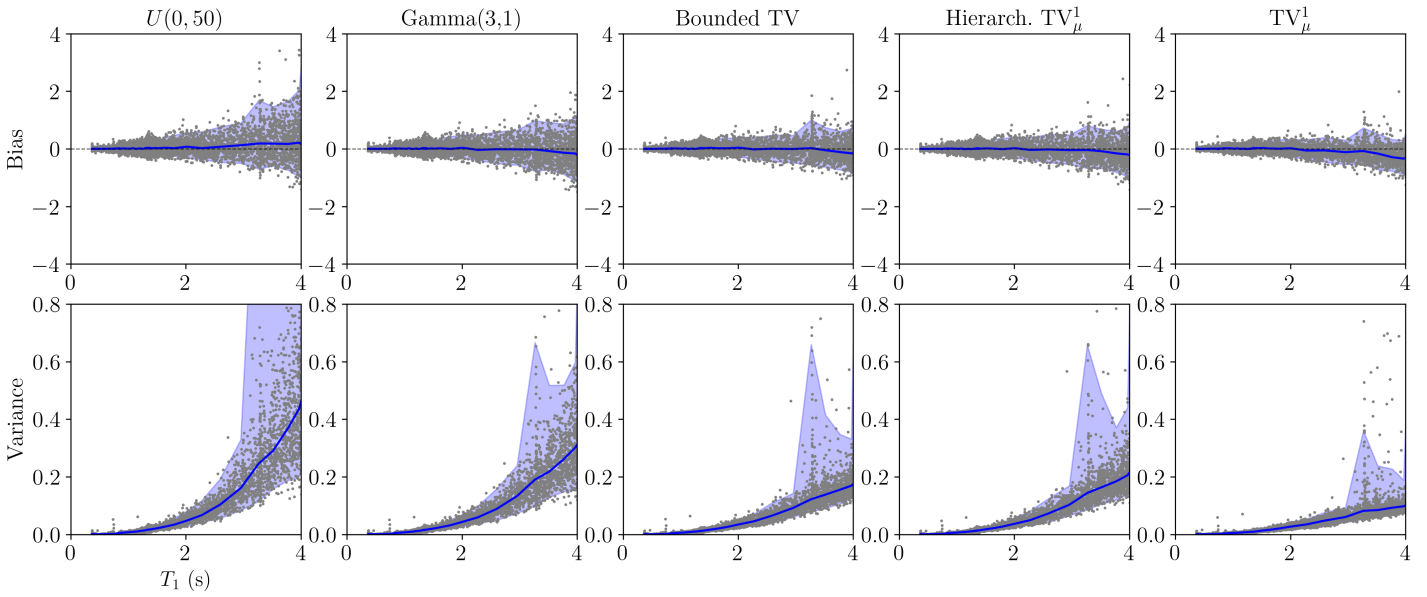


Figure 4: Bias and variance analysis for the Bayesian methods. In contrast to Figure 3, which presents spatially resolved maps, this figure illustrates the relationship between GT T_1 values and the corresponding bias or variance of the estimates. The horizontal axis shows the GT T_1 values, while the vertical axis indicates the bias/variance observed at each voxel. The blue curves represent the average bias/variance for a certain GT T_1 value. The blue shaded area contains 95% of the data points.

5.2 Cardiac T_1 Mapping

For the cardiac T_1 Mapping, the TV_μ^p model with $p = 1$ had the highest epld. Figure 5 shows the PDFs of the estimated T_1 values within the selected ROIs. The corresponding summary statistics are provided in Table 2. The bias and variance for each Bayesian method are visualized as spatial maps in Figure 6, alongside the mean T_1 maps, and as overall comparison metrics across all voxels in Figure 7.

Figure 5 illustrates that, across the three selected ROIs, the TV_μ^1 method produces the most concentrated posterior PDFs, whereas the Hierarchical TV_μ^1 , Bounded TV, Gamma, and Uniform models yield progressively wider distributions. This behaviour reflects a notable reduction in posterior uncertainty in the T_1 estimates achieved by the proposed TV_μ^1 method.

Table 2 shows that the standard deviations corresponding to the TV_μ^1 prior are smaller than all others across different tissues.

Figure 6 shows that the average T_1 maps obtained using the TV_μ^1 method appear smoother than those obtained with other priors. This observation is consistent with the bias maps. Additionally, the variance maps indicate a lower variance when using the TV_μ^1 method.

Figure 7 shows that the samples from the model using the TV_μ^1 prior have a smaller negative bias compared to other methods, especially when the T_1 values are high. Also, the variance from the proposed method is clearly much lower than for the other methods.

5.3 Breast T_1 Mapping

Similar to the synthetic experiment, the TV_μ^p model with $p = 1$ had the highest epld for the Breast T_1 Mapping. Figure 8 presents the PDFs of the estimated T_1 values within the ROIs. Figure 8 shows that the Hierarchical TV_μ^1 method produces the most concentrated posterior PDFs, whereas the TV_μ^1 , Bounded TV, Gamma, and Uniform models yield progressively wider distributions. The multimodal posterior distributions observed for ROIs 2, 3, and 6 reflect heterogeneous tissue composition within the analyzed regions. The corresponding summary statistics are provided in Table 3. Table 3 shows that the standard deviations corresponding to the Hierarchical TV_μ^1 prior are smaller than all others across different tissues.

Figure 9 shows the estimated T_1 map from MLE and the estimated mean T_1 maps from Uniform, Gamma, bounded TV, and proposed Hierarchical TV_μ^1 and TV_μ^1 models. The estimated T_1 maps from the proposed models are smoother than those from classical Bayesian alternatives.

The TV_μ^1 and Hierarchical TV_μ^1 methods exhibit different behavior on the real dataset compared with the synthetic ones. This suggests that the choice between these methods depends on the level of prior knowledge available for a given application. Moreover, the consistently better performance of $p = 1$ compared to $p = 2$ in terms of epld may be attributed to the lower sensitivity of the ℓ_1 norm to large T_1 values. We also noticed that methods with the TV_μ^1 prior do not always yield the most accurate mean

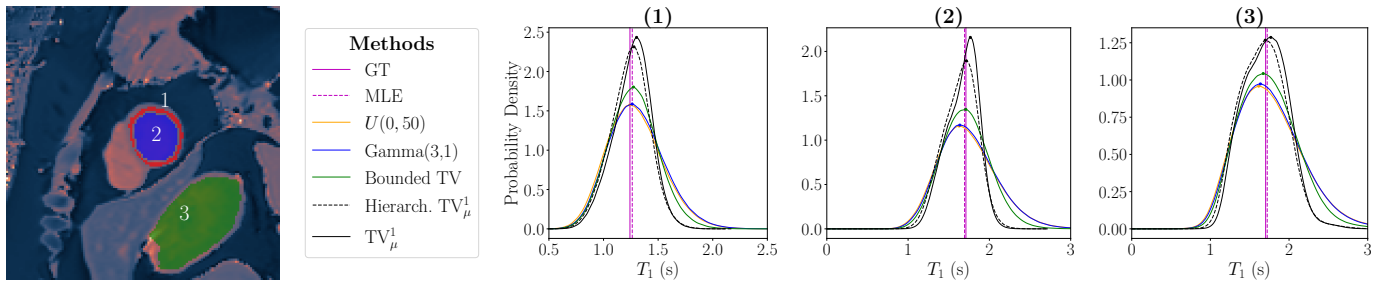


Figure 5: PDFs of T_1 values at selected ROIs (highlighted in red associated with myocardium, blue, and green on the anatomical image). The estimated PDFs from different Bayesian methods are shown, along with the GT and MLE T_1 values for comparison.

Tissue	GT	MLE	U(0, 50)	Gamma(3, 1)	Bounded TV	Hierarch. TV_μ^1	TV_μ^1
(1)	1.24	1.26	1.24, 1.29 ± 0.51	1.26, 1.30 ± 0.51	1.27, 1.28 ± 0.45	1.28, 1.24 ± 0.36	1.30, 1.26 ± 0.35
(2)	1.71	1.69	1.63, 1.73 ± 0.72	1.64, 1.74 ± 0.71	1.71, 1.71 ± 0.59	1.72, 1.64 ± 0.43	1.77, 1.67 ± 0.40
(3)	1.70	1.72	1.61, 1.76 ± 0.88	1.64, 1.77 ± 0.86	1.67, 1.74 ± 0.76	1.70, 1.67 ± 0.59	1.77, 1.70 ± 0.57

Table 2: Statistics of the estimated T_1 values in ROIs across the different methods. The table lists the GT T_1 values and the MLEs, alongside the results from different Bayesian methods, as indicated in the top column. For each Bayesian model, the table reports the posterior mode, mean, and two standard deviations as “mode, mean $\pm 2\sigma$ ”. All T_1 values are reported in seconds (s).

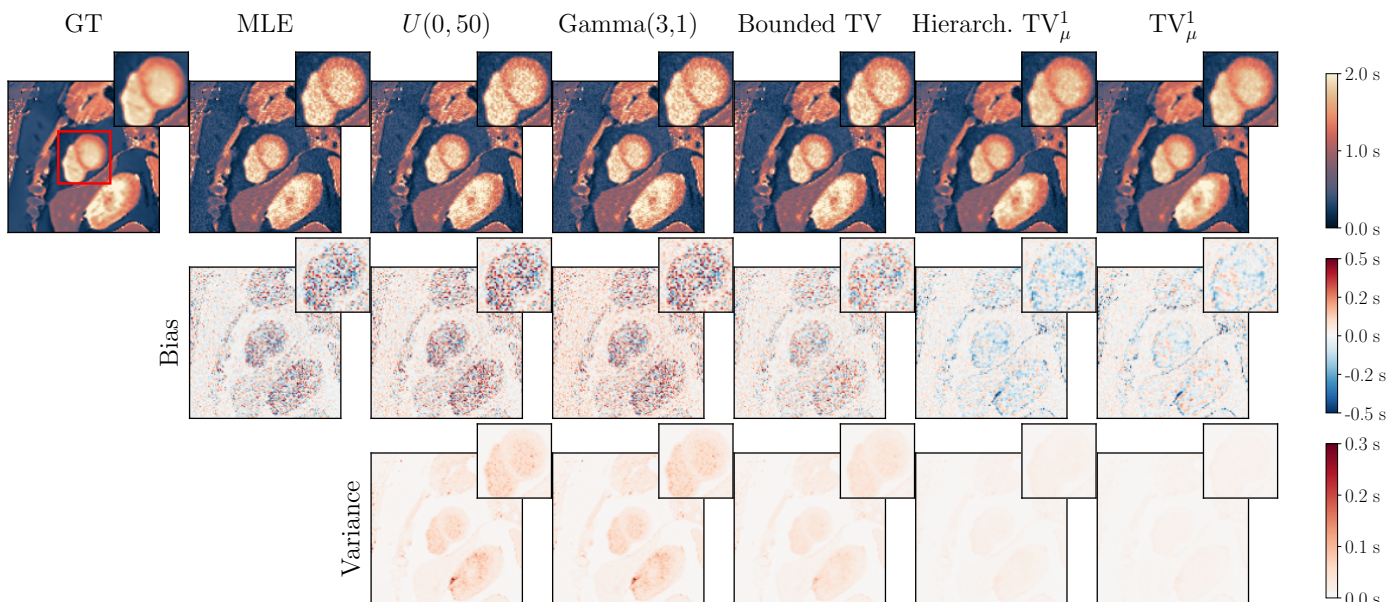


Figure 6: Comparison of mean T_1 , bias, and variance maps from different methods. The first row displays the GT T_1 map, the MLE point estimate, and the mean T_1 maps obtained from each Bayesian model, as labeled at the top of the figure. The second row presents the corresponding bias maps, while the third row shows the variance maps. Each subplot includes a zoomed-in view of a region (marked by the red square) in the top-right corner for detailed comparison.

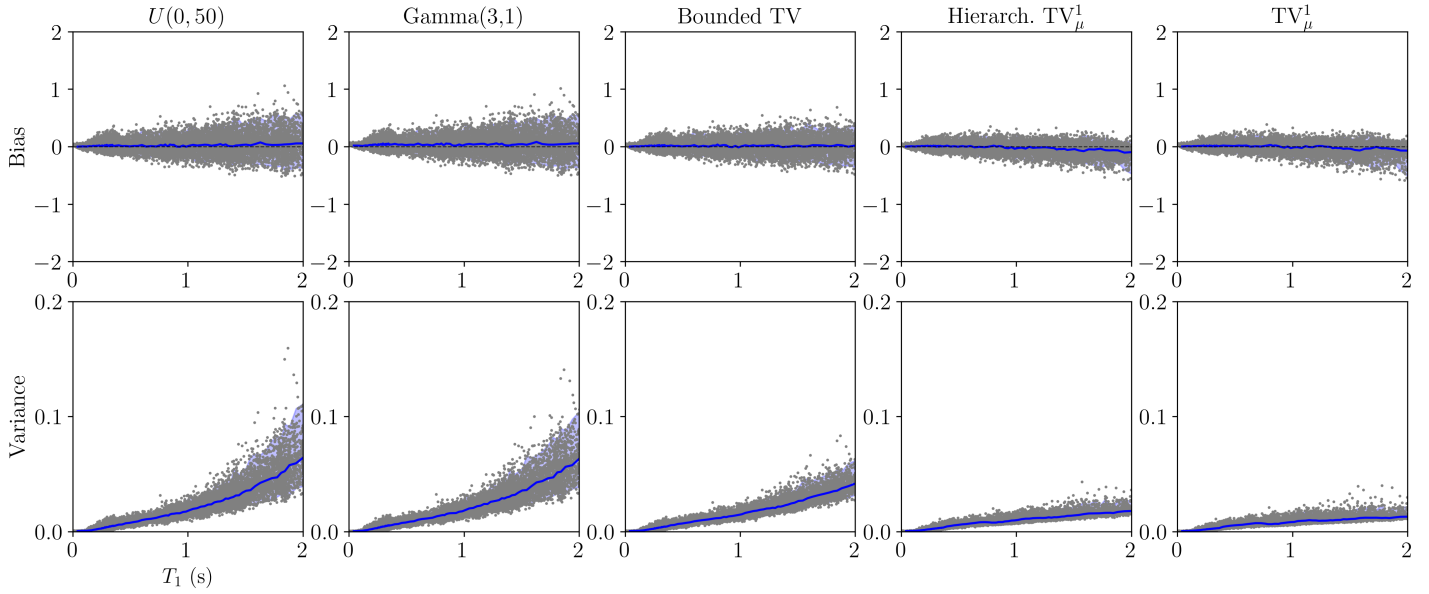


Figure 7: Bias and variance analysis for the Bayesian methods. The horizontal axis shows the GT T_1 values, while the vertical axis indicates the bias/variance observed at each voxel. The blue curves represent the average bias/variance for a certain GT T_1 value. The blue shaded area contains 95% of the data points.

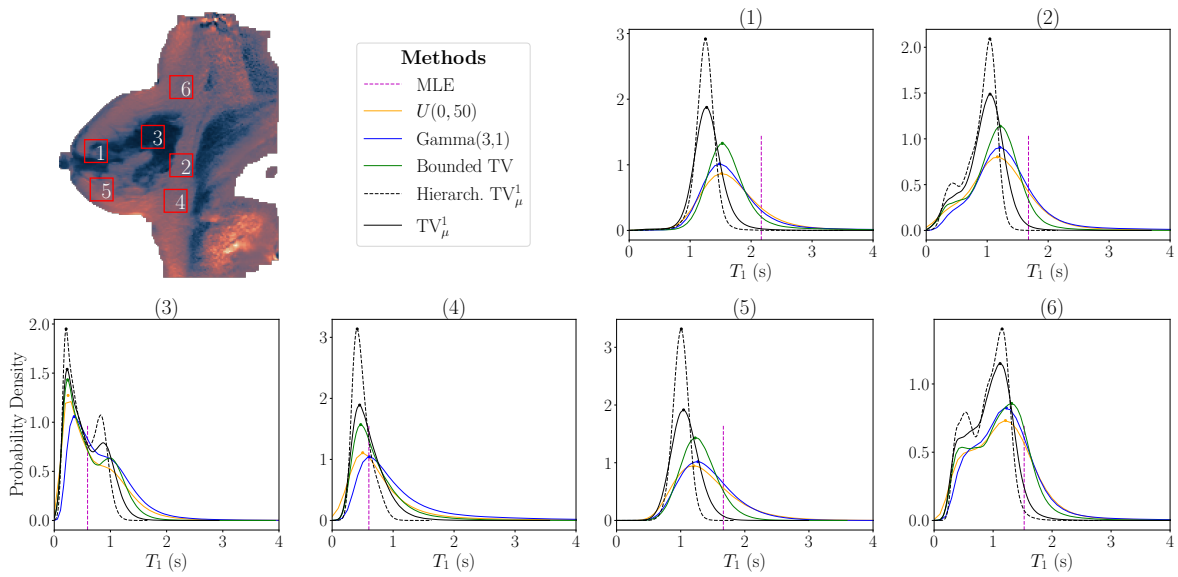


Figure 8: PDFs of T_1 values at selected ROIs (highlighted in red on the anatomical image, top-left). The estimated PDFs from different Bayesian methods are shown, along with the MLEs, for comparison.

Patch	MLE	$U(0, 50)$	$\text{Gamma}(3, 1)$	Bounded TV	Hierarch. TV_μ^1	TV_μ^1
Patch 1	2.17	1.52, 2.18 ± 4.83	1.48, 1.77 ± 1.63	1.53, 1.67 ± 1.11	1.25, $1.24 \pm \mathbf{0.34}$	1.27, 1.30 ± 0.53
Patch 2	1.68	1.17, 1.61 ± 4.40	1.21, 1.38 ± 1.60	1.22, 1.14 ± 0.87	1.04, $0.88 \pm \mathbf{0.53}$	1.05, 0.96 ± 0.64
Patch 3	0.59	0.25, 0.71 ± 1.22	0.36, 0.83 ± 1.05	0.24, 0.64 ± 0.80	0.21, $0.51 \pm \mathbf{0.55}$	0.23, 0.57 ± 0.66
Patch 4	0.60	0.50, 1.32 ± 4.87	0.61, 1.14 ± 1.87	0.47, 0.72 ± 0.82	0.41, $0.49 \pm \mathbf{0.31}$	0.45, 0.60 ± 0.53
Patch 5	1.67	1.20, 1.41 ± 1.00	1.27, 1.42 ± 0.88	1.23, 1.29 ± 0.59	1.01, $1.00 \pm \mathbf{0.25}$	1.04, 1.07 ± 0.43
Patch 6	1.53	1.21, 1.35 ± 3.24	1.22, 1.24 ± 1.38	1.31, 1.12 ± 1.00	1.15, $0.90 \pm \mathbf{0.63}$	1.12, 0.95 ± 0.72

Table 3: Quantitative comparison of ML and Bayesian estimates for each selected ROI. For each Bayesian model, the table reports the posterior mode, mean, and uncertainty as “mode, mean $\pm 2\sigma$ ”. All T_1 values are reported in seconds (s).

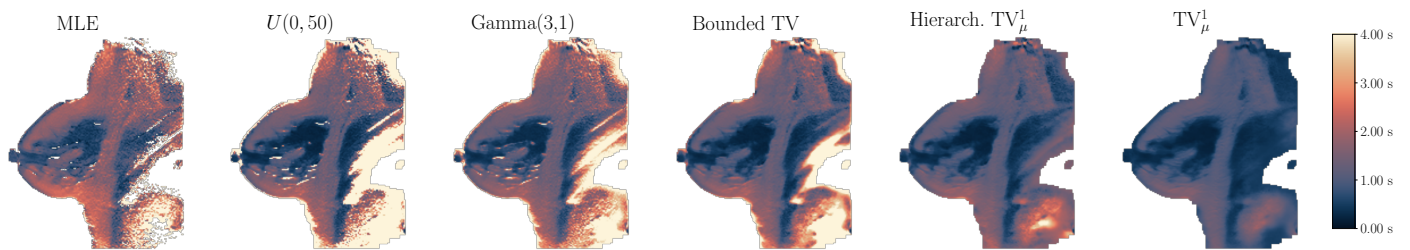


Figure 9: Comparison of mean T_1 maps from different methods as labeled at the top of the figure.

estimates compared to the Bounded TV method, reflecting a bias–variance trade-off.

6. Conclusion

This work introduces a new family of structured priors that integrate the TV function with ℓ_p norms, providing a proper and flexible prior distribution for Bayesian T_1 mapping. Unlike earlier TV-based formulations, the proposed construction preserves the desirable spatial regularization properties of TV while ensuring a normalizable prior that assigns meaningful probability across the full parameter space. Embedded within a Bayesian framework and combined with NUTS sampling, this formulation enables both stable parameter estimation and rigorous uncertainty quantification.

Across synthetic brain, cardiac, and real in-vivo breast datasets, the $TV-\ell_p$ prior consistently produced more concentrated posterior densities than the other methods, leading to reduced uncertainty, lower variance, and only a small negative bias. As a result, the parameter maps exhibit enhanced spatial coherence and more reliable credibility assessments, even under challenging noise conditions and limited data.

Overall, the results demonstrate that incorporating TV-based spatial structure together with ℓ_p norms within the Bayesian framework provides a robust foundation for accurate T_1 estimation and uncertainty quantification, and can be naturally extended to other qMRI modalities or to joint reconstruction–estimation strategies in future work. While our experiments focus on medical imaging, the method is not limited to this domain and can be applied to a wide range of problems that require structured signal recovery or spatially informed priors.

Acknowledgments

This work was supported by the Swedish Research Council (Vetenskapsrådet; grant number 2021-04810) and Lion’s Cancer Research Foundation in Northern Sweden (grant numbers LP 22-2319, LP 24-2367, and AMP 26-1265).

Ethical Standards

The work followed appropriate ethical standards in conducting research and writing the manuscript, following all applicable laws and regulations regarding treatment of animals or human subjects.

Conflicts of Interest

We declare we don’t have conflicts of interest.

Data availability

All data used in this study are either publicly available or were synthetically generated as described in the manuscript. The synthetic brain and cardiac T_1 mapping datasets can be fully reproduced using the simulation procedures provided in the Data section. The real breast MRI data were obtained from the publicly accessible *QIN-BREAST-02* collection available through The Cancer Imaging Archive (TCIA). No additional restrictions apply to the availability or reproduction of the datasets used in this work. The code associated with this work is publicly available at: <https://github.com/Disi2022/Proper-TV-Structured-Priors>

References

- Artin Armagan, David B. Dunson, and Jaeyong Lee. Generalized double Pareto shrinkage. *Statistica Sinica*, 23(1):119, 2013.
- Simon Baudrexel, Ulrike Nöth, Jan-Rüdiger Schüre, and Ralf Deichmann. T1 mapping with the variable flip angle technique: a simple correction for insufficient spoiling of transverse magnetization. *Magnetic resonance in medicine*, 79(6):3082–3092, 2018.
- Kirsten Miriam Becker, Jeanette Schulz-Menger, Tobias Schaeffter, and Christoph Kolbitsch. Multi-parametric cardiac MRI for T1 mapping and cine imaging using iterative model-based image reconstruction. In *Proceedings*

- of the *International Society for Magnetic Resonance in Medicine*, volume 25, page 2722, 2017.
- Quinten Beirinckx, Ben Jeurissen, Michele Nicastrò, Dirk H. J. Poot, Marleen Verhoye, Arnold J. den Dekker, and Jan Sijbers. Model-based super-resolution reconstruction with joint motion estimation for improved quantitative MRI parameter mapping. *Computerized Medical Imaging and Graphics*, 100:102071, 2022.
- Alexandre Carré, Guillaume Klausner, Myriam Edjlali, Marvin Lerousseau, Jade Briend-Diop, Roger Sun, Samy Ammari, Sylvain Reuzé, Emilie Alvarez Andres, Théo Estienne, Stéphane Niyoteka, Enzo Battistella, Maria Vakalopoulou, Frédéric Dhermain, Nikos Paragios, Eric Deutsch, Catherine Oppenheim, Johan Pallud, and Charlotte Robert. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Scientific reports*, 10(1):12340, 2020.
- Carlos M. Carvalho, Nicholas G. Polson, and James G. Scott. The horseshoe estimator for sparse signals. *Biometrika*, pages 465–480, 2010.
- Kenner A. Christensen, David M. Grant, Edward M. Schulman, and Cheves Walling. Optimal determination of relaxation times of fourier transform nuclear magnetic resonance. determination of spin-lattice relaxation times in chemically polarized species. *The Journal of Physical Chemistry*, 78(19):1971–1977, 1974.
- Sean C. L. Deoni, Brian K. Rutt, and Terry M. Peters. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 49(3):515–526, 2003.
- Edgar A. DeYoe, Peter Bandettini, Jay Neitz, David Miller, and Paula Winans. Functional magnetic resonance imaging (fMRI) of the human brain. *Journal of neuroscience methods*, 54(2):171–187, 1994.
- Selim Esedoğlu and Stanley J Osher. Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. *Communications on Pure and Applied Mathematics*, 57(12), 2004.
- Neville D. Gai and John A. Butman. Modulated repetition time look-locker (MORTLL): A method for rapid high resolution three-dimensional T1 mapping. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 30(3):640–648, 2009.
- Andrew Gelman and Donald B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 1992.
- Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995.
- Daniel Gräfe, Jens Frahm, Andreas Merkschlager, Dirk Voit, and Franz Wolfgang Hirsch. Quantitative T1 mapping of the normal brain from early infancy to adulthood. *Pediatric Radiology*, 51:450–456, 2021.
- Cristina Granziera, Jens Wuerfel, Frederik Barkhof, Massimiliano Calabrese, Nicola De Stefano, Christian Enzinger, Nikos Evangelou, Massimo Filippi, Jeroen J. G. Geurts, Daniel S. Reich, Stefan Rocca, Maria A. amd Ropele, Àlex Rovira, Pascal Sati, Ahmed T. Toosy, Hugo Vrenken, Claudia A. M. Gandini Wheeler-Kingshott, Ludwig Kappos, and MAGNIMS Study Group. Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, 144(5):1296–1311, 2021.
- Fouad Hadj-Seleem, Tommy Löfstedt, Elvis Dohmatob, Vincent Frouin, Mathieu Dubois, Vincent Guillemot, and Edouard Duchesnay. Continuation of nesterov's smoothing for regression with structured sparsity in high-dimensional neuroimaging. *IEEE Transactions on Medical Imaging*, 37(11):2403–2413, 2018.
- Max Hellström. *Noise reduction and uncertainty estimation in quantitative Magnetic Resonance Imaging*. PhD thesis, Umeå University, 2025.
- Matthew D. Hoffman and Andrew Gelman. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, 2014.
- Shuai Huang, James J. Lah, Jason W Allen, and Deqiang Qiu. Accelerated model-based T1, T2* and proton density mapping using a Bayesian approach with automatic hyperparameter estimation. *Magnetic Resonance in Medicine*, 93(2):563–583, 2025.
- Hemant Ishwaran and J. Sunil Rao. Spike and slab variable selection: frequentist and Bayesian strategies. *The Annals of Statistics*, 33(2), 2005.
- Hernán Jara, Osamu Sakai, Ezequiel Farrher, Ana-Maria Oros-Peusquens, N. Jon Shah, David C. Alsop, and Kathryn E. Keenan. Primary multiparametric quantitative brain MRI: state-of-the-art relaxometric and proton density mapping techniques. *Radiology*, 305(1):5–18, 2022.

- David E. Jones, Robert N. Trangucci, and Yang Chen. Quantifying observed prior impact. *Bayesian Analysis*, 17(3):737–764, 2022.
- Elizaveta Lavrova, Emilie Lommers, Henry C. Woodruff, Avishek Chatterjee, Pierre Maquet, Eric Salmon, Philippe Lambin, and Christophe Phillips. Exploratory radiomic analysis of conventional vs. quantitative brain MRI: toward automatic diagnosis of early multiple sclerosis. *Frontiers in neuroscience*, 15:679941, 2021.
- Johnathan V. Le, Jason K. Mendes, Nicholas McKibben, Brent D. Wilson, Mark Ibrahim, Edward V. R. DiBella, and Ganesh Adluru. Accelerated cardiac T1 mapping with recurrent networks and cyclic, model-based loss. *Medical physics*, 49(11):6986–7000, 2022.
- Disi Lin, Anders Garpebring, and Tommy Löfstedt. A proper structured prior for Bayesian T_1 Mapping. In C. H. Sudre et al., editors, *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging (UNSURE 2025)*, volume 16166 of *Lecture Notes in Computer Science*. Springer, Cham, 2025.
- Dong C. Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1), 1989.
- Tommy Löfstedt, Max Hellström, Mikael Bylund, and Anders Garpebring. Bayesian non-linear regression with spatial priors for noise reduction and error estimation in quantitative MRI with an application in T1 estimation. *Physics in Medicine & Biology*, 65(22):225036, 2020.
- David C. Look and Donald R. Locker. Time saving in measurement of NMR and EPR relaxation times. *Review of scientific instruments*, 41(2):250, 1970.
- Daniel R. Messroghli, Aleksandra Radjenovic, Sebastian Kozerke, David M. Higgins, Mohan U. Sivananthan, and John P. Ridgway. Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 52(1):141–146, 2004.
- Vincent Michel, Alexandre Gramfort, Gaël Varoquaux, Evelyn Eger, and Bertrand Thirion. Total variation regularization for fmri-based prediction of behavior. *IEEE transactions on medical imaging*, 30(7):1328–1340, 2011.
- Shraddha Pandey, A. David Snider, Wilfrido A. Moreno, Harshan Ravi, Ali Bilgin, and Natarajan Raghunand. Joint total variation-based reconstruction of multiparametric magnetic resonance images for mapping tissue types. *NMR in Biomedicine*, 34(12):e4597, 2021.
- Christine Preibisch and Ralf Deichmann. Influence of RF spoiling on the stability and accuracy of T1 mapping based on spoiled FLASH with varying flip angles. *Magnetic Resonance in Medicine*, 61(1):125–135, 2009.
- I. L. Pykett, B. R. Rosen, F. S. Buonanno, and T. J. Brady. Measurement of spin-lattice relaxation times in nuclear magnetic resonance imaging. *Physics in Medicine & Biology*, 28(6):723, 1983.
- Prabhakar Shantha Rajiah, Christopher J. François, and Tim Leiner. Cardiac MRI: state of the art. *Radiology*, 307(3):e223008, 2023.
- Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.
- Banafshe Shafieizargar, Riway Byanju, Jan Sijbers, Stefan Klein, Arnold J. den Dekker, and Dirk H. J. Poot. Systematic review of reconstruction techniques for accelerated quantitative MRI. *Magnetic Resonance in Medicine*, 90(3):1172–1208, 2023.
- Bernard W. Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018.
- Kewei Tang, Zhixun Su, Jie Zhang, Lihong Cui, Wei Jiang, Xiaonan Luo, and Xiyan Sun. Bayesian rank penalization. *Neural Networks*, 116:246–256, 2019.
- Temel Tirkes, Xuandong Zhao, Chen Lin, Alex Jordan Stuckey, Liang Li, Shivriman Giri, and Dominik Nickel. Evaluation of variable flip angle, MOLLI, SASHA, and IR-SNAPSHOT pulse sequences for T1 relaxometry and extracellular volume imaging of the pancreas and liver. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 32:559–566, 2019.
- Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27:1413–1432, 2017.
- Patrick Virtue and Michael Lustig. The empirical effect of Gaussian noise in undersampled MRI reconstruction. *Tomography*, 3(4):211, 2017.
- Xiaoqing Wang, Sebastian Rosenzweig, Volkert Roeloffs, Moritz Blumenthal, Nick Scholand, Zhengguo Tan, H Christian M Holme, Christina Unterberg-Buchwald, Rabea Hinkel, and Martin Uecker. Free-breathing myocardial T1 mapping using inversion-recovery radial FLASH and motion-resolved model-based reconstruction. *Magnetic resonance in medicine*, 89(4):1368–1384, 2023.

Sumio Watanabe and Manfred Opper. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of machine learning research*, 11(12), 2010.

Demi Wekking, Michele Porcu, Pushpamali De Silva, Luca Saba, Mario Scartozzi, and Cinzia Solinas. Breast MRI: clinical indications, recommendations, and future applications in breast cancer diagnosis. *Current oncology reports*, 25(4):257–267, 2023.

TE Yankeelov, GS Karczmar, and Abramson RG. Data from QIN-BREAST-02[Dataset]. The Cancer Imaging Archive, <https://doi:10.7937/TCIA.2019.4cfm06rr>, 2019.

Tao Zhang, John M. Pauly, and Ives R. Levesque. Accelerating parameter mapping with a locally low rank constraint. *Magnetic resonance in medicine*, 73(2):655–661, 2015.