SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

A perfect reconstruction property for PDE-constrained total-variation minimization with application in Quantitative Susceptibility Mapping

K. Bredies D. Vicente

SFB-Report No. 2017-001 January 2017

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

A perfect reconstruction property for PDE-constrained total-variation minimization with application in Quantitative Susceptibility Mapping

Kristian Bredies David Vicente*

January 23, 2017

Abstract

We study the recovery of piecewise constant functions of finite bounded variation (BV) from their image under a linear partial differential operator with unknown boundary conditions. It is shown that minimizing the total variation (TV) semi-norm subject to the associated PDE-constraints yields perfect reconstruction up to a global constant under a mild geometric assumption on the jump set of the function to reconstruct. The proof bases on establishing a structural result about the jump set associated with BV-solutions of the homogeneous PDE. Furthermore, we show that the geometric assumption is satisfied up to a negligible set of orthonormal transformations. The results are then applied to Quantitative Susceptibility Mapping (QSM) which can be formulated as solving a two-dimensional wave equation with unknown boundary conditions. This yields in particular that total variation regularization is able to reconstruct piecewise constant susceptibility distributions, explaining the high-quality results obtained with TV-based techniques for QSM.

1 Introduction

We investigate how it is possible to solve a linear partial differential equation in a domain lacking boundary conditions but with the a-priori knowledge that the desired solution is a piecewise constant function of bounded variation, i.e., the gradient is "sparse" in a certain sense. To make the problem precise, let $d \geq 1$ be a spatial dimension, let $\alpha = (\alpha_1, \ldots, \alpha_d)$ denote a multi-index in $\mathbb{N}^d$ and let $\mathcal{A}$ be the linear differential operator of order $k \geq 0$ according to

$$\mathcal{A} = \sum_{\alpha\colon |\alpha| \leq k} a_\alpha \partial^\alpha,$$

where $(a_\alpha)_\alpha$ is a family of constant real coefficients such that $a_\alpha \neq 0$ for an $|\alpha| = k$. Further, let $\Omega \subset \mathbb{R}^d$ be a bounded domain and ψ a distribution in Ω. We assume that there exists an unknown piecewise constant function $\chi \in \mathrm{BV}(\Omega)$ which satisfies $\mathcal{A}\chi = \psi$ in the distributional sense and aim at finding χ given ψ.

For that purpose, we consider solutions of the minimization problem

$$\min_{\chi \in \mathrm{BV}(\Omega)} \mathrm{TV}(\chi) \qquad \text{subject to} \qquad \mathcal{A}\chi = \psi. \tag{$\mathcal{P}$}$$

As this is a convex problem (minimization of a convex functional on a convex constraint set), it is accessible regarding both analysis and numerical solution. The intuition behind this approach is the following. As $\mathcal{A}$ is a linear differential operator, the constraint set is given by $\{\chi + u\colon u \in$

*Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, 8010 Graz, Austria. Email: kristian.bredies@uni-graz.at, david.vicente@uni-graz.at. The Institute of Mathematics and Scientific Computing is a member of NAWI Graz (www.nawigraz.at) and BioTechMed Graz (www.biotechmedgraz.at).

$\mathrm{BV}(\Omega),\ \mathcal{A}u = 0\}$. The exact solution χ is piecewise constant and of bounded variation meaning that its gradient vanishes up to a set of finite $(d-1)$-dimensional Hausdorff measure, i.e., is "sparse". The gradient of a solution $u \in \mathrm{BV}(\Omega)$ of the homogeneous equation $\mathcal{A}u = 0$ now can be expected to possess enough structure to make it distinguishable from $\nabla\chi$ in terms of giving an increase of the TV functional when added to χ. The aim of this paper is to show that this is indeed true under generic geometrical assumptions on the interfaces of χ.

The main result about the minimization problem ($\mathcal{P}$) recovering χ (possibly up to a global constant) is stated as Theorem 3.5. It relies on the following auxiliary result on the structure of the singularities that can occur for $u \in \mathrm{BV}(\Omega)$ such that $\mathcal{A}u = 0$. Introducing the homogeneous polynomial P with degree k in $\xi_1, \ldots, \xi_d$ according to

$$P(\xi) = \sum_{\alpha\colon |\alpha|=k} a_\alpha \xi^\alpha.$$

it holds that u satisfies

$$P(\nu_u(x)) = 0 \qquad \text{for } \mathcal{H}^{d-1}\text{-almost every } x \in J_u,$$

where J_u is the jump set of u, ν_u is its measure-theoretic normal vector and $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure, see Theorem 3.3. This may be interpreted as a generalization of the notion of wavefront in BV where the wavefront is supported in J_u in the spatial domain and in the direction of ν_u in the frequency domain. With these notions, the condition for perfect reconstruction (up to a global constant) in Theorem 3.5 then reads as the geometrical condition

$$\mathcal{H}^{d-1}(\{x \in J_\chi\colon P(\nu_\chi(x)) = 0\}) = 0.$$

For a given piecewise constant $\chi \in \mathrm{BV}(\Omega)$, this condition is generic with respect to orthonormal transformations, or, more precisely, it is satisfied for $\chi \circ Q$ where $Q \in \mathrm{O}_d(\mathbb{R})$, i.e, in the orthogonal group, up to a Lebesgue null set in $\mathrm{O}_d(\mathbb{R})$, see Theorem 3.4. Thus, in view of the Haar probability measure on $\mathrm{O}_d(\mathbb{R})$, the perfect recovery property holds almost surely.

To demonstrate the practical relevance of these results, we also discuss the application to Quantitative Susceptibility Mapping (QSM), a recently established Magnetic Resonance Imaging (MRI) technique which aims at providing the spatial susceptibility distribution χ of tissues inside the human body [1]. The quantitative determination of magnetic susceptibility is an important issue to characterize diseased tissue. For instance, neurodegenerative diseases such as Alzheimer's or Parkinson's cause deposits of iron, which are very responsive to magnetic fields [2]. QSM is made possible by considering the phase of the measured complex-valued MR image which corresponds, up to a factor, to the field inhomogeneity b_0 caused by the susceptibility distribution χ. The latter obeys the convolution relation $\chi * \delta = b_0$, where δ is the component of the magnetic dipole kernel along the main magnetic field. The problem can, however, also be written in terms of the two-dimensional wave equation

$$\Box\chi = -\Delta b_0 \qquad \text{in } \Omega,$$

where $\Omega \subset \mathbb{R}^3$ corresponds to the region of interest, Δ is the Laplacian in $\mathbb{R}^3$ and $\Box$ the wave operator with the main magnetic field axis z as time and the two transversal axes x and y as spatial variables:

$$\Box = \frac{2}{3}\frac{\partial^2}{\partial z^2} - \frac{1}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2}.$$

The problem is ill-posed in the sense that the lack of boundary conditions leads to a non-uniqueness of solutions that becomes apparent in terms of artifacts. Recently developed single-step reconstruction algorithms for QSM [3, 4] aim at resolving this non-uniqueness by employing regularizing functionals that allow for edges such as TV or the total generalized variation (TGV) of order 2 [5]. In this context, TV-regularization for the noise-free case corresponds exactly to a problem of type ($\mathcal{P}$). The perfect reconstruction property established in this paper thus implies that such an approach is able to generically reconstruct piecewise constant susceptibility distributions and might explain the surprising efficiency of TV-regularization for this application.

The minimization problem ($\mathcal{P}$) can be interpreted as an optimization problem with partial differential equations. Research on the latter is currently a very active branch of applied mathematics and there is a vast amount of literature available [6, 7, 8, 9]. However, using the total variation as a penalty term in this context seems to be addressed only recently and only few works are available today. Indeed, to the best knowledge of the authors, this is the first time that perfect reconstruction properties are discussed for PDE-constrained TV-minimization. These types of results are nevertheless closely connected with the theory of compressed sensing [10]: There, minimizing, in finite dimensions, some sparsifying functional (typically the 1-norm) subject to affine linear constraints allows for perfect reconstruction provided that the solution is *sparse*, i.e., only a few coefficients are non-zero. In this context, the maximal number of non-zero coefficients for which this can be guaranteed is associated with the so-called *restricted isometry property* (RIP) for the linear operator associated with the affine linear constraint. For many types of random matrices, the probability for the latter turns out to be overwhelming.

Let us point out similarities and differences between the perfect reconstruction properties in compressed sensing and this paper. In the present paper, the sparsifying functional is the total variation semi-norm rather than the 1-norm and the piecewise constancy assumption corresponds to a sparsity assumption in the continuous setting. Indeed, in the discrete setting, results concerning perfect reconstruction by minimizing a discrete TV-functional are also available [10, 11]. However, to the best knowledge of the authors, there are no such results for the continuous setting considered in this paper. Nevertheless, while compressed sensing is mainly studied in the finite-dimensional setting, there are works dealing with perfect reconstruction in infinite-dimensional spaces, such as spaces associated with representations in a wavelets basis or a frame [12, 13, 14] or spaces of Radon measures [15, 16]. Furthermore, the RIP in compressed sensing, which is a structural assumption, can be connected with the assumption that the linear constraints in this paper are a partial differential equation. For the latter, however, the concept of sparsity does not play a role for the perfect reconstruction result, but rather a geometrical assumption on the data to recover. Nevertheless, similarities still arise in the point that perfect reconstruction is very likely — a fact which is reflected in this paper by the genericity of the geometrical assumption and by the overwhelming probability of certain random matrices having the RIP in compressed sensing.

Finally, let us comment on the two auxiliary theorems leading to the perfect reconstruction result which are also interesting on their own. First, Theorem 3.3 can be interpreted as a statement about the structure of wavefronts associated with the homogeneous PDE $\mathcal{A}u = 0$. Such wavefronts, i.e., singularities of the solutions of PDEs, are extensively studied throughout the literature. In this context, there are different notions of singularities such as the classical and the Sobolev sense [17]. As already mentioned, this work refers to singularities in BV, i.e., jump singularities in the measure-theoretic sense. During the writing of this manuscript, the authors became aware of a recent similar result concerning singularities in the measure sense, i.e., measures that are singular with respect to the Lebesgue measure [18, Theorem 1.1]. It can be expected that Theorem 3.3 follows from this result, however, as the latter deals with the jump part of BV functions (and not its Cantor part), the proof which is presented here is significantly simpler and more direct. Furthermore, Theorem 3.4 might be interpreted as a result on how normal vectors associated with the jump set of a BV-function are distributed. Its proof utilizes elementary measure theory and real algebraic geometry [19], the latter as the critical set of normal directions $\{\xi \in \mathbb{S}^{d-1} \colon P(\xi) = 0\}$ defines a (semi-) algebraic set. The result is mainly owed to the specific structure of semi-algebraic sets, exploiting the fact that such sets can always be decomposed into a finite number of smooth manifolds, each with a fixed dimension.

The paper is organized as follows. Section 2 is devoted to the introduction of some results on BV-functions we utilize in the course of the paper. In Section 3, we introduce our TV-minimization model for the recovery of piecewise constant functions and its adapted functional setting. We briefly verify that the associated minimization problem is well-posed. Then, we prove our main result: the perfect reconstruction property for piecewise constant functions in BV. As announced, we need to two auxiliary results. Section 4 is devoted to prove the first one which deals with the localization of the jump set for functions in BV which are in the kernel of the wave operator. Eventually, Section 5 is devoted to the proof of the second one which establishes that the geometric condition that is sufficient for perfect reconstruction is generic with respect to rotations of a fixed BV-function. We present the application to QSM in Section 6 while conclusions and an outlook is given in Section 7.

2 Preliminaries

Throughout this paper, we assume that $\Omega \subset \mathbb{R}^d$, $d \geq 1$ is a bounded Lipschitz domain, i.e., a non-empty, open, connected and bounded set with Lipschitz boundary.

The following collects the basic results about functions of bounded variation we will need in order to prove the prefect reconstruction property, the monograph [20] might serve as a detailed reference for these results. We adopt the following notations: $B_\varepsilon(x)$ denotes the ball with center $x \in \mathbb{R}^d$ and radius $\varepsilon > 0$, for any $\nu \in \mathbb{S}^{d-1}$, $B_\varepsilon^+(x,\nu)$ (resp. $B_\varepsilon^-(x,\nu)$) denotes the half ball $\{y \in B_\varepsilon(x) \colon (y-x)\cdot\nu > 0\}$ (resp. $\{y \in B_\varepsilon(x) \colon (y-x)\cdot\nu < 0\}$), the Lebesgue measure in $\mathbb{R}^d$ is denoted by $\mathcal{L}^d$ and the Hausdorff measure with dimension $s \geq 0$ by $\mathcal{H}^s$. Further, $\mu \llcorner A$ denotes the trace of a measure μ onto a measurable set A, that is, $\mu \llcorner A(\cdot) = \mu(A \cap \cdot)$. The space of Radon measure in Ω is denoted by $\mathcal{M}(\Omega)$, its norm is given by $\|\mu\|_{\mathcal{M}} = |\mu|(\Omega)$ where $|\mu|$ is the total-variation measure associated with $\mu \in \mathcal{M}(\Omega)$. A function $u \in L^1(\Omega)$ is of bounded variation, denoted by $u \in \mathrm{BV}(\Omega)$, if its distributional derivative Du is a vectorial Radon measure, *i.e.*, $Du \in \mathcal{M}(\Omega;\mathbb{R}^d)$. The total variation of u is defined by the total mass of Du and is denoted by $\mathrm{TV}(u) = \|Du\|_{\mathcal{M}}$. For $u \in \mathrm{BV}(\Omega)$, the derivative of Du satisfies the following decomposition (see [20, Section 3.9])

$$Du = \nabla u \mathcal{L}^d + (u^+ - u^-)\nu_u \mathcal{H}^{d-1} \llcorner J_u + D^c u, \tag{2.1}$$

where ∇u is the density of Du with respect to the Lebesgue measure, J_u is the jump set which is the collection of points $x \in \Omega$ for which there exist lower and upper approximate limits $u^-(x) < u^+(x)$ and a measure-theoretic normal vector $\nu_u(x)$ such that

$$\lim_{\varepsilon \to 0^+} \fint_{B_\varepsilon^+(x,\nu_u(x))} |u - u^+(x)| \,\mathrm{d}y = 0, \qquad \lim_{\varepsilon \to 0^+} \fint_{B_\varepsilon^-(x,\nu_u(x))} |u - u^-(x)| \,\mathrm{d}y = 0,$$

and $D^c u$ is the Cantor part. At almost any point of its jump set, a function in BV satisfies the following blow-up result.

Lemma 2.1. *For $u \in \mathrm{BV}(\Omega)$, $x \in \Omega$ and $\varepsilon > 0$, we write $u_{x,\varepsilon}(y) = u(x + \varepsilon y)$ for $y \in \mathbb{R}^d$ where the right-hand side is defined and $u_{x,\varepsilon}(y) = 0$ else. Then, for $\mathcal{H}^{d-1}$-a.e. $x \in J_u$, $u_{x,\varepsilon}$ is weakly-$\star$ convergent in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$ when ε goes to 0^+ and its limit ω_x is given by*

$$\omega_x(y) = \begin{cases} u^+(x) & \text{if } \nu_u(x)\cdot y \geq 0, \\ u^-(x) & \text{if } \nu_u(x)\cdot y < 0. \end{cases}$$

Proof. We fix $x \in J_u$, $R > 0$ and aim at proving the result in $\mathrm{BV}(B_R(0))$. For $\varepsilon > 0$ sufficiently small, the function $u_{x,\varepsilon}$ is determined in $B_R(0)$ by evaluation of u in Ω. Writing $\nu = \nu_u(x)$ we have

$$\int_{B_R(0)} |u_{x,\varepsilon} - \omega_x| \,\mathrm{d}y = \varepsilon^{-d} \left(\int_{B_{\varepsilon R}^+(x,\nu)} |u - u^+(x)| \,\mathrm{d}y + \int_{B_{\varepsilon R}^-(x,\nu)} |u - u^-(x)| \,\mathrm{d}y \right) \to 0$$

as $\varepsilon \to 0^+$ since x is in the jump set J_u and $\mathcal{L}^d(B^+_{\varepsilon R}(x,\nu)) = \mathcal{L}^d(B^-_{\varepsilon R}(x,\nu)) \backsim \varepsilon^d$. This establishes $\lim_{\varepsilon\to 0^+} u_{x,\varepsilon} = \omega_x|_{B_R(0)}$ in $L^1(B_R(0))$.

Moreover, the values $|Du_{x,\varepsilon}|(B_R(0))$ are bounded as $\varepsilon \to 0^+$. To see this, note that for $\varphi \in \mathcal{D}(B_R(0))$, i.e., φ a test function in $B_R(0)$, $\varphi_\varepsilon(y) = \varphi((y-x)/\varepsilon)$ and $D_i u$, $D_i u_{x,\varepsilon}$ denoting the i-th component of Du, $Du_{x,\varepsilon}$, respectively, we have

$$\begin{aligned}\int_\Omega \varphi_\varepsilon \,\mathrm{d}D_i u = -\int_\Omega u(y)\left(\frac{\partial}{\partial x_i}\varphi_\varepsilon\right)(y)\,\mathrm{d}y = -\frac{1}{\varepsilon}\int_\Omega u(x)\frac{\partial \varphi_\varepsilon}{\partial x_i}\left(\frac{y-x}{\varepsilon}\right)\,\mathrm{d}y \\ = -\varepsilon^{d-1}\int_{(\Omega-x)/\varepsilon} u_{x,\varepsilon}(y)\frac{\partial\varphi}{\partial x_i}(y)\,\mathrm{d}y = \varepsilon^{d-1}\int_{\mathbb{R}^d}\varphi\,\mathrm{d}D_i u_{x,\varepsilon},\end{aligned} \tag{2.2}$$

so by the definition of the total variation measure, $|Du_{x,\varepsilon}|(B_R(0)) = |Du|(B_{\varepsilon R}(x))/\varepsilon^{d-1}$. According to [20, Theorem 2.56, (2.40)], for $\mathcal{H}^{d-1}$-almost every $x \in J_u$, the right-hand side stays bounded, hence, the family $(u_{x,\varepsilon})_\varepsilon$ is bounded in $\mathrm{BV}(B_R(0))$ as $\varepsilon \to 0^+$. By compactness in BV [20, Theorem 3.23], there is a sequence $(\varepsilon_n)_n$ with $\varepsilon_n \to 0^+$ as $n \to \infty$ for which $(u_{x,\varepsilon_n})_n$ converges to $v \in \mathrm{BV}(B_R(0))$ in the weak-$\star$ sense as $n \to \infty$. As we also have L^1-convergence, it follows $v = \omega_x|_{B_R(0)}$. The usual subsequence argument then gives the stated weak-$\star$ convergence as $\varepsilon \to 0^+$, proving the lemma. □

3 The perfect reconstruction property

In this section we study the minimization problem ($\mathcal{P}$) and prove its ability to provide perfect reconstruction for piecewise constant functions of bounded variation after stating two auxiliary theorems that are proven in Sections 4 and 5. In the following, $\mathcal{A}$ is a linear differential operator of order $k \geq 0$ with constant coefficients:

$$\mathcal{A} = \sum_{|\alpha|\leq k} a_\alpha \partial^\alpha, \qquad a_\alpha \neq 0 \text{ for some } |\alpha| = k. \tag{3.1}$$

The formal adjoint is given by $\mathcal{A}^* = \sum_{|\alpha|\leq k}(-1)^{|\alpha|}a_\alpha\partial^\alpha$. Moreover, let ψ be a distribution in Ω. By definition of the distributional derivative, $\mathcal{A}\chi = \psi$ holds for $\chi \in \mathrm{BV}(\Omega)$ in the distributional sense if and only if

$$\int_\Omega \chi\mathcal{A}^*\varphi\,\mathrm{d}x = \psi(\varphi) \qquad \text{for all } \varphi \in \mathcal{D}(\Omega),$$

where $\mathcal{D}(\Omega)$ is the space of test functions on Ω. Minimizing the total variation subject to this constraint, i.e., solving problem ($\mathcal{P}$), amounts to

$$\min_{\chi\in\mathrm{BV}(\Omega)} \mathrm{TV}(\chi) \qquad \text{subject to} \qquad \int_\Omega \chi\mathcal{A}^*\varphi\,\mathrm{d}x = \psi(\varphi) \qquad \text{for all } \varphi\in\mathcal{D}(\Omega). \tag{3.2}$$

The problem is well-posed if the constraint set is non-empty.

Proposition 3.1. *Let ψ be a distribution in Ω such that there is a $\chi \in \mathrm{BV}(\Omega)$ with $\mathcal{A}\chi = \psi$ in the distributional sense. Then, the minimization problem ($\mathcal{P}$) admits a solution with finite total variation.*

Proof. Obviously, the functional in (3.2) admits finite values, so let $(\chi_n)_n$ be a minimizing sequence for TV satisfying the linear constraints in $\mathrm{BV}(\Omega)$. We set $\tilde{\chi}_n = \chi_n - m_n$, where m_n is the mean value of χ_n in Ω.

First, suppose that $a_0 = 0$. Then, each $\tilde{\chi}_n$ satisfies also $\mathcal{A}\tilde{\chi}_n = \psi$ in the distributional sense and $(\tilde{\chi}_n)_n$ is still a minimizing sequence for TV. According to the Poincaré inequality for functions of bounded variation [20, Theorem 3.50], the sequence $(\tilde{\chi}_n)_n$ is bounded in $L^1(\Omega)$ and thus, it is also bounded in $\mathrm{BV}(\Omega)$. So, we may extract a subsequence, still denoted by $(\tilde{\chi}_n)_n$, which is weak-$\star$ convergent to some $\chi \in \mathrm{BV}(\Omega)$. As TV is sequentially lower semi-continuous with respect to its weak-$\star$ topology, we get

$$\mathrm{TV}(\chi) \leq \liminf_n \mathrm{TV}(\tilde{\chi}_n).$$

In particular, $(\tilde{\chi}_n)_n$ converges in $L^1(\Omega)$, so passing to the limit for $\varphi \in \mathcal{D}(\Omega)$ gives

$$\int_\Omega \chi \mathcal{A}^* \varphi \, \mathrm{d}x = \lim_{n\to\infty} \int_\Omega \tilde{\chi}_n \mathcal{A}^* \varphi \, \mathrm{d}x = \psi(\varphi),$$

so $\mathcal{A}\chi = \psi$ in the distributional sense which implies that χ is indeed a solution of ($\mathcal{P}$).

Finally, let us discuss the case $a_0 \neq 0$. We may choose a test function $\varphi \in \mathcal{D}(\Omega)$ such that $\int_\Omega \varphi \, \mathrm{d}x \neq 0$. As $\mathcal{A}\chi_n = \psi$, this gives

$$\int_\Omega \tilde{\chi}_n \mathcal{A}^* \varphi \, \mathrm{d}x = \int_\Omega \chi_n \mathcal{A}^* \varphi \, \mathrm{d}x - m_n \int_\Omega \mathcal{A}^* \varphi \, \mathrm{d}x = \psi(\varphi) - a_0 m_n \int_\Omega \varphi \, \mathrm{d}x.$$

Analogous to the case $a_0 = 0$, one gets that the sequence $(\tilde{\chi}_n)_n$ is bounded in $\mathrm{BV}(\Omega)$, so by the above identity, the sequence $(m_n)_n$ is bounded. Hence, $(\chi_n)_n$ is bounded in $\mathrm{BV}(\Omega)$ and the arguments for the case $a_0 = 0$ applied to $(\chi_n)_n$ instead of $(\tilde{\chi}_n)_n$ yield the result. □

Next, let us define the class of functions which should be recovered by minimizing ($\mathcal{P}$). We say that a $\chi \in \mathrm{BV}(\Omega)$ is *piecewise constant* if it essentially admits discrete values, i.e., there is a null-set E such that $\chi(\Omega \setminus E)$ is discrete. For piecewise constant χ, the jump set J_χ corresponds exactly to (confer [21, Section 5.9, Theorem 1])

$$J_\chi = \bigcup_{t\in\mathbb{R}} J_{\mathbf{1}_{\{\chi<t\}}},$$

i.e., the union of all interfaces between the constant regions of χ. The Coarea formula for BV-functions [20, Theorem 3.40] then yields that $|D\chi|(\Omega \setminus J_\chi) = 0$, i.e., $D\chi$ only admits a jump part.

The perfect reconstruction property will be obtained under a geometric assumption on the jump set J_χ. For the purpose of stating this assumption, consider, for $\xi \in \mathbb{R}^d$,

$$P(\xi) = \sum_{|\alpha|=k} a_\alpha \xi^\alpha \tag{3.3}$$

which defines a non-zero k-homogeneous polynomial in $\mathbb{R}[X_1, \ldots, X_d]$.

Definition 3.2. *We say that $\chi \in \mathrm{BV}(\Omega)$ is* regular *for P if we have*

$$\mathcal{H}^{d-1}(\{x \in J_\chi \colon P(\nu_\chi(x)) = 0\}) = 0.$$

Otherwise, χ is singular *for P. Moreover, a $u \in \mathrm{BV}(\Omega)$ is called* purely singular *for P, if*

$$P(\nu_u(x)) = 0 \qquad \text{for } \mathcal{H}^{d-1}\text{-almost every } x \in J_u.$$

For piecewise constant χ, one could alternatively define regularity for P via the characteristic functions of the sublevel sets $\mathbf{1}_{\{\chi<t\}}$. However, as the respective normals coincide $\mathcal{H}^{d-1}$-almost everywhere, χ is regular for P if and only if $\mathbf{1}_{\{\chi<t\}}$ is regular for P for each $t \in \mathbb{R}$.

With the notions of regularity and pure singularity for P, we can state the first auxiliary theorem which is proven in Section 4.

Theorem 3.3. *Let $u \in \mathrm{BV}(\Omega)$ be a weak solution of $\mathcal{A}u = 0$. Then, u is purely singular for P.*

In order to establish the perfect reconstruction property, we assume that the data $\psi = \mathcal{A}\chi$ is coming from a piecewise constant function χ that is regular for P. While this is not satisfied for general piecewise constant functions, it is still generic in the sense that fixing χ and varying $Q \in \mathrm{O}_d(\mathbb{R})$, the transformed function $\chi \circ Q$ will almost surely be regular for P. The latter statement will be made precise in the following.

Recall that one can define Lebesgue null sets on an n-dimensional smooth manifold M as follows: Choosing $(\phi_i, U_i)_i$ a C^∞-atlas for M, a $N \subset M$ is called a *Lebesgue null set* if for each i, the set $\phi_i(N \cap U_i) \subset \mathbb{R}^n$ is a Lebesgue null set. Hence, Lebesgue null sets on the $(d-1)d/2$-dimensional manifold $\mathrm{O}_d(\mathbb{R})$ have a meaning. As the latter is also a compact Lie group, it is moreover natural to consider the Haar probability measure, i.e., the regular Borel probability measure that is invariant to the group's (left) action. For such groups, the Haar probability measure is always a volume form [22] and hence, Haar null sets and Lebesgue null sets coincide. The Haar probability measure on $\mathrm{O}_d(\mathbb{R})$ can be interpreted as a random choice in $\mathrm{O}_d(\mathbb{R})$ without giving preference to any direction. Thus, if a property holds in $\mathrm{O}_d(\mathbb{R})$ up to a Lebesgue null set, then this property is true almost surely.

Now, for $\chi \in \mathrm{BV}(\Omega)$ and $Q \in \mathrm{O}_d(\mathbb{R})$, the transformed function $\chi \circ Q$ belongs to $\mathrm{BV}(Q^{-1}(\Omega))$. Regarding the regularity with respect to P, we have the following result which is the second auxiliary theorem.

Theorem 3.4. *For fixed $\chi \in \mathrm{BV}(\Omega)$, the function $\chi \circ Q$ is regular for P almost surely for $Q \in \mathrm{O}_d(\mathbb{R})$.*

The proof can be found in Section 5. Now, we use Theorem 3.3 to establish that one can not improve the TV semi-norm of a piecewise constant function under the assumption of regularity of the solution. This leads to a perfect reconstruction property as follows.

Theorem 3.5. *Let $\chi \in \mathrm{BV}(\Omega)$ be piecewise constant and regular for P and let $\psi = \mathcal{A}\chi$. Then, χ is a solution of ($\mathcal{P}$), i.e., we have*

$$\mathrm{TV}\,(\chi) \leq \mathrm{TV}\,(\chi + u)\,, \tag{3.4}$$

for any weak solution $u \in \mathrm{BV}(\Omega)$ of $\mathcal{A}u = 0$. Moreover, we have an equality in (3.4) if and only if u is a constant in the case $a_0 = 0$ and $u = 0$ in the case $a_0 \neq 0$.

Proof. First, we prove that $\mathcal{H}^{d-1}(J_u \cap J_\chi) = 0$ for any $u \in \mathrm{BV}(\Omega)$ with $\mathcal{A}u = 0$. According to [20, Theorem 3.78], J_u (resp. J_χ) is a countable $\mathcal{H}^{d-1}$-rectifiable set oriented by ν_u (resp. ν_χ). The existence result of traces for BV functions on interior rectifiable sets [20, Theorem 3.77] gives

$$D\chi \llcorner J_u = (\chi^+_{J_u} - \chi^-_{J_u})\nu_u \mathcal{H}^{d-1} \llcorner J_u,$$

where

$$\lim_{\varepsilon \to 0^+} \fint_{B^+_\varepsilon(x,\nu_u(x))} |\chi - \chi^+_{J_u}(x)|\,\mathrm{d}y = 0, \qquad \lim_{\varepsilon \to 0^+} \fint_{B^-_\varepsilon(x,\nu_u(x))} |\chi - \chi^-_{J_u}(x)|\,\mathrm{d}y = 0,$$

for $\mathcal{H}^{d-1}$-almost every $x \in J_u$. For such x, the property $\chi^+_{J_u}(x) \neq \chi^-_{J_u}(x)$ implies $x \in J_\chi$ by definition with either $\chi^+(x) = \chi^+_{J_u}(x)$, $\chi^-(x) = \chi^-_{J_u}(x)$ and $\nu_\chi(x) = \nu_u(x)$ or $\chi^+(x) = \chi^-_{J_u}(x)$, $\chi^-(x) = \chi^+_{J_u}(x)$ and $\nu_\chi(x) = -\nu_u(x)$. It follows in particular that $\nu_\chi(x) = \pm\nu_u(x)$ for $\mathcal{H}^{d-1}$-almost every $x \in J_u \cap J_\chi$. By hypothesis, $P(\nu_\chi(x)) \neq 0$ for $\mathcal{H}^{d-1}$-almost every $x \in J_\chi$. Further, according to Theorem 3.3 and the homogeneity of P, we have $P(\pm\nu_u(x)) = 0$ for $\mathcal{H}^{d-1}$-almost every $x \in J_u$. So, we deduce that $\mathcal{H}^{d-1}(J_u \cap J_\chi) = 0$.

In view of the decomposition (2.1) of Du, J_χ is a null set for both $\mathcal{L}^d$ and $|D^c u|$, thus $|Du|(J_\chi) = (u^+ - u^-)\mathcal{H}^{d-1}(J_u \cap J_\chi) = 0$. Also, since χ is piecewise constant, $|D\chi|(\Omega \setminus J_\chi) = 0$.

These two equalities imply that

$$\begin{aligned}\mathrm{TV}(\chi+u) &= |D\chi + Du|(\Omega)\\ &= |D\chi + Du|(J_\chi) + |D\chi + Du|(\Omega\setminus J_\chi)\\ &= |D\chi|(J_\chi) + |Du|(\Omega\setminus J_\chi)\\ &= |D\chi|(\Omega) + |Du|(\Omega)\\ &= \mathrm{TV}(\chi) + \mathrm{TV}(u).\end{aligned}$$

So, the inequality (3.4) follows immediately and equality holds if and only if $\mathrm{TV}(u) = 0$, that is, if and only if u is a constant. If $a_0 \neq 0$, then $\mathcal{A}u = 0$ moreover implies that this constant is zero, i.e., $u = 0$. □

We refer to Section 6 to the application of this theorem to perfect reconstruction of piecewise constant susceptibility distributions in Quantitative Susceptibility Mapping.

We conclude this section with a brief discussion of possible generalizations of Theorem 3.5 to total generalized variation (TGV) of order 2. The latter is a penalty functional that incorporates also the second-order information of a function while maintaining its ability to admit jump discontinuities. While the original definition is as a dual functional [5], it can also be written in a primal form as follows (see [23, 24]):

$$\mathrm{TGV}^2_\lambda(\chi) = \min_{v\in \mathrm{BD}(\Omega)} \lambda_1 |Du - v\mathcal{L}^d|(\Omega) + \lambda_0 |Ev|(\Omega) \tag{3.5}$$

where λ_0, λ_1 are positive constants and $\mathrm{BD}(\Omega)$ denotes the space of vector fields of *bounded deformation*, i.e., the collection of all $v \in L^1(\Omega, \mathbb{R}^d)$ such that $Ev \in \mathcal{M}(\Omega, \mathbb{R}^{d\times d})$ with $Ev = \frac{1}{2}(Dv + Dv^T)$ being the distributional symmetrized derivative. For the associated space

$$\mathrm{BGV}^2_\lambda(\Omega) = \{\chi \in L^1(\Omega)\colon\ \mathrm{TGV}^2_\lambda(\chi) < \infty\}, \qquad \|\chi\|_{\mathrm{BGV}^2_\lambda} = \|\chi\|_1 + \mathrm{TGV}^2_\lambda(\chi)$$

it holds that $\mathrm{BGV}^2_\lambda(\Omega) = \mathrm{BV}(\Omega)$ (see again [23, 24]). One can moreover show that TGV^2_λ is lower semi-continuous in $L^1(\Omega)$, $\mathrm{TGV}^2_\lambda(\chi) = 0$ if and only if χ is an affine-linear function in Ω and that a Poincaré-type estimate $\|\chi\|_1 \leq C\,\mathrm{TGV}^2_\lambda(\chi)$ holds for χ satisfying $\int_\Omega \chi\varphi \,\mathrm{d}x = 0$ for each φ linear-affine in Ω (see [24]). Inspecting the proof of Proposition 3.1, one sees that these properties are sufficient to prove existence of solutions of

$$\min_{\chi\in \mathrm{BV}(\Omega)} \mathrm{TGV}^2_\lambda(\chi) \qquad \text{subject to} \quad \mathcal{A}\chi = \psi \tag{3.6}$$

under the same assumptions as Proposition 3.1. However, it is not clear whether an analogous version of the perfect reconstruction property holds.

The following aims at giving a positive answer for a very special class of true data $\chi \in \mathrm{BV}(\Omega)$. Suppose that χ is piecewise constant and such that $\mathrm{TGV}^2_\lambda(\chi) = \lambda_1 \mathrm{TV}(\chi)$. Nontrivial χ with that property exist, such as, for instance, characteristic functions of $C^{1,1}$-sets in Ω (see [5, Proposition 3.11]). Now, if χ is moreover regular for P and $u \in \mathrm{BV}(\Omega)$ satisfies $\mathcal{A}u = 0$, then we still get that $|Du|(J_\chi) = 0$ and $|D\chi|(\Omega\setminus J_\chi) = 0$ as argued in the proof of Theorem 3.5. Therefore, for $v \in \mathrm{BD}(\Omega)$ arbitrary,

$$\begin{aligned}\lambda_1|D\chi + Du - v\mathcal{L}^d|(\Omega) + \lambda_0|Ev|(\Omega)\\ &= \lambda_1|D\chi + Du - v\mathcal{L}^d|(J_\chi) + \lambda_1|D\chi + Du - v\mathcal{L}^d|(\Omega\setminus J_\chi) + \lambda_0|Ev|(\Omega)\\ &= \lambda_1|D\chi|(J_\chi) + \lambda_1|Du - v\mathcal{L}^d|(\Omega\setminus J_\chi) + \lambda_0|Ev|(\Omega)\\ &= \lambda_1|D\chi|(\Omega) + \lambda_1|Du - v\mathcal{L}^d|(\Omega) + \lambda_0|Ev|.\end{aligned}$$

Minimizing over all $v \in \mathrm{BD}(\Omega)$ and using that $\mathrm{TGV}^2_\lambda(\chi) = \lambda_1 \mathrm{TV}(\chi)$ yields $\mathrm{TGV}^2_\lambda(\chi + u) = \mathrm{TGV}^2_\lambda(\chi) + \mathrm{TGV}^2_\lambda(u)$. Therefore, we have

$$\mathrm{TGV}^2_\lambda(\chi) \leq \mathrm{TGV}^2_\lambda(\chi + u)$$

for all $u \in \mathrm{BV}(\Omega)$ with $\mathcal{A}u = 0$ with equality if and only if $\mathrm{TGV}^2_\lambda(u) = 0$, i.e., u is linear-affine in Ω. This means that we are able to recover, via TGV^2_λ-minimization, the data χ up to linear-affine functions that satisfy the homogeneous PDE. Observe, however, that the assumptions on χ are quite strong.

4 The pure singularity of BV-solutions of the PDE

This section is devoted to the proof of Theorem 3.3, i.e., the statement that whenever $u \in \mathrm{BV}(\Omega)$ satisfies $\mathcal{A}u = 0$ in the weak sense for $\mathcal{A}$ according to (3.1), then u is purely singular for P or, in other words, for $\mathcal{H}^{d-1}$-almost every $x \in J_u$, it holds that $P(\nu_u(x)) = 0$.

Proof of Theorem 3.3. First, note that the case $k = 0$ is trivial: In this case, $\mathcal{A}u = 0$ implies $u = 0$, so there is nothing to show.

Therefore, assume $k \geq 1$. According to Lemma 2.1, for $\mathcal{H}^{d-1}$-almost every $x \in J_u$, the family $(u_{x,\varepsilon})_\varepsilon$, $\varepsilon > 0$, weakly-$\star$ converges to ω_x in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$. Choose such an $\bar{x} \in J_u$, denote by $\bar{\nu} = \nu_u(\bar{x})$ and let θ be a test function in $\mathcal{D}(\mathbb{R}^d)$ such that $\theta(0) = 1$, $\theta \geq 0$ and $\operatorname{supp}\theta \subset B_1(0)$. For $\varepsilon > 0$ we set

$$\varphi_\varepsilon(x) = \frac{\big(\bar{\nu} \cdot (x - \bar{x})\big)^{k-1}}{\varepsilon^{d-1}} \theta\left(\frac{x - \bar{x}}{\varepsilon}\right)$$

and aim at testing with φ_ε. As $u \in \mathrm{BV}(\Omega)$ is a solution of $\mathcal{A}u = 0$ in the distributional sense, we get that

$$\sum_{|\alpha| \leq k} (-1)^{|\alpha|} a_\alpha \int_\Omega u \partial^\alpha \varphi_\varepsilon \,\mathrm{d}x = 0. \tag{4.1}$$

Let us examine the limit of each integral as $\varepsilon \to 0^+$.

Claim 1. For $|\alpha| \leq k - 1$, we have $\int_\Omega u \partial^\alpha \varphi_\varepsilon \,\mathrm{d}x \to 0$ as $\varepsilon \to 0^+$.

To prove this, we apply Leibniz's derivation rule in order to get

$$\partial^\alpha \varphi_\varepsilon(x) = \varepsilon^{k-1-|\alpha|} \sum_{\beta \leq \alpha} \frac{1}{\varepsilon^{d-1}} c_\beta \left(\frac{\bar{\nu} \cdot (x - \bar{x})}{\varepsilon}\right)^{k-1-|\beta|} \partial^{\alpha-\beta}\theta\left(\frac{x - \bar{x}}{\varepsilon}\right)$$

where each coefficient c_β depends only on α and $\bar{\nu}$. For ε small enough, a change of variables yields

$$\int_\Omega u \partial^\alpha \varphi_\varepsilon \,\mathrm{d}x = \varepsilon^{k-|\alpha|} \sum_{\beta \leq \alpha} c_\beta \int_{\mathbb{R}^d} u_{\bar{x},\varepsilon}(y) (\bar{\nu} \cdot y)^{k-1-|\beta|} \partial^{\alpha-\beta}\theta(y) \,\mathrm{d}y,$$

so each integral stays bounded as $\varepsilon \to 0^+$ since in particular, $u_{\bar{x},\varepsilon} \to \omega_{\bar{x}}$ in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and each $y \mapsto (\bar{\nu} \cdot y)^{k-1-|\beta|} \partial^{\alpha-\beta}\theta(y)$ is bounded with compact support. Claim 1 then follows by letting $\varepsilon \to 0^+$ and observing that $\varepsilon^{k-|\alpha|} \to 0$.

Claim 2. If $|\alpha| = k$, then we have with $H_{\bar{\nu}} = \{y \in \mathbb{R}^d : y \cdot \bar{\nu} = 0\}$ that

$$\lim_{\varepsilon \to 0^+} \int_\Omega u \partial^\alpha \varphi_\varepsilon \,\mathrm{d}x = \left((k-1)!(u^-(\bar{x}) - u^+(\bar{x})) \int_{H_{\bar{\nu}}} \theta\,(y) \,\mathrm{d}\mathcal{H}^{d-1}(y)\right) \bar{\nu}^\alpha.$$

To prove this, choose i such that $\alpha_i \neq 0$, denote by $\alpha' = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_d)$ and apply again Leibniz's derivation rule in order to get

$$\partial^{\alpha'} \varphi_\varepsilon(x) = \frac{(k-1)!\bar{\nu}^{\alpha'}}{\varepsilon^{d-1}} \theta\left(\frac{x - \bar{x}}{\varepsilon}\right) + \sum_{\beta < \alpha'} c_\beta \frac{1}{\varepsilon^{d-1}} \left(\frac{\bar{\nu} \cdot (x - \bar{x})}{\varepsilon}\right)^{k-1-|\beta|} \partial^{\alpha'-\beta}\theta\left(\frac{x - \bar{x}}{\varepsilon}\right)$$

where each c_β now only depends on α' and $\bar\nu$. Denoting by $D_i u$ the i-th component of Du, we obtain, using weak differentiation,

$$\begin{aligned}\int_\Omega u\partial^\alpha\varphi_\varepsilon\,\mathrm{d}x &= -\int_\Omega \partial^{\alpha'}\varphi_\varepsilon\,\mathrm{d}D_i u\\ &= -\frac{(k-1)!\bar\nu^{\alpha'}}{\varepsilon^{d-1}}\int_\Omega \theta\left(\frac{x-\bar x}{\varepsilon}\right)\,\mathrm{d}D_i u\\ &\quad -\sum_{\beta<\alpha'}\frac{c_\beta}{\varepsilon^{d-1}}\int_\Omega\left(\frac{\bar\nu\cdot(x-\bar x)}{\varepsilon}\right)^{k-1-|\beta|}\partial^{\alpha'-\beta}\theta\left(\frac{x-\bar x}{\varepsilon}\right)\,\mathrm{d}D_i u. \end{aligned}\tag{4.2}$$

Moreover, as $\varepsilon\to 0^+$, the weak-$\star$ convergence of $u_{\bar x,\varepsilon}\to\omega_{\bar x}$ in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$ implies the weak-$\star$ convergence of $D_i u_{\bar x,\varepsilon}$ to $D_i\omega_{\bar x} = (u^+(\bar x)-u^-(\bar x))\bar\nu_i\mathcal{H}^{d-1}\llcorner H_{\bar\nu}$ in $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)$. Using this and the change-of-coordinates identity (2.2) for the first integral in (4.2) implies that

$$\lim_{\varepsilon\to0^+}\frac{1}{\varepsilon^{d-1}}\int_\Omega\theta\left(\frac{x-\bar x}{\varepsilon}\right)\,\mathrm{d}D_i u = (u^+(\bar x)-u^-(\bar x))\bar\nu_i\int_{H_{\bar\nu}}\theta(y)\,\mathrm{d}\mathcal{H}^{d-1}(y)$$

and that for each $\beta<\alpha'$ the corresponding integral obeys

$$\lim_{\varepsilon\to0^+}\frac{1}{\varepsilon^{d-1}}\int_\Omega\left(\frac{\bar\nu\cdot(x-\bar x)}{\varepsilon}\right)^{k-1-|\beta|}\partial^{\alpha'-\beta}\theta\left(\frac{x-\bar x}{\varepsilon}\right)\,\mathrm{d}D_i u = 0$$

since $(\bar\nu\cdot y)=0$ for $y\in H_{\bar\nu}$ and $k-1-|\beta|>0$. Hence, Claim 2 follows from sending (4.2) to the limit as $\varepsilon\to0^+$.

According to Claims 1 and 2, when $\varepsilon\to0^+$, the equality (4.1) gives

$$\left((-1)^{k+1}(k-1)!(u^+(\bar x)-u^-(\bar x))\int_{H_{\bar\nu}}\theta\,(y)\;\mathrm{d}\mathcal{H}^{d-1}(y)\right)\sum_{|\alpha|=k}a_\alpha\bar\nu^\alpha = 0.$$

As $(-1)^{k+1}(k-1)!(u^+(\bar x)-u^-(\bar x))\int_{H_{\bar\nu}}\theta\,(y)\;\mathrm{d}\mathcal{H}^{d-1}(y)\neq0$, we deduce that $P(\bar\nu)=P(\nu_u(\bar x))=0$. By the above choice of $\bar x$, this is true for $\mathcal{H}^{d-1}$-almost every $\bar x\in J_u$. □

5 Genericity of regularity with respect to orthogonal transforms

This section is devoted to the proof of Theorem 3.4 which will be done using real algebraic geometry. We will briefly introduce the relevant results from real algebraic geometry and refer to the literature for further details, for instance, [19].

First, for a $S\subset\mathbb{R}[X_1,\dots,X_d]$, we denote by $\mathcal{Z}(S)$ the following subset of $\mathbb{R}^d$:

$$\mathcal{Z}(S)=\{\xi\in\mathbb{R}^d\colon f(\xi)=0 \text{ for any } f\in S\}.$$

A set V is called *algebraic* if there exists $S\subset\mathbb{R}[X_1,\dots,X_d]$ such that $V=\mathcal{Z}(S)$. A set V is called *semi-algebraic* if it has a representation

$$V=\bigcup_{i=1}^{s}\bigcap_{j=1}^{r_i}\{\xi\in\mathbb{R}^d\colon f_{i,j}(\xi)\star_{i,j}0\},$$

where s and $r_1,\dots,r_s$ are integers, each $f_{i,j}\in\mathbb{R}[X_1,\dots,X_d]$ and $\star_{i,j}$ is either $<,>$ or $=$. An algebraic set is semi-algebraic and any semi-algebraic set can be decomposed as a finite union of smooth manifolds [19, Proposition 2.9.10]:

Proposition 5.1. *Let $V\subset\mathbb{R}^d$ be a semi-algebraic set. Then, V is the disjoint union of a finite number of semi-algebraic C^∞-submanifolds V_i, each diffeomorphic to an open hypercube $]0,1[^{\dim V_i}$ (with the special case $]0,1[^0=\{0\}$).*

In this context, $\dim V_i$ denotes the dimension of the respective submanifold. In contrast, recall that the *Hausdorff dimension* of a set $V \subset \mathbb{R}^d$, denoted by $\dim_{\mathcal{H}} V$, is defined by $\inf \{s \geq 0 \colon \mathcal{H}^s(V) = 0\}$. If V is semi-algebraic and the decomposition $V = \bigcup_i V_i$ is given by Theorem 5.1, then we have $\dim_{\mathcal{H}} V = \max_i \dim V_i$.

In order to prepare for the proof, let $\chi \in \mathrm{BV}(\Omega)$ and $Q \in \mathrm{O}_d(\mathbb{R})$ be given. Then, as straightforward computations show, $\chi \circ Q \in \mathrm{BV}(Q^{-1}(\Omega))$ with jump set $J_{\chi\circ Q} = Q^{-1}(J_\chi)$ and normal $\nu_{\chi\circ Q}(x) = Q^{-1}\nu_\chi(Qx)$ for $\mathcal{H}^{d-1}$-almost every $x \in J_{\chi\circ Q}$. Thus, $P(\nu_{\chi\circ Q}(x)) = 0$ for $\mathcal{H}^{d-1}$-almost every $x \in J_{\chi\circ Q}$ if and only if $P(Q^{-1}\nu_\chi(x)) = 0$ for $\mathcal{H}^{d-1}$-almost every $x \in J_\chi$. Consequently,

$$\chi \circ Q \text{ is regular for } P \quad \Leftrightarrow \quad \chi \text{ is regular for } P \circ Q^{-1}. \tag{5.1}$$

We are therefore interested in orthogonal transformations Q for which $P \circ Q^{-1}$ vanishes on certain sets. We first prove the following

Lemma 5.2. *Let P be a non-zero homogeneous polynomial in $\mathbb{R}^d$, $d \geq 2$ and let $\xi \in \mathbb{S}^{d-1}$ be fixed. Then, the set $\mathcal{P}_\xi = \{Q \in \mathrm{O}_d(\mathbb{R}) \colon P(Q^{-1}\xi) = 0\}$ is Lebesgue negligible in $\mathrm{O}_d(\mathbb{R})$.*

Proof. We first establish the following property: Let $0 < \varepsilon < 1$, $\xi \in \mathbb{S}^{d-1}$ and $Q \in \mathrm{O}_d(\mathbb{R})$ be fixed. Then, for any $\tilde{\xi} \in B_{\varepsilon/\sqrt{2}}(Q^{-1}\xi) \cap \mathbb{S}^{d-1}$, there exists a $\widetilde{Q} \in \mathrm{O}_d(\mathbb{R})$ such that $\widetilde{Q}^{-1}\xi = \tilde{\xi}$ and $\|Q - \widetilde{Q}\|_F < \varepsilon$ where $\|\cdot\|_F$ denotes the Frobenius matrix norm.

To see that this is true, let $\tilde{\xi} \in B_{\varepsilon/\sqrt{2}}(Q^{-1}\xi) \cap \mathbb{S}^{d-1}$ be fixed. If $Q^{-1}\xi = \tilde{\xi}$, the result is obvious, otherwise, as $\varepsilon < 1$, we have $Q^{-1}\xi \neq \pm\tilde{\xi}$. So, we can assume without loss of generality that $\{Q^{-1}\xi, \tilde{\xi}\}$ are linearly independent. Let Π be the plane in $\mathbb{R}^d$ spanned by $\{Q^{-1}\xi, \tilde{\xi}\}$ and $R \in \mathrm{O}_d(\mathbb{R})$ be the rotation that maps $RQ^{-1}\xi = \tilde{\xi}$ and is the identity on $\Pi^\perp$. We set $\widetilde{Q} = QR^{-1}$ and denote by $\theta \in \mathbb{R}$ be an angle of rotation associated with R, such that by construction we have

$$|Q^{-1}\xi - \tilde{\xi}| = \sqrt{|Q^{-1}\xi|^2 - 2\cos\theta|Q^{-1}\xi||\tilde{\xi}| + |\tilde{\xi}|^2} = \sqrt{2(1-\cos\theta)}.$$

By the choice of $\tilde{\xi}$, we get in particular that $\sqrt{2(1-\cos\theta)} < \varepsilon/\sqrt{2}$. Moreover, we have

$$\|Q - \widetilde{Q}\|_F = \|Q^{-1} - \widetilde{Q}^{-1}\|_F = \|\operatorname{id} - R\|_F = \left\| \begin{pmatrix} 1-\cos\theta & \sin\theta \\ -\sin\theta & 1-\cos\theta \end{pmatrix} \right\|_F = 2\sqrt{1-\cos\theta}.$$

So, we deduce that $\|Q - \widetilde{Q}\| < \varepsilon$.

Now, for $\xi \in \mathbb{S}^{d-1}$ the relation $P(Q^T\xi) = 0$ is algebraic with respect to $Q \in \mathbb{R}^{d\times d}$. Moreover, $Q \in \mathrm{O}_d(\mathbb{R})$ if and only if the algebraic relation $Q^TQ - \operatorname{id} = 0$ is satisfied and for $Q \in \mathrm{O}_d(\mathbb{R})$ we have $Q^{-1} = Q^T$, so the set $\mathcal{P}_\xi$ is algebraic and according to Proposition 5.1, it can be decomposed into a disjoint union of a finite number of C^∞-submanifolds V_i, each diffeomorphic to an open hypercube $]0,1[^{\dim V_i}$. We aim at leading the existence of an i such that $\dim V_i = \dim \mathrm{O}_d(\mathbb{R}) = d(d-1)/2$ to a contradiction. If this is the case, then V_i is non-empty and open in $\mathrm{O}_d(\mathbb{R})$, so there exists a $Q \in V_i$ and $0 < \varepsilon < 1$ such that $\|Q - \widetilde{Q}\|_F < \varepsilon$ implies $P(\widetilde{Q}^{-1}\xi) = 0$. According to the above, we deduce that for each $\tilde{\xi} \in B_{\varepsilon/\sqrt{2}}(Q\xi) \cap \mathbb{S}^{d-1}$, we can find a $\tilde{Q}$ with $\tilde{Q}^{-1}\xi = \tilde{\xi}$ and $P(\tilde{Q}^{-1}\xi) = 0$. Hence, P vanishes on $B_{\varepsilon/\sqrt{2}}(Q\xi) \cap \mathbb{S}^{d-1}$ and by homogeneity, on the whole cone generated by this set. As the latter is non-empty, open and since P is a polynomial, we deduce that $P = 0$, which is a contradiction. So, we have proved that for each i we have $\dim V_i < \dim \mathrm{O}_d(\mathbb{R})$ and, consequently, each V_i is Lebesgue negligible in $\mathrm{O}_d(\mathbb{R})$. Thus, $\mathcal{P}_\xi$ is a Lebesgue null set in $\mathrm{O}_d(\mathbb{R})$. □

Now, we are ready to establish the proof.

Proof of Theorem 3.4. Let $\chi \in \mathrm{BV}(\Omega)$ be fixed. If $d = 1$, the statement is trivial as P is homogeneous in $\mathbb{R}$ and then it vanishes in $\mathbb{S}^0$ if and only if $P = 0$.

In the case $d \geq 2$, we establish the following property: For each compact $K \subset \mathbb{R}^d$ with $\mathcal{H}^{d-1}(K) < \infty$ and $\nu : K \to \mathbb{S}^{d-1}$ measurable with respect to $\mathcal{H}^{d-1}$, there exists an at most countable set $\mathcal{S}_{K,\nu}$ such that

$$\Big\{Q \in \mathrm{O}_d(\mathbb{R}) : \mathcal{H}^{d-1}(\{x \in K\colon P(Q^{-1}\nu(x)) = 0\}) > 0\Big\} \subset \bigcup_{\xi \in \mathcal{S}_{K,\nu}} \mathcal{P}_\xi. \tag{5.2}$$

Then, the statement of the theorem can be deduced as follows: According to [20, Theorem 3.78], the jump set J_χ is countably $\mathcal{H}^{d-1}$-rectifiable, i.e., up to a $\mathcal{H}^{d-1}$-null set, the countable union of compact sets K with finite $\mathcal{H}^{d-1}$-measure and associated normal $\nu = \nu_\chi|_K$. Setting $\mathcal{S}$ as the union of all the corresponding $\mathcal{S}_{K,\nu}$ still yields an at most countable set. Now, if $\mathcal{H}^{d-1}(\{x \in J_\chi\colon P(Q^{-1}\nu_\chi(x)) = 0\}) > 0$, then $\mathcal{H}^{d-1}(\{x \in K\colon P(Q^{-1}\nu(x)) = 0\}) > 0$ for at least one of the (K, ν) in the above decomposition of (J_χ, ν_χ). Hence,

$$\Big\{Q \in \mathrm{O}_d(\mathbb{R}) : \mathcal{H}^{d-1}(\{x \in J_\chi\colon P(Q^{-1}\nu_\chi(x)) = 0\}) > 0\Big\} \subset \bigcup_{\xi \in \mathcal{S}} \mathcal{P}_\xi.$$

The set on the left-hand side are exactly the $Q \in \mathrm{O}_d(\mathbb{R}^d)$ for which χ is singular for $P \circ Q^{-1}$. The latter is, however, equivalent to $\chi \circ Q$ being singular for P, see (5.1). By Lemma 5.2, each $\mathcal{P}_\xi$ is a Lebesgue null set and so is the the countable union over all $\xi \in \mathcal{S}$. Thus, the $Q \in \mathrm{O}_d(\mathbb{R}^d)$ for which $\chi \circ Q$ is singular for P are Lebesgue-negligible and the statement of the theorem follows.

Now, in order to establish (5.2), we construct the set $\mathcal{S}_{K,\nu}$ according to the following algorithm:

Data: $K \subset \mathbb{R}^d$ compact, $\mathcal{H}^{d-1}(K) < \infty$, $\nu : K \to \mathbb{S}^{d-1}$ measurable w.r.t. $\mathcal{H}^{d-1}$;
Result: $\mathcal{S}_{K,\nu} \subset \mathbb{S}^{d-1}$ at most countable;
Initialization: $\mathcal{S} \leftarrow \emptyset$;
for $j = 0, \dots, d-2$ **do**
  for $n = 1, 2, 3, \dots$ **do**
    $\mathcal{W} \leftarrow \{V \subset \mathbb{S}^{d-1}\colon V$ is semi-algebraic, $\dim_{\mathcal{H}} V = j$,
      $\mathcal{H}^{d-1}(\nu^{-1}(V)) \geq 1/n,\ V \cap \mathcal{S}_{K,\nu} = \emptyset\}$;
    while $\mathcal{W} \neq \emptyset$ **do**
      Choose $V \in \mathcal{W}$;
      Decompose $V = \bigcup_i V_i$ as in Theorem 5.1;
      $\mathcal{I} \leftarrow \{i\colon \dim V_i = j\}$;
      For each $i \in \mathcal{I}$, choose $\mathcal{S}_i$ at most countable and dense in V_i;
      $\mathcal{S}_{K,\nu} \leftarrow \mathcal{S}_{K,\nu} \cup \bigcup_{i \in \mathcal{I}} \mathcal{S}_i$;
      $\mathcal{W} \leftarrow \{V \subset \mathbb{S}^{d-1}\colon V$ is semi-algebraic, $\dim_{\mathcal{H}} V = j$,
        $\mathcal{H}^{d-1}(\nu^{-1}(V)) \geq 1/n,\ V \cap \mathcal{S}_{K,\nu} = \emptyset\}$;
    end
  end
end

We will see that the **while**-loop always terminates after finitely many steps. Moreover, the **for**-loop for n has to be interpreted as a recursion updating $\mathcal{S}_{K,\nu}$ for each n and taking the countable union afterwards. In particular, the resulting $S_{K,\nu}$ will be at most countable.

Considering $\mathcal{S}_{K,\nu}$ at the end of each step in the outermost loop for $j = 0, \dots, d-2$, we claim that

$$\Big(V \subset \mathbb{S}^{d-1},\ V \text{ is semi-algebraic},\ \dim_{\mathcal{H}} V \leq j,\ \mathcal{H}^{d-1}(\nu^{-1}(V)) > 0\Big) \quad \Rightarrow \quad V \cap \mathcal{S}_{K,\nu} \neq \emptyset. \tag{$\mathcal{C}_j$}$$

This claim will be verified via finite induction. Suppose that ($\mathcal{C}_0$), ..., ($\mathcal{C}_{j-1}$) are true. Fix $n \geq 1$ and denote by V^r the set V in the r-th step of the **while**-loop. We aim at showing that for $1 \leq r < s$, it holds that $\mathcal{H}^{d-1}(\nu^{-1}(V^r \cap V^s)) = 0$. For that purpose, observe that for $j = 0$, the 0-dimensional semi-algebraic set V^r is a finite union of points. Thus, at the s-th step of the

while-loop, $V^r \subset \mathcal{S}_{K,\nu}$ and hence, $V^r \cap V^s = \emptyset$ since $V^s \cap S_{K,\nu} = \emptyset$ by construction, yielding the result.

For $j \geq 1$, we state that $\dim_{\mathcal{H}} V^r \cap V^s < j$. Clearly, $\dim_{\mathcal{H}} V^r \cap V^s \leq j$, so assume that $\dim_{\mathcal{H}} V^r \cap V^s = j$. Denoting by $(V_i^r)_i$ the decomposition of V^r used in the algorithm and by $(W_{i'})_{i'}$ a decomposition of $V^r \cap V^s$ according to Theorem 5.1 yields the existence of i, i' such that $\dim_{\mathcal{H}} V_i^r \cap W_{i'} = j$ (otherwise, $V^r \cap V^s$ would be the finite union of sets of Hausdorff-dimension less than j). Then, necessarily, $\dim V_i^r = \dim W_{i'} = j$ and $V_i^r \cap W_{i'}$ is non-empty, open in V_i^r. By construction, $\mathcal{S}_{K,\nu}$ contains, in the s-th step of the **while**-loop, a dense subset of V_i^r and thus, also a point of $W_{i'} \subset V^s$. However, this contradicts $V^s \cap \mathcal{S}_{K,\nu} = \emptyset$, so it must hold that $\dim_{\mathcal{H}} V^r \cap V^s < j$. Invoking the induction hypothesis ($\mathcal{C}_{j-1}$) then yields $\mathcal{H}^{d-1}(\nu^{-1}(V^r \cap V^s)) = 0$: Otherwise, $V^r \cap V^s \cap S_{K,\nu} \neq \emptyset$ for the $S_{K,\nu}$ at the end of the previous step of the outermost loop. This contradicts both the choice of V^r and V^s of the algorithm.

In any case, $\mathcal{H}^{d-1}(\nu^{-1}(V^r \cap V^s)) = 0$ holds for any distinct r and s such that we deduce that the following additivity property holds true:

$$\mathcal{H}^{d-1}(\nu^{-1}(\textstyle\bigcup_r V^r)) = \displaystyle\sum_r \mathcal{H}^{d-1}(\nu^{-1}(V^r)).$$

As by construction, we have for each r that $\mathcal{H}^{d-1}(\nu^{-1}(V^r)) \geq 1/n$ and since $\mathcal{H}^{d-1}(\nu^{-1}(\bigcup_r V^r)) < \infty$, the sum must be finite and hence, the **while**-loop must terminate.

Finally, regarding $\mathcal{S}_{K,\nu}$ at the end of step j of the outermost loop, the statement ($\mathcal{C}_j$) is true: Choosing a semi-algebraic $V \subset \mathbb{S}^{d-1}$ with $\dim_{\mathcal{H}} V = j$ and such that $\mathcal{H}^{d-1}(\nu^{-1}(V)) = c > 0$, we fix n such that $c \geq 1/n$. After the **while**-loop terminated for this n, we have $\mathcal{W} = \emptyset$, so it follows that $V \cap \mathcal{S}_{K,\nu} \neq \emptyset$, which establishes the induction step.

With ($\mathcal{C}_{d-2}$) being true, we may apply it to the algebraic set $V = \{\xi \in \mathbb{S}^{d-1} \colon (P \circ Q^{-1})(\xi) = 0\}$ for $Q \in \mathrm{O}_d(\mathbb{R})$. As P is homogeneous and non-zero, $\dim_{\mathcal{H}} V \leq d-2$, so whenever $\mathcal{H}^{d-1}(\nu^{-1}(V)) > 0$, we get $V \cap S_{K,\nu} \neq \emptyset$. Consequently, there is a $\xi \in S_{K,\nu}$ with $P(Q^{-1}\xi) = 0$, so $Q \in \mathcal{P}_\xi$. This establishes (5.2), so the proof is complete. □

6 Application to Quantitative Susceptibility Mapping

As already mentioned in the introduction, Quantitative Susceptibility Mapping (QSM) is a recent Magnetic Resonance Imaging (MRI) technique which aims at providing the spatial susceptibility distribution χ of tissues inside the human body. With MRI, one is able to measure the magnetic field inhomogeneity b_0 along the direction of the main magnetic field. This inhomogeneity is caused by the susceptibility distribution χ. Maxwell's equation and MR Physics [25] lead to the following relation for χ and b_0:

$$\frac{k_x^2 + k_y^2 - 2k_z^2}{3} \mathcal{F}(\chi)(\mathbf{k}) = |\mathbf{k}|^2 \mathcal{F}(b_0)(\mathbf{k}), \qquad \mathbf{k} = (k_x, k_y, k_z) \in \mathbb{R}^3, \tag{6.1}$$

where $\mathcal{F}$ is the Fourier transform in $\mathbb{R}^3$. This equation can be viewed as the Fourier transform of the following partial differential equation

$$\Box \chi = -\Delta b_0, \tag{6.2}$$

where Δ is the Laplacian in $\mathbb{R}^3$ and $\Box$ is a certain wave operator, where the coordinates x, y (associated with axes transversal to the main magnetic field) correspond to the space and the z-coordinate (whose axis is aligned with the main magnetic field) corresponds to the time:

$$\Box = \frac{2}{3}\frac{\partial^2}{\partial z^2} - \frac{1}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2}.$$

Usually, QSM does not base on solving (6.2). Instead, the problem is reformulated as a deconvolution problem as a result of the following formal derivation. Starting from (6.1), we divide

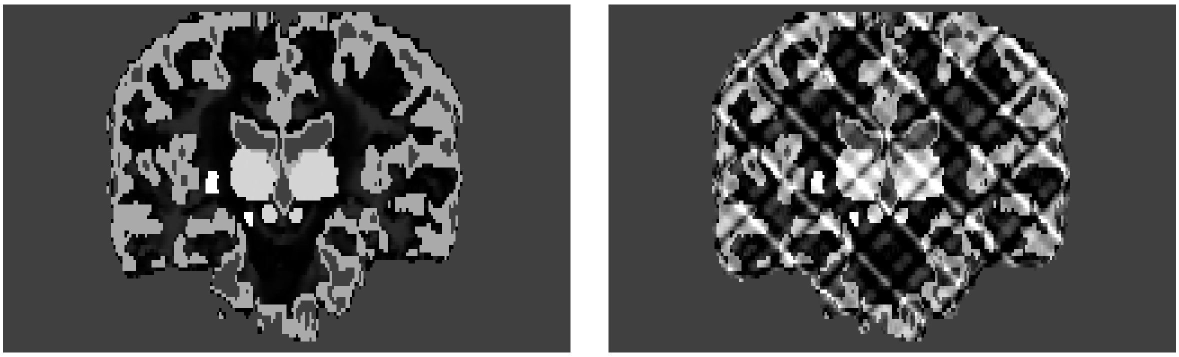

Figure 6.1: A coronal slice of a piecewise constant susceptibility phantom χ on a region of interest Ω (left) and $\chi + u$ (right) with $\Box u = 0$. The solution u causes structured artifacts that are typical for QSM.

$|\mathbf{k}|^2\mathcal{F}(b_0)(\mathbf{k})$ by $(k_x^2 + k_y^2 - 2k_z^2)/3$ and set

$$\delta = \mathcal{F}^{-1}(\hat{\delta}) \qquad \text{where} \qquad \hat{\delta}(\mathbf{k}) = \frac{k_x^2 + k_y^2 - 2k_z^2}{3|\mathbf{k}|^2}.$$

Using the relation $\mathcal{F}(\delta)\mathcal{F}(\chi) = \mathcal{F}(\delta * \chi)$, the equation (6.1) becomes

$$\chi * \delta = b_0. \tag{6.3}$$

The function δ is called the *magnetic dipole kernel* and the equation (6.3) explains why QSM reconstruction is usually regarded as a deconvolution problem [26, 1]. However, despite of its theoretical simplicity, this approach suffers from several drawbacks. First, the equation $\chi * \delta = b_0$ makes sense only for data which are given in the whole space $\mathbb{R}^3$, while in practice, b_0 is only known on a bounded domain $\Omega \subset \mathbb{R}^3$. This is due to the fact that MRI measurements of the field inhomogeneity b_0 can only be performed in a region of interest where proton spins may be excited by radio-frequency pulses, which is basically only inside parts of the body (usually the brain in case of QSM).

A second default of this approach is due to the fact that the Fourier-transformed kernel $\hat{\delta}$ vanishes on the critical cone $\mathcal{C} = \{(k_x, k_y, k_z) \colon k_x^2 + k_y^2 - 2k_z^2 = 0\}$. As a consequence, deconvolution via division by $\hat{\delta}$ in Fourier space is an operation which is well-defined only for functions that are factorizable by $\hat{\delta}$, *i.e.*, for functions whose Fourier transforms vanish in a accurate way on the cone $\mathcal{C}$. Thus, if the measurement b_0 is perturbed by some additional noise, this leads to an amplification of noise phenomenon, an ill-posedness effect which is typical for deconvolution problems. In this case, the dominant part of the amplified noise's Fourier transform is supported on the critical cone $\mathcal{C}$, which means that this part satisfies the homogeneous wave equation associated with $\Box$ in $\mathbb{R}^3$, resulting in characteristic *streaking artifacts*, see Figure 6.1 for a numerical example.

These drawbacks are well-known in the QSM literature and there is a plethora of methods dealing with each of the above difficulties. First, the fact that b_0 is only known on the bounded region of interest Ω is accounted for by performing *background field correction*, i.e., estimating and eliminating all contributions of χ outside of Ω from b_0. There are several ways of doing this, we refer to the review article [27] for an overview. Second, the ill-posedness of the deconvolution is dealt with by employing regularization. Along classical linear approaches, Tikhonov regularization with TV penalties can be found in the literature [28]. Recently, approaches for QSM were proposed that perform these two tasks simultaneously by solving a single minimization problem [3, 29]. These approaches also employ minimization involving TV or TGV_λ^2 and turn out to be very robust as well as computationally efficient, allowing in particular for fast MR acquisition schemes with typically low signal-to-noise ratio.

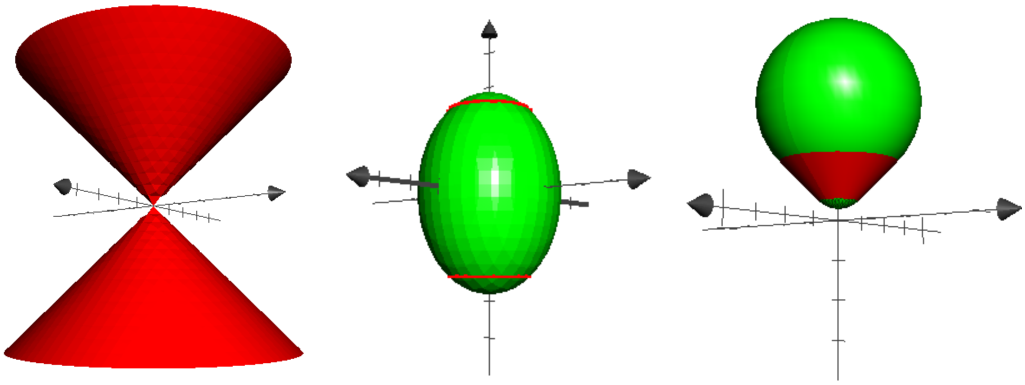

Figure 6.2: The critical cone $\{\xi \colon P(\xi) = 0\}$ associated with QSM (left), a level-set of a function regular (center) and singular (right) for P, respectively (also confer Definition 3.2).

The underlying idea of the latter single-step techniques for QSM is not to solve the deconvolution problem (6.3) in $\mathbb{R}^3$ but rather to employ (6.2) which can naturally be restricted to a bounded Ω where the data b_0 is given. However, at this point, the underlying model for QSM does not provide boundary conditions, so a susceptibility distribution χ can only be recovered up to solutions u of the wave equation $\Box u = 0$. In order to single out a meaningful solution, we employ a-priori assumptions on the true data χ. Within image processing, it is nowadays well-accepted that $\chi \in \mathrm{BV}(\Omega)$ is a suitable image prior [30, 31]. In this context, one often assumes that χ is piecewise constant and in consequence possessing a low total variation. This naturally leads to the minimization problem

$$\min_{\chi \in \mathrm{BV}(\Omega)} \ \mathrm{TV}(\chi) \qquad \text{subject to} \qquad \Box \chi = -\Delta b_0. \tag{6.4}$$

This is exactly a problem of structure ($\mathcal{P}$) with $\mathcal{A} = \Box$ and $\psi = -\Delta b_0$. So we may apply Theorem 3.5 and get that χ can exactly be recovered up to a constant almost surely with respect to rotations (and mirrored rotations) in $\mathbb{R}^3$. The latter can be interpreted as follows: For a fixed piecewise constant χ, the regularity for P with $P(\xi) = 2\xi_z^2 - \xi_x^2 - \xi_y^2$ yields a condition on the structure of the interfaces of the susceptibility distribution that may or may not be satisfied (see Figure 6.2). However, performing a QSM measurement in the MR scanner, one will actually acquire data from randomly distributed slightly rotated versions of χ when performing repeated measurements, giving almost surely regular true data and, consequently, perfect reconstruction. We can thus summarize:

Theorem 6.1. *For b_0 generated by a piecewise constant susceptibility distribution $\chi \in \mathrm{BV}(\Omega)$, the QSM method (6.4) recovers χ up to a constant almost surely with respect to rotations.*

As discussed at the end of Section 3, this is still true if TV is replaced with TGV_λ^2 under more restrictive assumptions on χ, a fact that justifies, at least in a specific situation, also the use of total generalized variation for QSM as it is done in [3, 4].

As the above model assumes noise-free measurements, let us finally comment on how noise can be incorporated into the model. Denoting by b_0 the true data and by $\eta \in L^2(\Omega)$ the noise in the measurement, plugging in the noisy data $\tilde{b}_0 = b_0 + \eta$ into the QSM equation (6.2) yields

$$\Box \chi + \Delta \eta = -\Delta \tilde{b}_0.$$

We now have to estimate, in the variational approach, η in addition to χ. Usually, the noise (which can be seen to be the discrepancy of the above equation) is measured in the squared L^2-norm which has to be balanced against TV in terms of a regularization parameter $\lambda > 0$, leading to the following Tikhonov regularization approach for QSM:

$$\min_{\substack{\chi \in \mathrm{BV}(\Omega),\\ \eta \in L^2(\Omega)}} \ \tfrac{1}{2}\|\eta\|_2^2 + \lambda\,\mathrm{TV}(\chi) \qquad \text{subject to} \qquad \Box \chi + \Delta \eta = -\Delta \tilde{b}_0. \tag{6.5}$$

The numerical performance of this and the analogous method with $\lambda\,\mathrm{TV}$ replaced by TGV^2_λ can be found in [3]. Note that the analogy of (6.4) and compressed sensing discussed in the introduction can be extended to (6.5): In compressed sensing, minimizing a Tikhonov functional consisting of the 2-norm of the discrepancy and the sparsifying norm is also the most common way for sparse recovery from noisy data [10]. In this situation, the perfect reconstruction property does no longer hold, however, stability of the reconstruction with precise error bounds can still be established. Whether an analogous result holds for (6.5) or, more generally, for the corresponding modification of ($\mathcal{P}$), is currently an open question and direction of future research.

7 Conclusions and outlook

As we have seen, TV-minimization is indeed able to recover piecewise constant functions, at least when subjected to linear PDE constraints. Moreover, we were discussing the application to QSM, justifying the use of TV-minimization methods for the reconstruction of susceptibility maps in MRI.

There are several possible directions of further research and some open problems. First of all, throughout the paper, we were assuming exact measurements which is reflected in the constraint $\mathcal{A}\chi = \psi$. However, in practice, one has to assume that ψ contains noise. In this context, a model for the regularity of the noise is important. For instance, as we were discussing for the QSM application at the end of Section 6, the noise on ψ could be the Laplacian of an L^2-function, implying some negative Sobolev regularity. This would lead to variational problems of the type

$$\min_{\substack{\chi\in \mathrm{BV}(\Omega),\\ \eta\in L^2(\Omega)}} \ \tfrac{1}{2}\|\eta\|_2^2 + \lambda\,\mathrm{TV}(\chi) \qquad \text{subject to} \qquad \mathcal{A}\chi + \mathcal{B}\eta = \psi,$$

where $\mathcal{B}$ is another linear differential operator with constant coefficients, for instance of elliptic type. It is currently not known what can be said about solutions of this problem, we expect, however, that the present perfect reconstruction results can be extended to yield some stable reconstruction property and desirable bounds for the reconstruction error. Such statements would be in analogy to respective results in compressed sensing.

Another open question is related to minimizing with respect to TGV^2_λ instead of TV in order to recover χ, an approach that was discussed at the end of Section 3, see (3.6). Indeed, minimizers still exist and for particular piecewise constants functions, the perfect reconstruction property is valid. However, TGV^2_λ is in particular a suitable model for piecewise affine functions, so it would be of interest to investigate whether perfect reconstruction analogous to Theorem 3.5 may be established for this class of functions. We expect that the proof of Theorem 3.5 has to be modified accordingly, which involves the study of higher-order jump singularities of solutions of $\mathcal{A}u = 0$ as well as the structure of optimal v in the minimum representation (3.5) for TGV^2_λ. An alternative approach in this direction would be to examine perfect reconstruction via TV^2-minimization for piecewise affine functions $\chi \in \mathrm{BV}^2(\Omega)$, i.e., without jump singularities. Results in that direction can be expected to be beneficial for the TGV^2_λ-case.

Finally, it might be possible to extend the main result of this paper to stochastic coefficients for the operator $\mathcal{A}$, i.e., considering $(a_\alpha)_\alpha$ as random variables with a certain distribution. This would lead to a generalization of considering $\chi \circ Q$ with the orthogonal transform Q being random and asking whether perfect reconstruction can be achieved. One would then expect an analogous statement to Theorem 3.4, for instance, that the regularity assumption of a fixed piecewise constant χ with respect to P (which is now random) is satisfied almost surely or with high probability.

Acknowledgments

This work was supported in part by the Austrian Science Fund (FWF) under grant SFB-F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences").

References

[1] Y. Wang and T. Liu. Quantitative Susceptibility Mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. *Magn. Res. Med.*, 73(1):82–101, 2015.

[2] E. M. Haacke, N. Y. Cheng, M. J. House, Q. Liu, J. Neelavalli, R. J. Ogg, A. Khan, M. Ayaz, W. Kirsch, and A. Obenaus. Imaging iron stores in the brain using magnetic resonance imaging. *Magn. Res. Imaging*, 23(1):1–25, 2005.

[3] C. Langkammer, K. Bredies, B.A. Poser, M. Barth, G. Reishofer, A.P. Fan, B. Bilgic, F. Fazekas, C. Mainero, and S. Ropele. Fast Quantitative Susceptibility Mapping using 3D EPI and Total Generalized Variation. *NeuroImage*, 111:622–630, 2015.

[4] I. Chatnuntawech, P. McDaniel, S.F. Cauley, B.A. Gagoski, C. Langkammer, A. Martin, P.E. Grant, L.L. Wald, K. Setsompop, E. Adalsteinsson, and B. Bilgic. Single-step quantitative susceptibility mapping with variational penalties. *NMR in Biomedicine*, 2016.

[5] K. Bredies, K. Kunisch, and T. Pock. Total Generalized Variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[6] F. Tröltzsch. *Optimal control of partial differential equations.* Americal Mathematical Society, 2010.

[7] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for optimal control problems. *SIAM Journal on Control and Optimization*, 37(4):1176–1194, 1999.

[8] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2006.

[9] C. Meyer, A. Rösch, and F. Tröltzsch. Optimal control problems of PDEs with regularized pointwise state constraints. *Computational Optimization and Applications*, 33:209–228, 2006.

[10] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[11] A. Chambolle, V. Duval, G. Peyré, and C. Poon. Geometric properties of solutions to the total variation denoising problem. *Inverse Problems*, 33(1):015002, 2016.

[12] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.

[13] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.

[14] B. Adcock, A. C. Hansen, and C. Poon. On optimal wavelet reconstructions from Fourier samples: linearity and universality of the stable sampling rate. *Appl. Comput. Harmon. Anal.*, 36(3):387–415, 2014.

[15] E. J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[16] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015.

[17] L. Hörmander. *The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis.* Berlin: Springer Verlag, 2nd edition, 1990.

[18] G. De Philippis and F. Rindler. On the structure of $\mathcal{A}$-free measures and applications. *Annals of Mathematics*, 184:1017–1039, 2016.

[19] J. Bochnak, M. Coste, and M.F. Roy. *Real Algebraic geometry.* Springer, 1998.

[20] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford Science Publications, 2000.

[21] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, 1992.

[22] W. Rossmann. *Lie groups: an introduction through linear groups.* Oxford University Press, 2002.

[23] K. Bredies and T. Valkonen. Inverse problems with second-order total generalized variation constraints. In *Proceedings of SampTA 2011 — 9th International Conference on Sampling Theory and Applications*, Singapore, 2011.

[24] K. Bredies and M. Holler. Regularization of linear inverse problems with Total Generalized Variation. *Journal of Inverse and Ill-posed Problems*, 22(6):871–913, 2014.

[25] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design.* John Wiley and Sons, New York, 1999.

[26] J. K. Choi, H. S. Park, S. Wang, and J. K. Seo. Inverse problem in quantitative susceptibility mapping. *SIAM Journal on Imaging Sciences*, 7(3):1669–1689, 2014.

[27] F. Schweser, S. D. Robinson, L. de Rochefort, W. Li, and K. Bredies. An illustrated comparison of processing methods for phase MRI and QSM: removal of background field contributions from sources outside the region of interest. *NMR in Biomedicine*, doi:10.1002/nbm.3604, 2016. To appear.

[28] L. de Rochefort, R. Brown, M. Prince, and Y. Wang. Quantitative MR susceptibility mapping using piecewise constant regularized inversion of the magnetic field. *Magn. Res. Med.*, 60:1003–1009, 2008.

[29] K. Bredies, C. Langkammer, and D. Vicente. TGV-based single-step QSM reconstruction with coil combination. In *4th International Workshop on MRI Phase Constrast and Quantitative Susceptibility Mapping.* Medical University of Graz, 2016.

[30] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[31] V. Caselles, A. Chambolle, and M. Novaga. The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Modeling and Simulation*, 6(3):879–894, 2007.