Over-parameterized Models for Vector Fields

Keren Rotker*, Dafna Ben Bashat†, and Alex M. Bronstein‡

Abstract. Vector fields arise in a variety of quantity measure and visualization techniques such as fluid flow imaging, motion estimation, deformation measures and color imaging, leading to a better understanding of physical phenomena. Recent progress in vector field imaging technologies has emphasized the need for efficient noise removal and reconstruction algorithms. A key ingredient in the success of extracting signals from noisy measurements is prior information, which can often be represented as a parameterized model. In this work, we extend the over-parameterization variational framework in order to perform model-based reconstruction of vector fields. The over-parameterization methodology combines local modeling of the data with global model parameter regularization. By considering the vector field as a linear combination of basis vector fields and appropriate scale and rotation coefficients, the denoising problem reduces to a simpler form of coefficient recovery. We introduce two versions of the over-parameterization framework: total variation-based method and sparsity-based method, relying on the cosparse analysis model. We demonstrate the efficiency of the proposed frameworks for two- and three-dimensional vector fields with linear and quadratic over-parameterization models.

Key words. vector fields, denoising, over-parameterization, variational methods, regularization, total-variation, sparsity, cosparsity, inverse problems

AMS subject classifications. 47N10, 35A15, 49N45, 68U10, 17B66, 37C10, 46N10

1. Introduction.

Reconstruction and denoising of vector fields have become a main subject of research in image and signal processing. This is partly due to their being the appropriate mathematical representation of objects such as displacement or deformation fields. Moreover, modern imaging technologies enable direct measurements of flows as vector quantities. Such imaging modalities include the particle image velocimetry (PIV), an optical method which provides velocity measurements in fluids, and the phase-contrast magnetic resonance imaging (PC-MRI) which produces an in-vivo time-resolved velocity field of the blood flow. Recent growth in computational power and capacity, enable to process large volumes of multidimensional data and design algorithms for analyzing the vector fields data.

Visualization and quantitative analysis of the flow and its pattern have a great significance in many disciplines. In medical imaging, for instance, blood flow patterns within the vessels are believed to be associated with the formation of several pathologies and their evolution [31, 11]. However, in many cases, the imaging techniques produce low signal to noise ratio measurements, motivating the need for flow-field denoising and analysis algorithms, which consider the flow pattern and its physical properties.

Several variational techniques have been considered for denoising and reconstruction of flow-fields. An n-dimensional flow-field with n components can be represented by the vector function $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ over $\mathbb{R}^n$, where $\mathbf{x} \in \mathbb{R}^d$. The most popular functional for noise removal includes a least-squares fitting data term and a regularization term which

*School of Electrical Engineering, Tel-Aviv University, Israel.
†Functional Brain Center, Tel Aviv Sourasky Medical Center, Israel.
‡Department of Computer Science, Technion - Israel Institute of Technology, Israel.

imposes certain characteristics on the signal, leading to the minimization problem of the form

$$\mathbf{f}^* = \underset{\mathbf{f}:\mathbb{R}^d \to \mathbb{R}^n}{\arg\min} \int_\Omega \|\mathbf{y}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\|^2 \, d\mathbf{x} + \lambda \Re(\mathbf{f}), \tag{1.1}$$

where Ω is the vector field domain, $\mathbf{y}$ is a noisy flow-field, $\mathbf{f}^*$ is the reconstructed field, $\|\cdot\|$ is the L_2-norm, $\Re$ is the regularization term and $\lambda > 0$ is a parameter weighing the regularization term in relation to the data term.

Variational reconstruction methods for vector fields can roughly be divided into two classes. The first class of methods extends known regularization techniques to multi-channel data. Relying on scalar setting observations that L_1 regularization and specifically total variation (TV) regularization preserves edges and discontinuities better than its L_2 counterparts [41], some efforts were made in order to extend the TV scalar regularization to vector fields. A straightforward extension, comprised of penalizing the variations of each channel separately, was introduced in [4]. While forming a simple optimization problem, this method ignores any dependencies that might exist among the vector's components. In order to provide coupling between different channels while preserving the discontinuities, a vectorial TV norm was defined as an L_2-norm of the channel-by-channel TV norm [6, 42]. Other regularization techniques including the nuclear TV norm, the beltrami flow or the Mumford-Shah regularizer were also used to couple the multi-channel data [7, 46, 45, 5].

The second class of variational techniques for vector field denoising comprises of methods that impose particular physical properties of the model on the given measurements. The irrotational and incompressible characteristics of fluid flow are governed by the curl and divergence operators, which makes the curl-divergence regularization very common in fluid flow-field reconstructions. Combined with the L_1 norm, this regularizer has been efficiently used for denoising vector-fields with discontinuities [48, 47, 9]. A different approach was taken in [35, 39] for optical flow estimation. The authors proposed to represent the optical flow vector at each pixel by the coefficients of a specific motion model, using an over-parameterized representation.

In this work, we propose a novel over-parameterized variational framework for vector field denoising, which relies on previous knowledge of the vector field pattern in order to perform model-based reconstruction of the signal from its noisy measurements. While most vector field recovery variational algorithms directly penalize the change of the flow, the over-parameterization representation has the advantage that the smoothness term penalizes deviations from the flow model. In the proposed framework, each vector field is represented as a linear combination of basis vector fields and appropriate scale and rotation coefficients, thus the denoising problem reduces to a simpler form of coefficient recovery.

The paper is organized as follows. The over- parameterization framework for scalar signals is presented in Section 2 and our extension of the framework to vector fields is discussed in Section 3. In order to overcome the drawbacks of using TV-based over-parameterization functional, Section 4 reviews our suggested denoising techniques based on the sparsity analysis model. The algorithm for the sparsity-based over-parameterization is specified in Section 5. Section 6 displays some experimental results of the sparsity-based over-parameterization for vector fields compared to the TV-based over-parameterization and the channel-coupled TV regularization. Finally, Section 7 concludes our work and discuses potential future research.

2. The over-parameterization framework. Recovering a function from its noisy and distorted samples is a fundamental problem in both image and signal processing fields. A key ingredient in the success of a recovery method is the set of assumptions that summarizes the prior knowledge of the signal properties, which differentiate it from the noise. The assumptions may range from some general signal properties such as smoothness or piecewise-smoothness to detailed information on the structure of the signal, and can often be expressed in the form of a parameterized model. The over-parameterization framework for model-based noise removal allows designing an objective functional that exploits local fitting of parameterized models with global assumption on their variations.

For the sake of simplicity, we present the over-parameterized model for scalar signals. Let f be a one-dimensional signal we wish to recover from a noisy set of measurements, $y(x) = f(x) + n(x)$, where $n(x)$ is an additive white Gaussian noise. Suppose that the ideal function f can be described by a linear combination of m basis signals selected a priori and defined globally across the signal domain,

$$f(x) = \sum_{i=1}^{m} u_i(x)\phi_i(x), \tag{2.1}$$

where $\{u_i\}_{i=1}^m$ is a set of coefficients and $\{\phi_i\}_{i=1}^m$ is a set of basis signals. For example, we can consider the Taylor approximation as a parameterized model with polynomial basis functions, $\phi_i(x) = x^{i-1}$. In general, for $m > 1$, infinite combinations of coefficients assignments can represent the ideal function at any specific location. This redundancy can be resolved by imposing global regularization on the coefficients of the considered model.

In the over-parameterization framework, appropriate basis signals are such that the true signal could be described by a linear combination of approximately piecewise-constant coefficients, so that most of the local changes in the signal are induced by changes in the basis functions, rather than variations of the coefficients. This can be achieved by imposing some global prior on the coefficients which favors their being piecewise-constant, for instance, a regularization that preserves sharp discontinuities on the variations of the parameters, such as the TV norm. The combination of the local fitting with the global prior on the variations of the coefficients yields the variational form of the over-parameterization functional,

$$u_i^* = \underset{u_i:\mathbb{R}\to\mathbb{R},1\le i\le m}{\arg\min} \int_\Omega \left[\frac{1}{2} \left\| y(x) - \sum_{i=1}^{m} u_i(x)\phi_i(x) \right\|^2 + \sum_{i=1}^{m} \lambda_i \left\| \nabla u_i(x) \right\| \right] dx, \tag{2.2}$$

where u_i^* is the recovered coefficients field for basis signal i, and λ_i is a parameter which weighs the penalty of coefficients field i with respect to the other coefficients fields. This approach utilizes the TV norm in order to treat each coefficients field separately, however, according to the proposed model, all coefficients should be encouraged to have joint discontinuity points. Therefore, [34, 35] suggested to reformulate the recovery problem as

$$u_i^* = \underset{u_i:\mathbb{R}\to\mathbb{R},1\le i\le m}{\arg\min} \int_\Omega \left[\frac{1}{2} \left\| y(x) - \sum_{i=1}^{m} u_i(x)\phi_i(x) \right\|^2 + \left\| [\nabla u_1(x), \nabla u_2(x), \ldots, \nabla u_m(x)] \right\| \right] dx, \tag{2.3}$$

where the prior is a channel-coupled TV norm, which couples the different coefficients fields. For simplicity's sake, we avoid writing explicit weights for the coefficients variations in this work, since we may equivalently scale the basis signals.

A first attempt at the over-parameterization variational methodology was made in [34], when an over-parameterized model based on total variation regularization was proposed for image denoising. This model was later extended to handle optical flow problems in [35] and [39], adjusting itself to various assumptions regarding the flow pattern. The technique was revised in [44] and great robustness and accuracy improvements were demonstrated on simulated one-dimensional signals and images by using a non-local data term combined with the Ambrosio-Tortorelli and TV regularizer. The latest work in this subject was introduced by [21] and combined the over-parameterized variational strategy with the sparse representation methodology. Assigning a sparse prior enabled improved results for denoising of simulated one-dimensional signals and cartoon images as well as segmentation of piecewise-linear images.

3. The over-parameterization framework for vector fields. A vector field is a set of vector objects that can be described by two properties: magnitude (or length) and direction. Elementary linear algebra describes how to add vectors to each other, scale vectors and rotate them to specific directions in the Euclidean space. In order to apply a generalized over-parameterization techniques suited for vector fields reconstruction, we propose to parameterize the vector field by a linear combination of rotated and scaled basis vector fields. Let $\mathbf{f}$ be an n-dimensional Euclidean vector field with n components we wish to recover from a noisy set of measurements $\mathbf{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{n}(\mathbf{x})$, where $\mathbf{n}(\mathbf{x})$ is an additive white Gaussian noise. Suppose that the ideal vector field f can be described by a linear combination of m basis vector fields selected a priori and defined globally across the signal domain, $\boldsymbol{\phi}_i$, $i = 1, 2, \ldots$. For example, we can consider the Taylor approximation as a parameterized model with polynomial magnitude basis vector fields. The vector field can be parameterized by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} s_i(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x}), \tag{3.1}$$

where m is the number of basis fields used for the approximation, $\{s_i\}_{i=1}^m$ is a set of scaling coefficients and $\{\mathbf{R}_i\}_{i=1}^m$ is a set of rotation transformations, which can be describe by the elements of the Lie group $SO(n)$, the special-orthogonal group in $\mathbb{R}^n$.

Following the over-parameterization framework, appropriate basis vector fields are ones that provide a good estimation for vector field $\mathbf{f}$ by a linear combination of approximately piecewise constant parameters $\{s_i\}_{i=1}^m$ and $\{\mathbf{R}_i\}_{i=1}^m$ with joint discontinuity points. This may be achieved by imposing some global prior on the parameters. This over-parameterized formulation has three main advantages over direct signal recovery: First, the variables are assumed to be piecewise constant, thus have a simpler form compared to the ideal signal, and are easier to recover. Second, the coefficients regularization process becomes meaningful, since many physical flow properties can be described by constant coefficient. For example: laminar flow properties can be described by second order polynomial basis fields and constant coefficients. And third, the recovery of the signal is obtained along with the recovery of its coefficients.

Figures 1 and 2 demonstrate the over-parameterization representation for vector fields. The color map represents the magnitude of each vector while the orientation is presented using the arrows. Figure 1 presents a two-dimensional vector field which can be well constructed by the first-order Taylor approximation. Given this prior knowledge we define the following two-dimensional basis vector fields: a constant magnitude vector field, a linearly varying magnitude in the horizontal direction and a linearly varying magnitude in the vertical direction, as displayed in Sub-figures 2a, 2b and 2c, respectively. This signal can be constructed by constant scaling and rotation coefficients, as presented in Sub-figures 2d, 2e, 2f, 2g, 2h and 2i, where the rotation transformations are represented by the angles by which the vectors are rotated counterclockwise about the z-axis. The original signal can be estimated by summing the scaled and rotated basis vector fields shown in Sub-figures 2j, 2k and 2l.

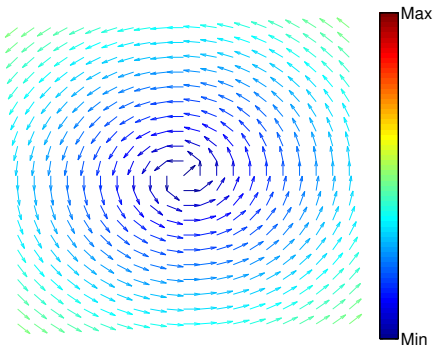


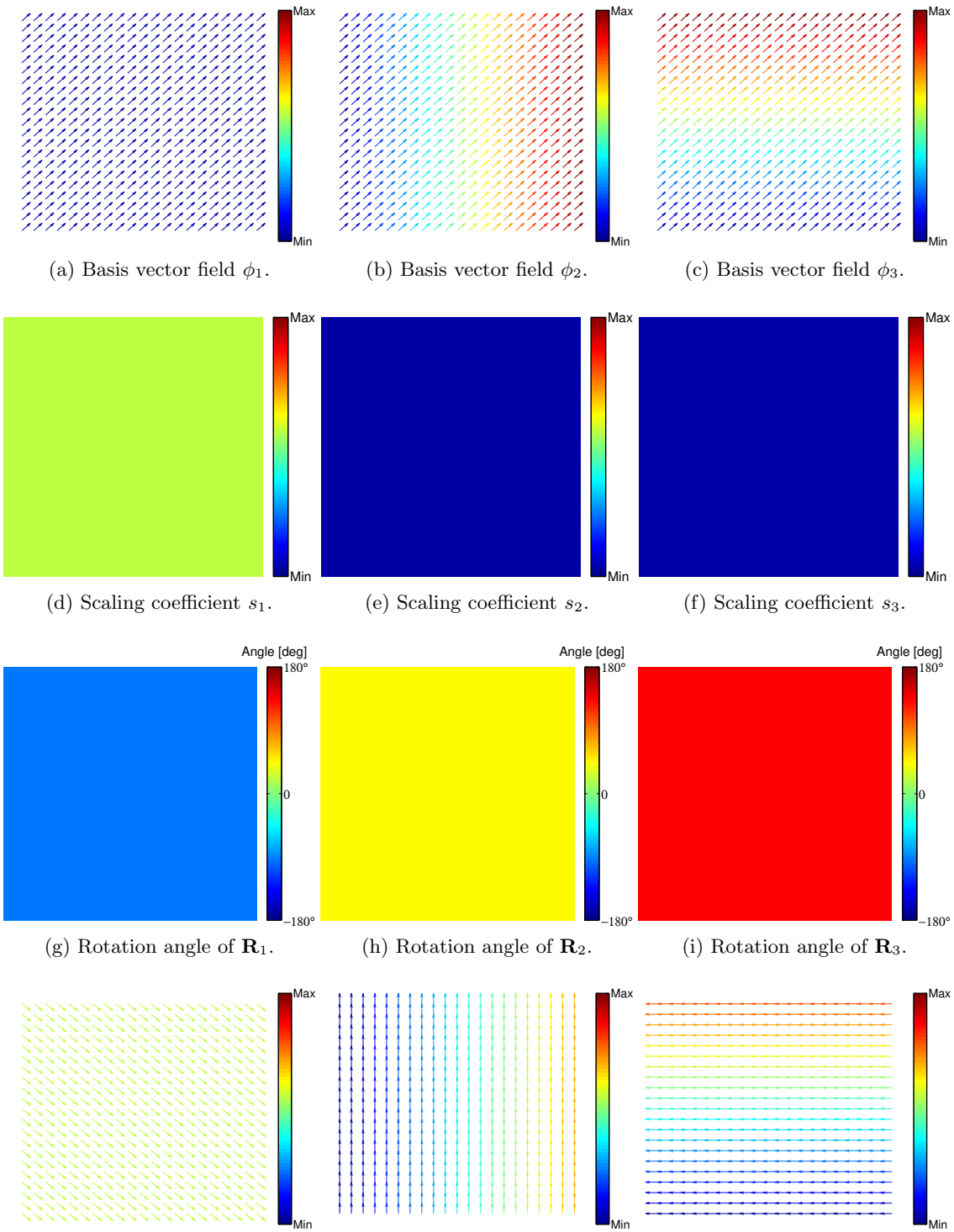
Figure 1. *Two-dimensional vector field.*

3.1. The special-orthogonal group. A Lie group G is a group endowed with the structure of a differentiable manifold, such that the inversion map, $s : G \to G$, $g \mapsto g^{-1}$ and the multiplication map, $m : G \times G \to G$, $(g, h) \mapsto gh$ are smooth. This structure allows Lie group elements and their neighborhoods to be mapped onto a neighborhood of the identity element by the group action with their inverses. Each Lie group has a corresponding Lie-algebra, $\mathfrak{g}$, which is defined on the tangent space to the Lie group at the identity. The Lie-algebra is a vector space equipped with an anti-symmetric bilinear operator, known as Lie-bracket, describing the non-commutative part of the group product. It can be considered as a linearization of the Lie group around the identity, thus allowing to define differentiation on the Lie group.

In this work, we are interested in a specific Lie group called the special-orthogonal group, $SO(n)$, and its related matrix manifold. Lie group $SO(n)$ is the group of rotations, describing all orientation-preserving isometries of the n-dimensional Euclidean space. Elements of the group can be represented by an $n \times n$ orthogonal matrix with determinant one:

$$SO(n) = \{\mathbf{R} \in \mathbb{R}_{n\times n}, \mathbf{R}^T\mathbf{R} = \mathbf{I}, \det(\mathbf{R}) = 1\} \tag{3.2}$$

The corresponding Lie-algebra of the $SO(n)$ group is the $so(n)$ space, which can be descried by the set of skew-symmetric matrices,



(a) Basis vector field ϕ_1. (b) Basis vector field ϕ_2. (c) Basis vector field ϕ_3.

(d) Scaling coefficient s_1. (e) Scaling coefficient s_2. (f) Scaling coefficient s_3.

(g) Rotation angle of $\mathbf{R}_1$. (h) Rotation angle of $\mathbf{R}_2$. (i) Rotation angle of $\mathbf{R}_3$.

(j) $s_1\mathbf{R}_1\phi_1$. (k) $s_2\mathbf{R}_2\phi_2$. (l) $s_3\mathbf{R}_3\phi_3$.

Figure 2. *Over-parameterized representation of the two-dimensional vector field presented in Figure 1.*

$$so(n) = \{\mathbf{A} \in \mathbb{R}_{n\times n}, \mathbf{A}^T = -\mathbf{A}\}. \tag{3.3}$$

For an in-depth discussion on Lie groups, Lie-algebra and the $SO(n)$ group, we refer the reader to standard literature [24]. The subject of regularization of Lie groups is discussed in [15, 23, 40].

3.2. TV-based over-parameterization for vector fields. Imposing the suggested model on the over-parameterization framework, the reconstructed field $\mathbf{f}^*$ can be found by minimizing the following functional:

$$\underset{\substack{s_i:\mathbb{R}^d\to\mathbb{R},\\ \mathbf{R}_i:\mathbb{R}^d\to SO(n),\\ 1\le i\le m}}{\arg\min} \int_\Omega \left[\tfrac{1}{2}\left\|\mathbf{y}(\mathbf{x}) - \sum_{i=1}^m s_i(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})\right\|^2 + \lambda_s \boldsymbol{\psi}_s(\mathbf{s}) + \lambda_R \boldsymbol{\psi}_R(\mathbf{R})\right] d\mathbf{x}, \tag{3.4}$$

where,

$$\boldsymbol{\psi}_s(\mathbf{s}) = \|[\nabla s_1(\mathbf{x}), \nabla s_2(\mathbf{x}), \ldots, \nabla s_m(\mathbf{x})]\| \tag{3.5}$$

induces a channel-coupled TV norm, and

$$\boldsymbol{\psi}_R(\mathbf{R}) = \|[\nabla \mathbf{R}_1(\mathbf{x}), \nabla \mathbf{R}_2(\mathbf{x}), \ldots, \nabla \mathbf{R}_m(\mathbf{x})]\| \tag{3.6}$$

induces a vectorial TV norm on the elements in the embedding of the Lie group into Euclidean space. In this prior, $\|\cdot\|$ is the Frobenius norm and $\nabla\mathbf{R}$ denotes the Jacobian of $\mathbf{R}$, described as a column-stacked vector $\mathbb{R}^{n^2}$. We note that the same notation is used throughout this work in order to represent the Lie group element, its matrix representation and the embedding into Euclidean space, as specified in each case. Ω is the vector field domain, and λ_s and λ_R describe the relative strength of the priors. The reconstructed vector field is therefore $\mathbf{f}^*(\mathbf{x}) = \sum_{i=1}^m s_i^*(\mathbf{x})\mathbf{R}_i^*(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})$.

The formulation of the suggested regularization term for the $SO(n)$ group is achieved by replacing the equivalent of the total-variation regularization approach in terms of Lie-algebra, $\left\|\mathbf{R}_i^{-1}(\mathbf{x})\nabla\mathbf{R}_i(\mathbf{x})\right\|$, with a regularization of an embedding of the Lie group into Euclidean space and adapting the channel-coupled TV norm to form the over-parameterization regularizer, $\|[\nabla\mathbf{R}_1(\mathbf{x}), \nabla\mathbf{R}_2(\mathbf{x}), \ldots, \nabla\mathbf{R}_m(\mathbf{x})]\|$. The replacement is justified by the fact that $SO(n)$ group is an isometry of the n-dimensional Euclidean space [40]. Therefore, if the constraint $\mathbf{R}_i(\mathbf{x}) \in SO(n), \forall\mathbf{x}\in\Omega$ is approximately fulfilled, then $\|\nabla\mathbf{R}(\mathbf{x})\| \approx \left\|\mathbf{R}_i^{-1}(\mathbf{x})\nabla\mathbf{R}_i(\mathbf{x})\right\|$.

Minimizing cost function (3.4) is not a straight-forward task as it combines the minimization of two variables, the scaling and the rotation coefficients. In addition, the rotation group domain is non-convex and the regularizers are defined on the border between convex and non-convex functions.

4. Sparsity-based over-parameterization model for vector fields. The main advantage of the over-parameterization framework is its wide solution domain. While using extra variables, the over-parameterized model is usually more naturally suited to describe the signal structure, thus often enabling convergence to excellent denoising solutions [34, 35, 39]. However,

the constraints imposed on the solution domain via the TV regularization cannot guaranty convergence to a piecewise-constant parameters solution, which may lead to poor recovery results of the signal and its parameters, as demonstrated in [44].

In addition to the known shortcomings, the total variation regularizer may also produce blunt edges. The regularizer considers the total signal difference while not differentiating smooth gradual signal change from sharp discontinuity. Hence, it will not provide an actual piecewise-constant solution as required by the over-parameterization model. Another weakness of the functional is termed "origin biasing". Since the basis signals are defined globally across the domain, there has to be some fixed arbitrary origin. Changing the origin will vary the reconstructed coefficients and will also affect the value of the regularization term. In an attempt to overcome these weaknesses and refine the discontinuities, we consider a sparsity-based technique.

4.1. The sparse analysis model. The sparse analysis model is a popular representation model used in many signal and image processing applications. For the sake of simplicity, we first present this model for scalar signals. Let f be an n-dimensional signal we wish to recover from a noisy set of measurements $\mathbf{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{n}(\mathbf{x})$, where $\mathbf{n}(\mathbf{x})$ is an additive white Gaussian noise, and $\mathbf{\Omega}$ is a possibly redundant analysis operator. The analysis model considers the behavior of the analysis vector $\mathbf{\Omega f}$ and assumes it is sparse. A common analysis operator is the finite difference operator $\mathbf{\Omega}_{DIF}$ which concatenates the directional derivatives of a signal and is closely related to total variation. Thus $\mathbf{f}$ can be recovered by solving

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} \|\mathbf{\Omega f}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{f}\|_2 \leq \epsilon. \tag{4.1}$$

The zeros in $\mathbf{\Omega f}^*$ correspond to row vectors in the analysis operator. It can be interpreted as a subspace characterized by the zeros, in which the recovered signal $\mathbf{f}^*$ is orthogonal to the appropriate row of analysis operator $\mathbf{\Omega}$. We say that $\mathbf{f}$ is cosparse under $\mathbf{\Omega}$ with a cosupport Λ if $\mathbf{\Omega}_\Lambda f = 0$, where $\mathbf{\Omega}_\Lambda$ is a sub-matrix of $\mathbf{\Omega}$ consisting of the rows corresponding to Λ. Many approximation techniques were suggested in the recent years in order to solve this NP-hard problem.

4.2. Over-parameterization for vector fields via the sparse analysis model. Once understanding the TV-based framework disadvantages, we revisit the proposed model and propose a novel sparsity-based over-parameterization framework for vector field denoising. Let us denote the fields of similarity-like transformation coefficients, consisting of rotation and scaling, as $\mathbf{A}_i(\mathbf{x}) = s_i(\mathbf{x})\mathbf{R}_i(\mathbf{x}), \forall i = 1, \ldots, m$. The desired vector field can now be represented as:

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} \mathbf{A}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x}). \tag{4.2}$$

where m is the number of basis fields used for the approximation.

Using a sparse analysis model combined with the over-parameterization framework, the reconstructed vector field $\mathbf{f}^*$ can be found by solving the following problem:

$$\arg\min_{\substack{\mathbf{a}_{i,j},\\ 1\leq i\leq m,\\ 1\leq j\leq n^2}} \left\| \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} |\mathbf{\Omega}_{DIF}\mathbf{a}_{i,j}| \right) \right\|_0 \quad s.t. \quad \left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\|_2 \leq \epsilon, \tag{4.3}$$

where $\mathbf{a}_{i,j}$ are column-stacked vector representations of the j-th matrix elements of transformation coefficients functions $\mathbf{a}_i$. The prior enforces by abuse of notation the finite-difference operator on the coefficients and its accurate formulation is discussed in the squeal. The parameter ϵ denotes an upper-bound for the noise energy level. The multiplication $\mathbf{C}_i\mathbf{a}_i^T$ is a generalization of $\mathbf{C}_i(\mathbf{x})\mathbf{a}_i^T(\mathbf{x})$ for each point $\mathbf{x}$, such that $\mathbf{C}_i(\mathbf{x})\mathbf{a}_i^T(\mathbf{x})$ equals $\mathbf{A}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})$, where $\mathbf{C}_i(\mathbf{x})$ is an $n \times n^2$ matrix and $\mathbf{a}_i(\mathbf{x})$ is the vectorization of a scaled-rotation matrix $\mathbf{A}_i(x)$. The reconstructed vector field is therefore $\mathbf{f}^* = \sum_{i=1}^m \mathbf{C}_i {\mathbf{a}_i}^{*T}$.

The first term in problem (4.3) enforces a jointly sparse solution of the model coefficients under the analysis finite-difference operator $\boldsymbol{\Omega}_{DIF}$, considering all the matrix coefficients components jointly. Each non-zero location in $\boldsymbol{\Omega}_{DIF}\mathbf{a}_{i,j}$ indicates a discontinuity in the piecewise-constant coefficient $\mathbf{a}_i$. By indicating the discontinuities in the signal model, the sparsity constraint prevents diffusivity interaction between different regions. The constraint uses the known upper-bound of the noise energy level in order to require that the recovered signal is as close to the noisy sample as the ideal signal, encouraging the solution to be similar to the original signal. The cost associated with the solution of problem (4.3) is directly due to the discontinuity set, hence the penalty no longer depends on the size of the variations at discontinuities.

In [21] it was suggested to solve the NP-hard problem of the basic scalar-valued over-parameterization problem by a generalization of the GAPn algorithm [33] called block GAPn (BGAPn). In this work we extend the BGAPn algorithm to handle scaled-rotation coefficients by adding auxiliary fields and appropriate constraints.

5. Sparsity-based over-parameterization variational algorithm. We generalize BGAPn algorithm in [21] to support other domains, specifically our over-paremeterization problem (4.3), including matrix-valued map regularization. The inner optimization step in the BGAPn algorithm is replaced with a similarity-like transformation coefficients estimation. This inner optimization problem is solved via ADMM. We introduce this extension in Algorithm 5.1.

The algorithm aims to solve the sparsity-based over-parameterization model (4.1) by identifying the cosupport Λ in a greedy way, as suggested in [21]. Since the same cosupport is used for all the elements of coefficients $\mathbf{a}_i(\mathbf{x})$, the coefficient parameters are encouraged to be piecewise constant while having joint discontinuity locations. The iterative scheme is initialized by including all the rows of the anaylsis operator $\boldsymbol{\Omega}$ in Λ. In each iteration a new solution is estimated and allows to find elements which correspond to a non-zero entry in $\boldsymbol{\Omega}_{\Lambda_k}\mathbf{a}_{i,j}^k$, where k is the iteration number. The row of $\boldsymbol{\Omega}$ correlating to the maximum magnitude of $\sum_{i=1}^m \left(\sum_{j=1}^{n^2} |\boldsymbol{\Omega}_{DIF}\mathbf{a}_{i,j}|\right)$ indicate a discontinuity and is thus removed from the cosupport. In order to accelerate the algorithm, multiple rows can be removed from the cosupport at each iteration.

Stopping criteria include the stability of the solution or the size of the cosupport. Another useful stopping criterion is $\|\boldsymbol{\Omega}_\Lambda \mathbf{a}_{i.j}\| < e$, where e is a small constant. The latter was used in our experiments. It is important to note that there are no known recovery guarantees for the greedy algorithm.

Algorithm 5.1 Sparsity-based over-parameterization for vector fields.

Input: noisy measurements $\mathbf{y}$; analysis operator $\mathbf{\Omega}_{DIF}$; noise level upper-bound ϵ; iteration numerator $k = 0$; cosupport Λ_0 consisting of all finite-difference operator rows; basis functions operators $\mathbf{C}_{i,j} : i = 1, \ldots, m, j = 1, \ldots, n^2$.

Output: reconstructed coefficients components $\mathbf{a}^*_{i,j}$; estimated cosupport Λ; estimated signal $\mathbf{f}^*$.

while stopping criteria are not met **do**

Estimate $\mathbf{a}^k_{i,j}$:

$$\mathbf{a}^k_{i,j} = \underset{\substack{\mathbf{a}_{i,j},\\ 1\le i\le m,\\ 1\le j\le n^2}}{\arg\min} \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} \|\mathbf{\Omega}_{\Lambda^k} \mathbf{a}_{i,j}\|_2^2 \right) \; s.t. \left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\|_2 \le \epsilon \tag{5.1}$$

Update the cosupport:

$$\Lambda^{k+1} = \Lambda^k \setminus \left\{ \arg\max \left\{ \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} \left\| \boldsymbol{\omega}^l \mathbf{a}^k_{i,j} \right\|^2 \right) : l \in \Lambda^k \right\} \right\}$$

Update iteration number: $k = k + 1$

end while

$\mathbf{a}^*_{i,j} = \mathbf{a}^k_{i,j}, \quad \forall i = 1, \ldots, m, j = 1, \ldots, n^2.$

Form the reconstructed signal:

$$\mathbf{f}^* = \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^{*T}$$

5.1. Optimization over scaled rotation set. The inner optimization step estimating the solution (5.1) can be simplified by solving the unconstrained problem,

$$\mathbf{a}^k_{i,j} = \underset{\substack{\mathbf{a}_{i,j},\\ 1\le i\le m,\\ 1\le j\le n^2}}{\arg\min} \left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\|_2^2 + \lambda \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} \|\mathbf{\Omega}_{\Lambda_k} \mathbf{a}_{i,j}\|_2^2 \right), \tag{5.2}$$

where $\lambda > 0$ determines the relative weight of the prior. The value of λ is in fact varied until $\mathbf{f}$ satisfies the constraint $\left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\| \le \epsilon$.

As discussed in Section 3.2, an efficient scheme for smoothing over matrix manifolds performs the regularization over group elements which are embedded into the Euclidean space. Adapting the approach suggested in [40] to the scaled-rotation coefficients, we add auxiliary variables $\mathbf{b}_i$, such that $\mathbf{b}_i(\mathbf{x}) = \mathbf{a}_i(\mathbf{x})$ and restrict $\mathbf{b}_i(\mathbf{x}) \in \mathbb{R}_+ \cdot SO(n)$ and $\mathbf{a}_i(\mathbf{x}) \in \mathbb{R}^{n^2}$. We obtain the equality constraints by using augmented Lagrangian terms added to the cost

Algorithm 5.2 ADMM algorithm for optimizing augmented Lagrangian (5.3).

Input: noisy vector field measurements $\mathbf{y}$; constant λ; initial guess $\mathbf{b}^0_{i,j}, \boldsymbol{\xi}^0_{i,j}, \forall i = 1, \ldots, m, j = 1, \ldots, n^2$

Output: recovered scaled rotation coefficients $\mathbf{a}_{i,j}, \forall i = 1, \ldots, m, j = 1, \ldots, n^2$

for $t = 1, 2, \ldots,$ until convergence **do**

Regularization update step:

$$\mathbf{a}^t_{i,j}(x) = \arg\min_{\mathbf{a}_{i,j}} L_c\left(\mathbf{a}_{1,1}, \ldots, \mathbf{a}_{m,n^2}, \mathbf{b}^{t-1}_{1,1}, \ldots, \mathbf{b}^{t-1}_{m,n^2}, \boldsymbol{\xi}^{t-1}_{1,1}, \ldots, \boldsymbol{\xi}^{t-1}_{m,n^2}\right), \quad \forall i = 1, \ldots, m, j = 1, \ldots, n^2 \tag{5.4}$$

Projection step:

$$\mathbf{b}^t_{i,j}(x) = \arg\min_{\mathbf{b}_{i,j}} L_c\left(\mathbf{a}^t_{1,1}, \ldots, \mathbf{a}^t_{m,n^2}, \mathbf{b}_{1,1}, \ldots, \mathbf{b}_{m,n^2}, \boldsymbol{\xi}^{t-1}_{1,1}, \ldots, \boldsymbol{\xi}^{t-1}_{m,n^2}\right), \quad \forall i = 1, \ldots, m, j = 1, \ldots, n^2 \tag{5.5}$$

Update Lagrange multipliers:

$$\boldsymbol{\xi}^t_{i,j} = \boldsymbol{\xi}^{t-1}_{i,j} + c\left(\mathbf{a}^t_{i,j} - \mathbf{b}^t_{i,j}\right), \quad \forall i = 1, \ldots, m, j = 1, \ldots, n^2 \tag{5.6}$$

end for

function. The resulting saddle-point problem is

$$\begin{aligned}
&\min_{\substack{\mathbf{a}_{i,j}, \mathbf{b}_{i,j}, \\ 1 \le i \le m, \\ 1 \le j \le n^2}} \max_{\substack{\boldsymbol{\xi}_{i,j} \\ 1 \le i \le m, \\ 1 \le j \le n^2}} L_c\left(\mathbf{a}_{1,1}, \ldots, \mathbf{a}_{m,n^2}, \mathbf{b}_{1,1}, \ldots, \mathbf{b}_{m,n^2}, \boldsymbol{\xi}_{1,1}, \ldots, \boldsymbol{\xi}_{m,n^2}\right) = \\
&\min_{\substack{\mathbf{a}_{i,j}, \mathbf{b}_{i,j}, \\ 1 \le i \le m, \\ 1 \le j \le n^2}} \max_{\substack{\boldsymbol{\xi}_{i,j} \\ 1 \le i \le m, \\ 1 \le j \le n^2}} \left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\|^2 + \lambda \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} \|\boldsymbol{\Omega}_{\Lambda_k} \mathbf{a}_{i,j}\|^2 \right) \\
&\qquad + \frac{c}{2} \sum_{i=1}^{m} \sum_{j=1}^{n^2} \|\mathbf{a}_{i,j} - \mathbf{b}_{i,j}\|^2 + \sum_{i=1}^{m} \sum_{j=1}^{n^2} \left\langle \boldsymbol{\xi}_{i,j}, \mathbf{a}_{i,j} - \mathbf{b}_{i,j} \right\rangle,
\end{aligned} \tag{5.3}$$

where $\boldsymbol{\xi}_{i,j}$ are the Lagrange multipliers of the j-th components of the i-th equality constraint in column stack representation, and c is a positive constant. This problem can be solved iteratively via the ADMM algorithm consisting of alternating minimization steps with respect to $\mathbf{a}_{i,j}$, $\mathbf{b}_{i,j}$ and $\boldsymbol{\xi}_{i,j}$, as described in Algorithm (5.2). This technique separates the optimization process into a regularization update step of a map onto an embedding space, and a per point projections step. The algorithm can be made locally convergent with minor modifications, as shown in [40].

5.1.1. Regularization update step. Minimizing the augmented Lagrangian (5.3) with respect to $\mathbf{a}_{i,j}$, we get the regularization update step of the embedded signal.

$$\mathbf{a}_{i,j} = \underset{\substack{\mathbf{a}_{i,j},\\ 1\le i\le m,\\ 1\le j\le n^2}}{\arg\min} \left\| \mathbf{y} - \sum_{i=1}^{m} \mathbf{C}_i \mathbf{a}_i^T \right\|^2 + \lambda \sum_{i=1}^{m} \left(\sum_{j=1}^{n^2} \left\| \mathbf{\Omega}_{\Lambda_k} \mathbf{a}_{i,j} \right\|^2 \right) + \frac{c}{2} \sum_{i=1}^{m} \sum_{j=1}^{n^2} \left\| \mathbf{a}_{i,j} - \mathbf{b}_{i,j} \right\|^2 + \sum_{i=1}^{m} \sum_{j=1}^{n^2} \left\langle \boldsymbol{\xi}_{i,j}, \mathbf{a}_{i,j} - \mathbf{b}_{i,j} \right\rangle . \tag{5.7}$$

The optimal $\mathbf{a}_{i,j}$ can be found by solving the appropriate Euler-Lagrange equation,

$$2\mathbf{C}_{i,j}^T \left(\sum_{l=1}^{m} \mathbf{C}_l \mathbf{a}_l^T - \mathbf{y} \right) + c\left(\mathbf{a}_{i,j} - \mathbf{b}_{i,j}\right) + \boldsymbol{\xi}_{i,j} + 2\lambda \mathbf{\Omega}_\Lambda^T \mathbf{\Omega}_\Lambda \mathbf{a}_{i,j} = 0, \tag{5.8}$$

5.1.2. Projection onto scaled-rotation domain. Once we found the embedded solution to problem (5.4), we proceed to the second update step in the ADMM algorithm presented, which turns the group constraint into a simple projection operator. The minimization of the Lagrangian with respect to $\mathbf{b}_{i,j}$ is

$$\mathbf{b}_{i,j} = \underset{\substack{\mathbf{b}_{i,j},\\ 1\le i\le m,\\ 1\le j\le n^2}}{\arg\min} \frac{c}{2} \sum_{i=1}^{m} \sum_{j=1}^{n^2} \left\| \mathbf{a}_{i,j} - \mathbf{b}_{i,j} \right\|^2 + \sum_{i=1}^{m} \sum_{j=1}^{n^2} \left\langle \boldsymbol{\xi}_{i,j}, \mathbf{a}_{i,j} - \mathbf{b}_{i,j} \right\rangle , \tag{5.9}$$

for fixed $\mathbf{a}_{i,j}$ and $\boldsymbol{\xi}_{i,j}$. This problem can be split into m separate sub-problems, one problem for each basis field. It is actually more reasonable to solve these problem for each matrix coefficient separately. Each of these sub-problems reduces to a projection problem per point on the domain, for each basis field separately:

$$\begin{aligned} \mathbf{b}_i &= \underset{\mathbf{b}_i(\mathbf{x}):\mathbb{R}^d\to\mathbb{R}_+\cdot SO(n)}{\arg\min} \frac{c}{2} \left\| \mathbf{a}_i(\mathbf{x}) - \mathbf{b}_i(\mathbf{x}) \right\|^2 + \left\langle \boldsymbol{\xi}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x}) - \mathbf{b}_i(\mathbf{x}) \right\rangle \\ &= \underset{\mathbf{b}_i(\mathbf{x}):\mathbb{R}^d\to\mathbb{R}_+\cdot SO(n)}{\arg\min} \frac{c}{2} \left\| \mathbf{b}_i(\mathbf{x}) - \left(\frac{\boldsymbol{\xi}_i(\mathbf{x})}{c} + \mathbf{a}_i(\mathbf{x}) \right) \right\|^2 \\ &= \mathrm{Proj}_{\mathbb{R}_+\cdot SO(n)} \left(\frac{\boldsymbol{\xi}_i(\mathbf{x})}{c} + \mathbf{a}_i(\mathbf{x}) \right). \end{aligned} \tag{5.10}$$

The projection onto $\mathbb{R}_+ \cdot SO(n)$ can be easily computed by means of singular value decomposition (SVD) per each point separately. Let $\mathbf{U}_i(\mathbf{x})\mathbf{S}_i(\mathbf{x})\mathbf{V}_i(\mathbf{x})^T = \left(\frac{\boldsymbol{\xi}_i(\mathbf{x})}{c} + \mathbf{a}_i(\mathbf{x})\right)$ be the SVD of $\frac{\boldsymbol{\xi}_i(\mathbf{x})}{c} + \mathbf{a}_i(\mathbf{x})$. The projection is

$$\mathbf{b}_i(\mathbf{x}) = \mathrm{Proj}_{\mathbb{R}_+\cdot SO(n)} \left(\frac{\boldsymbol{\xi}_i(\mathbf{x})}{c} - \mathbf{a}_i(\mathbf{x}) \right) = \mathbf{U}_i(\mathbf{x})\tilde{\mathbf{S}}_i(\mathbf{x})\mathbf{V}_i(\mathbf{x})^T, \tag{5.11}$$

where $\tilde{\mathbf{S}}_i(\mathbf{x}) = \sqrt{\left(\sum_{l=1}^{n} \left(\mathbf{S}_i(\mathbf{x})\right)_{ll}^2\right)/n} \cdot \mathbf{I}_n$, and $\mathbf{I}_n$ is the identity matrix of size $n \times n$.

5.2. Convergence properties. Global convergence of the algorithm is difficult to prove due to the non-convex nature of the optimization domain. The discontinuous nature of the projection on a non-convex set may cause the algorithm to oscillate and makes it difficult to prove convergence of the iterates. In order to ensure local convergence for this algorithm, we take the approach suggested by [3], and change the steps of the dual decomposition to

$$\begin{aligned}\mathbf{a}_{i,j}^k(\mathbf{x}) = \arg\min_{\mathbf{a}_{i,j}} & L_c\left(\mathbf{a}_{1,1},\ldots,\mathbf{a}_{m,n^2},\mathbf{b}_{1,1}^{k-1},\ldots,\mathbf{b}_{m,n^2}^{k-1},\boldsymbol{\xi}_{1,1}^{k-1},\ldots,\boldsymbol{\xi}_{m,n^2}^{k-1}\right)\\ &+\frac{1}{2\theta_k}\left\|\mathbf{a}_{i,j}-\mathbf{a}_{i,j}^{k-1}\right\|, \quad \forall i=1,\ldots,m, j=1,\ldots,n^2\end{aligned} \tag{5.12}$$

$$\begin{aligned}\mathbf{b}_{i,j}^k(\mathbf{x}) = \arg\min_{\mathbf{b}_{i,j}} & L_c\left(\mathbf{a}_{1,1}^k,\ldots,\mathbf{a}_{m,n^2}^k,\mathbf{b}_{1,1},\ldots,\mathbf{b}_{m,n^2},\boldsymbol{\xi}_{1,1}^{k-1},\ldots,\boldsymbol{\xi}_{m,n^2}^{k-1}\right)\\ &+\frac{1}{2\theta_k}\left\|\mathbf{b}_{i,j}-\mathbf{b}_{i,j}^{k-1}\right\|, \quad \forall i=1,\ldots,m, j=1,\ldots,n^2\end{aligned} \tag{5.13}$$

$$\boldsymbol{\xi}_{i,j}^k = \boldsymbol{\xi}_{i,j}^{k-1} + c\left(\mathbf{a}_{i,j}^k - \mathbf{b}_{i,j}^k\right), \quad \forall i=1,\ldots,m, j=1,\ldots,n^2 \tag{5.14}$$

where $\frac{1}{2\theta_k}$ denotes the coupling between each iterate and its previous value. The optimization steps in the modified algorithm remain a projection step and a smoothing step, with minor changes in the parameters. The optimal $\mathbf{a}_{i,j}$ can now be found by using the following Euler-Lagrange equation:

$$2\mathbf{C}_{i,j}^T\left(\sum_{l=1}^m \mathbf{C}_l\mathbf{a}_l^T - \mathbf{y}\right) + c\left(\mathbf{a}_{i,j}-\mathbf{b}_{i,j}\right) + \boldsymbol{\xi}_{i,j} + 2\lambda\boldsymbol{\Omega}_\Lambda^T\boldsymbol{\Omega}_\Lambda\mathbf{a}_{i,j} + \frac{1}{\theta_k}\left(\mathbf{a}_{i,j}-\mathbf{a}_{i,j}^{k-1}\right) = 0. \tag{5.15}$$

The minimization with respect to $\mathbf{b}_i(\mathbf{x})$ reduces to the following projection:

$$\mathbf{b}_i(\mathbf{x}) = \mathbf{U}_i(\mathbf{x})\tilde{\mathbf{S}}_i(\mathbf{x})\mathbf{V}_i(\mathbf{x})^T, \tag{5.16}$$

where $\mathbf{U}_i(\mathbf{x})\mathbf{S}_i(\mathbf{x})\mathbf{V}_i(\mathbf{x})^T$ is the SVD of matrix $\frac{\boldsymbol{\xi}_i^{k-1}(\mathbf{x})+c\mathbf{a}_i^k(\mathbf{x})+\frac{\mathbf{b}_i^{k-1}(\mathbf{x})}{\theta_k}}{c+\frac{1}{\theta_k}}$.

A complete convergent analysis is not straightforward and is left as future work. Empirical results demonstrate strong convergence properties for a large variety of θ values.

6. Experimental results. In this section we present several numerical experiments of the sparsity-based over-parameterization noise removal method for vector fields. We consider examples of two-dimensional fields with different models and a varying number of basis fields. Specifically, we examine the suggested framework on two-dimensional vector fields using a first- and second-order Taylor approximation model. We compare our results with the ones of the channel-coupled TV regularization, referred to as TV denoising throughout this section and the TV-based over-parameterization version presented in section 3.2. Finally, we regularize real 3-dimensional data and address the problem of denoising magnetic resonance imaging (MRI) vector field.

The scheme we used to obtain the solution of the TV-based over-parameterization functional includes the alternating minimization technique, where we iterate over minimization

with respect to the rotation transformations and minimization with respect to the scaling coefficients. In order to estimate the rotation transformations, auxiliary variables were added according to the approach suggested in [40] and an ADMM algorithm was incorporated. The minimization process can then be separated into a projection step and a TV regularization step over an embedded field, similar to the method we presented for optimization over the scaled rotation set in 5.1. The TV regularization step is simplified by another ADMM procedure which separates the problem into two additional sub-problems. A related approach is used for the TV regularization of the scaling coefficients. The algorithm is introduced in Appendix A.

Both sparsity-based and TV-based over-parameterization frameworks for vector fields are highly parallelizable, enabling efficient implementation on parallel hardware such as graphics process units (GPU). The frameworks were implemented using Matlab and Python. The parameters of the TV denoising and the TV-based over-parameterization method were tuned separately for each signal, while the same setup of the sparsity-based over-parameterization method was used for all experiments. In addition, the basis fields used for the TV-based over-parameterization were carefully scaled in order to avoid writing explicit relative weights in the prior.

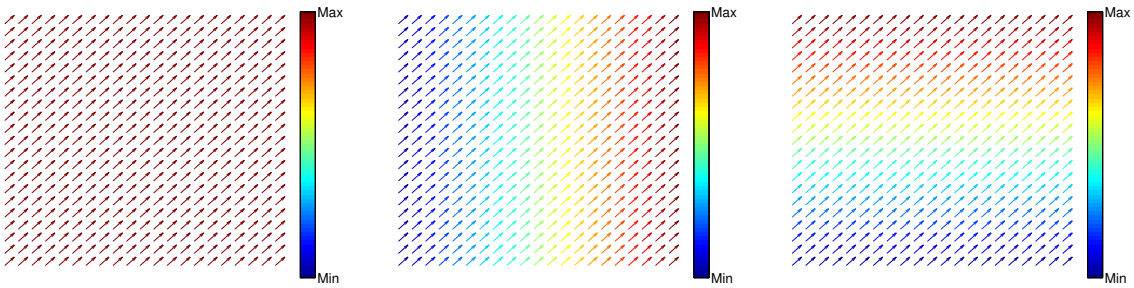
6.1. Two-dimensional vector fields. In order to evaluate the efficiency of our proposed method, we perform several tests on simulated two-dimensional vector fields, where each pixel contains a vector with two components. The simulations are designed to be naturally constructed by a combination of three scaled and rotated basis vector fields: a vector field with a constant magnitude and two vector fields with linearly growing magnitudes, one in the horizontal direction and the other in the vertical direction. The normalized basis fields are presented in Figure 3. After contaminating each channel with white Gaussian noise, we may approximate the scaled rotation coefficients and the desired vector field by solving the sparsity-based over-parameterization problem,

$$\underset{\substack{\mathbf{a}_{i,j},\\ 1\leq i\leq 3,\\ 1\leq j\leq 4}}{\arg\min} \left\| \sum_{i=1}^{3} \left(\sum_{j=1}^{4} |\mathbf{\Omega}_{DIF}\mathbf{a}_{i,j}| \right) \right\|_0 \quad s.t. \quad \left\| \mathbf{y} - \sum_{i=1}^{4} \mathbf{C}_i \mathbf{a}_i^T \right\|_2 \leq \epsilon, \tag{6.1}$$

where $\mathbf{a}_{i,j}$ are column-stacked vector representations of the j-th matrix elements of transformation coefficients function $\mathbf{a}_i$. The prior enforces by abuse of notation the finite-difference operator $\mathbf{\Omega}_{DIF}$ on the coefficients. In addition, $\mathbf{C}_i = [\mathbf{C}_{i,1}, \ldots, \mathbf{C}_{i,4}], m = 1, \ldots, 3$ and $\mathbf{a}_i = \left[\mathbf{a}_{i,1}^T, \ldots, \mathbf{a}_{i,4}^T\right]$. The matrices $\mathbf{C}_i(x)$ contain the basis field data as explained in Section 4.2. The reconstructed signal can be estimated as

$$\mathbf{f}^*(\mathbf{x}) = \sum_{i=1}^{3} \mathbf{a}_i^*(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x}). \tag{6.2}$$

Figures 4 and 5 show the recovery performance of the methods on two-dimensional simulated vector fields. Example 1 demonstrates the reconstruction of a piecewise-linear vector field, divided into four sections, where each section is a combination of a constant vector field,



(a) Basis vector field $\boldsymbol{\phi}_1$ with a constant magnitude.

(b) Basis vector field $\boldsymbol{\phi}_2$ with a linearly growing magnitude in the horizontal direction.

(c) Basis vector field ϕ_3 with a linearly growing magnitude in the vertical direction.

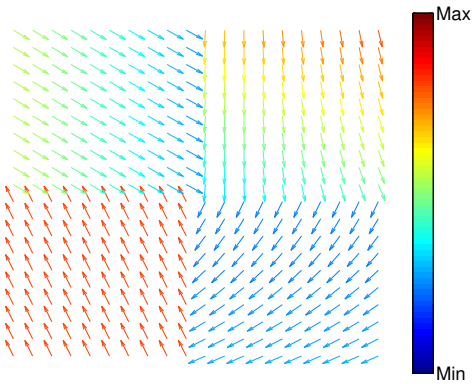
Figure 3. *Normalized basis vector fields for two-dimensional vector field recovery using a first-order Taylor approximation model.*

a horizontally changing vector field and a vertically changing vector field, each possessing different orientation and scaling. Example 2 combines these types of basis vector fields in order to form other patterns.

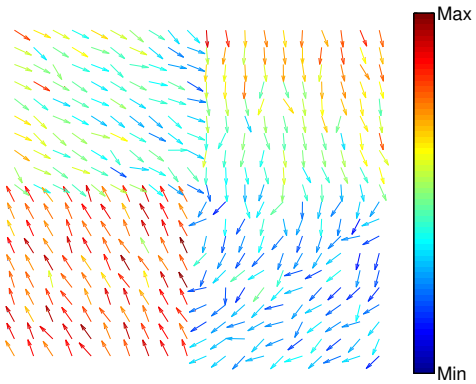
Figures 6 and 7 display the magnitude of the recovered vector fields of examples 1 and 2, while 8 and 9 present the histogram and the cumulative probability of the angular errors, respectively. The sparsity-based over-parameterization framework, outperformed both the TV-based over-parameterization and the channel-coupled TV methods and produced lower angular errors and higher SNR for the two examples. Using the over-parameterization technique with TV priors typically produced higher angular errors and lower SNR, in comparison to the noise removal obtained by the classic channel-coupled TV framework. These results coincide with the previously discussed shortcomings of the TV-based over-parameterization denoising scheme. Specifically, the boundary areas in the TV-based recovery show evident errors which might occur as a result of blunt edges in the reconstructed coefficients map. The staircasing and pixelization effects are also visible in smooth regions. Evidence of these drawbacks can also be found in the literature for some cases of scalar TV-based over-parameterization [44, 21].

6.2. Basis fields selection. Given prior knowledge of the ideal signal, one can design an over-parameterization model comprised of a linear combination of known basis fields that best suit the signal. In the following experiment, we examine the effect of vector field denoising using different basis fields.

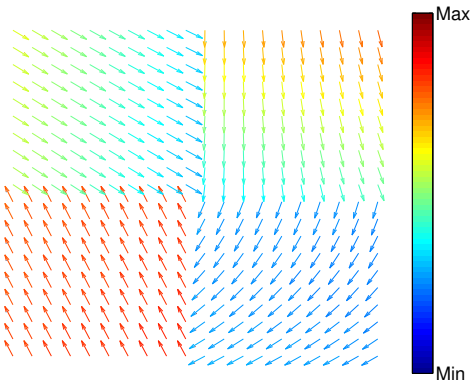
Figure 10 illustrates a synthetic laminar flow field. Laminar flow occurs when a fluid flows through a pipe or between two flat plates, and results in parabolic velocity profile. We compare the recovery performance of the sparsity-based over-parameterization using a first- and a second-order Taylor approximation. The first concluding in 3 basis functions as described in the previous section and the second concluding in 6 two-dimensional basis functions. The recovery results and the resulted angular errors are displayed in figures 10, 11 and 12. For convenience, the vector field plots display only the pixels containing flow information, though the noise was added to the entire field area. Both methods maintained the discontinuities in



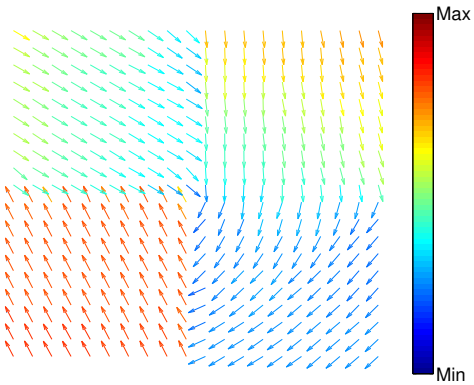
(a) Original vector field.



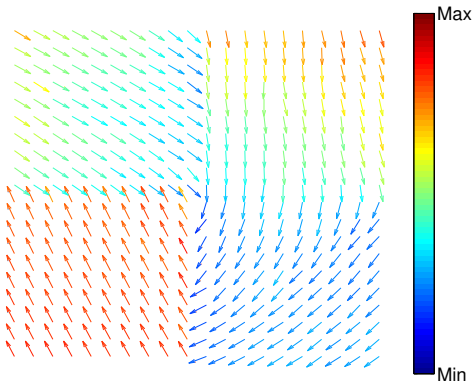
(b) Noisy vector field.
SNR=13.137dB, AAE=6.426±6.001°.



(c) Recovery via sparsity-based over-parameterization.
SNR=28.191dB, AAE=1.662±1.751°.

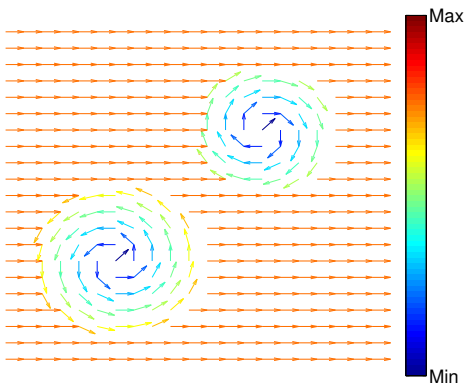


(d) Recovery via TV regularization.
SNR=22.028dB, AAE=2.452±2.870°.

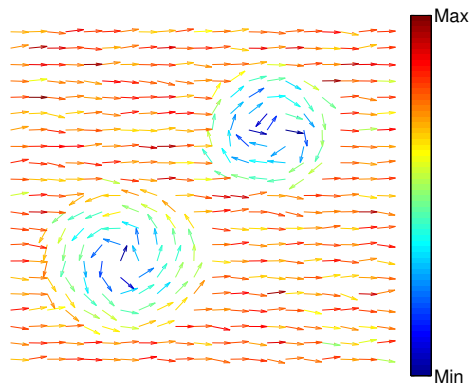


(e) Recovery via TV-based over-parameterization.
SNR=20.504dB, AAE=2.937±2.773°.

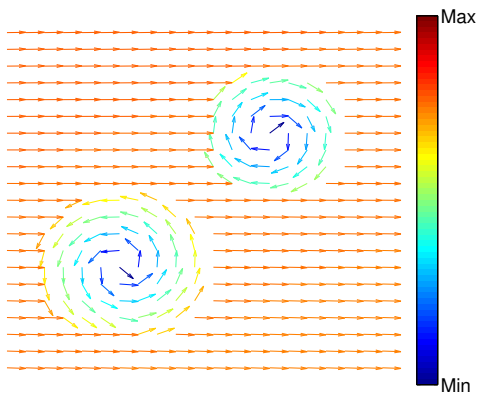
Figure 4. *Example 1 - Recovery of a piecewise-linear two-dimensional vector field.*



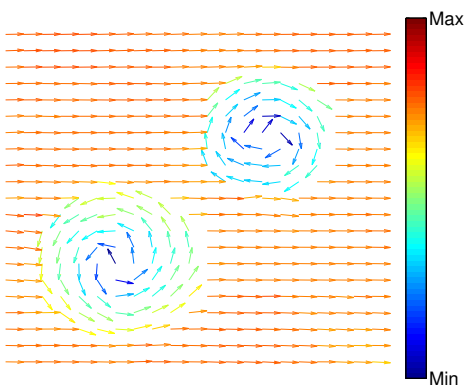
(a) Original vector field.



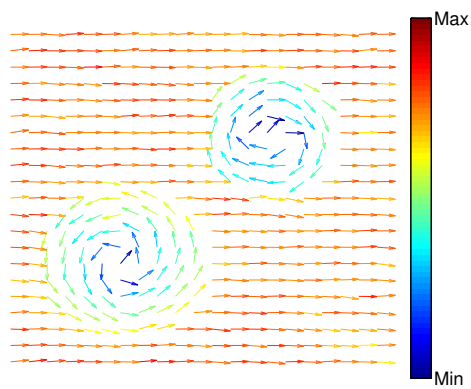
(b) Noisy vector field.
SNR=11.223dB, AAE=4.791±6.935°.



(c) Recovery via sparsity-based over-parameterization.
SNR=22.565dB, AAE=1.262±3.312°.

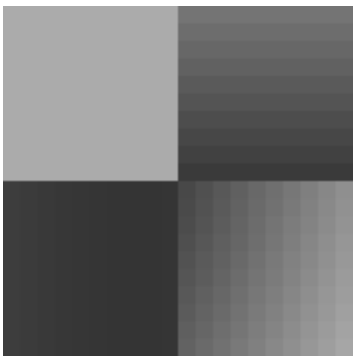


(d) Recovery via TV regularization.
SNR=16.911dB, AAE=2.331±4.386°.

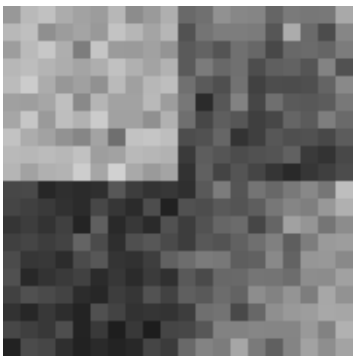


(e) Recovery via TV-based over-parameterization.
SNR=14.556dB, AAE=3.336±5.134°.

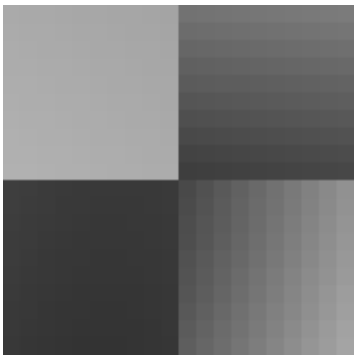
Figure 5. *Example 2 - Recovery of a two-dimensional vector field.*



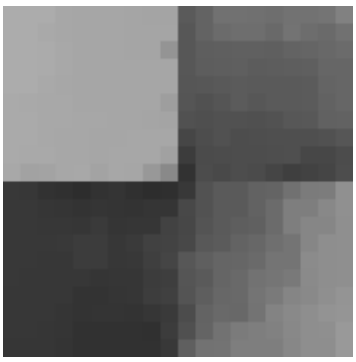
(a) Original magnitude.



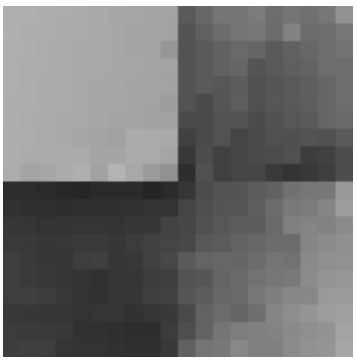
(b) Noisy magnitude.
SNR=13.137dB,
AAE=6.426±6.001°.



(c) Recovery via sparsity-based over-parameterization.
SNR=28.191dB,
AAE=1.662±1.751°.

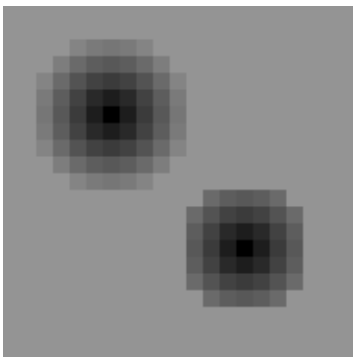


(d) Recovery via TV regularization.
SNR=22.028dB,
AAE=2.452±2.870°.

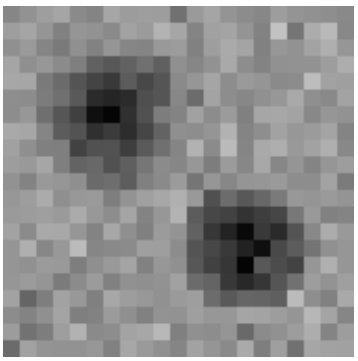


(e) Recovery via TV-based over-parameterization.
SNR=20.504dB,
AAE=2.937±2.773°.

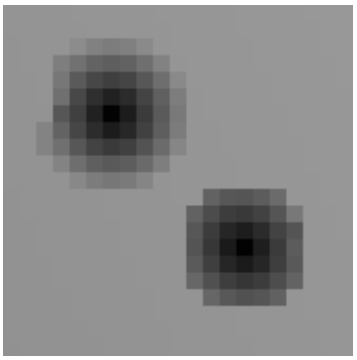
Figure 6. *Example 1 - Magnitude.*



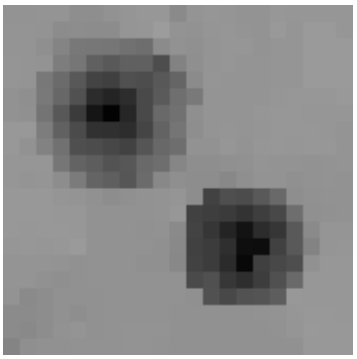
(a) Original magnitude.



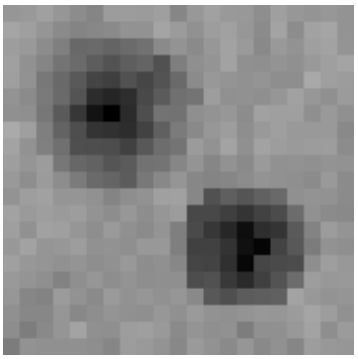
(b) Noisy magnitude.
SNR=11.223dB,
AAE=4.791±6.935°.



(c) Recovery via sparsity-based over-parameterization.
SNR=22.565dB,
AAE=1.262±3.312°.

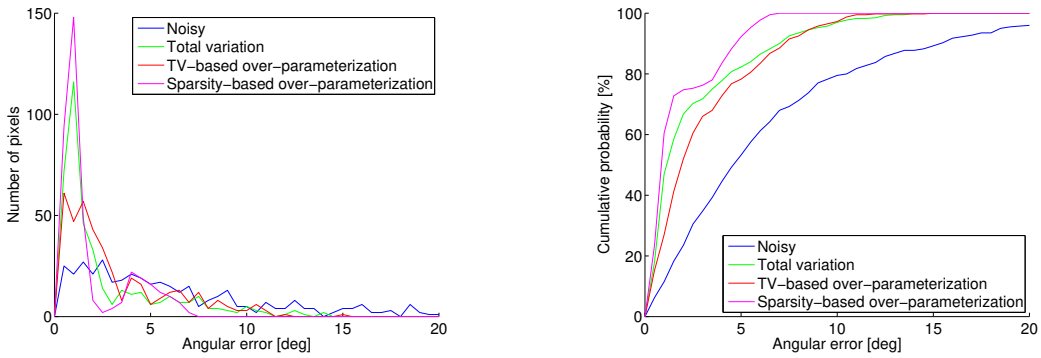


(d) Recovery via TV regularization.
SNR=16.911dB,
AAE=2.331±4.386°.



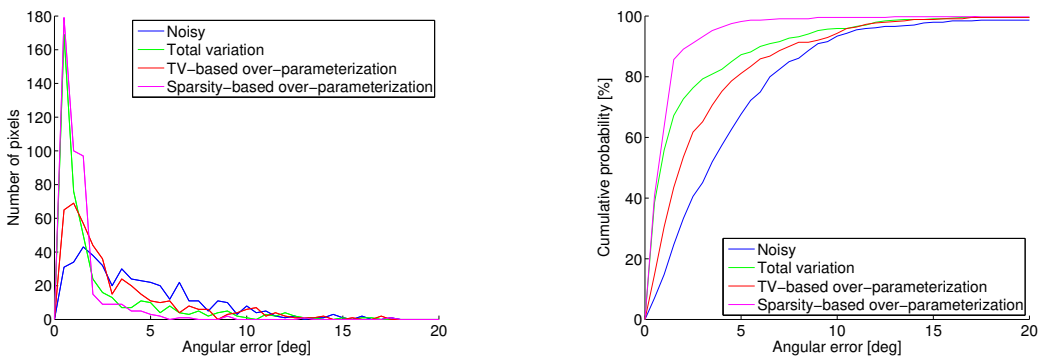
(e) Recovery via TV-based over-parameterization.
SNR=14.556dB,
AAE=3.336±5.134°.

Figure 7. *Example 2 - Magnitude.*



(a) Histogram. (b) Cumulative probability.

Figure 8. *Angular error analysis for Example 1.*



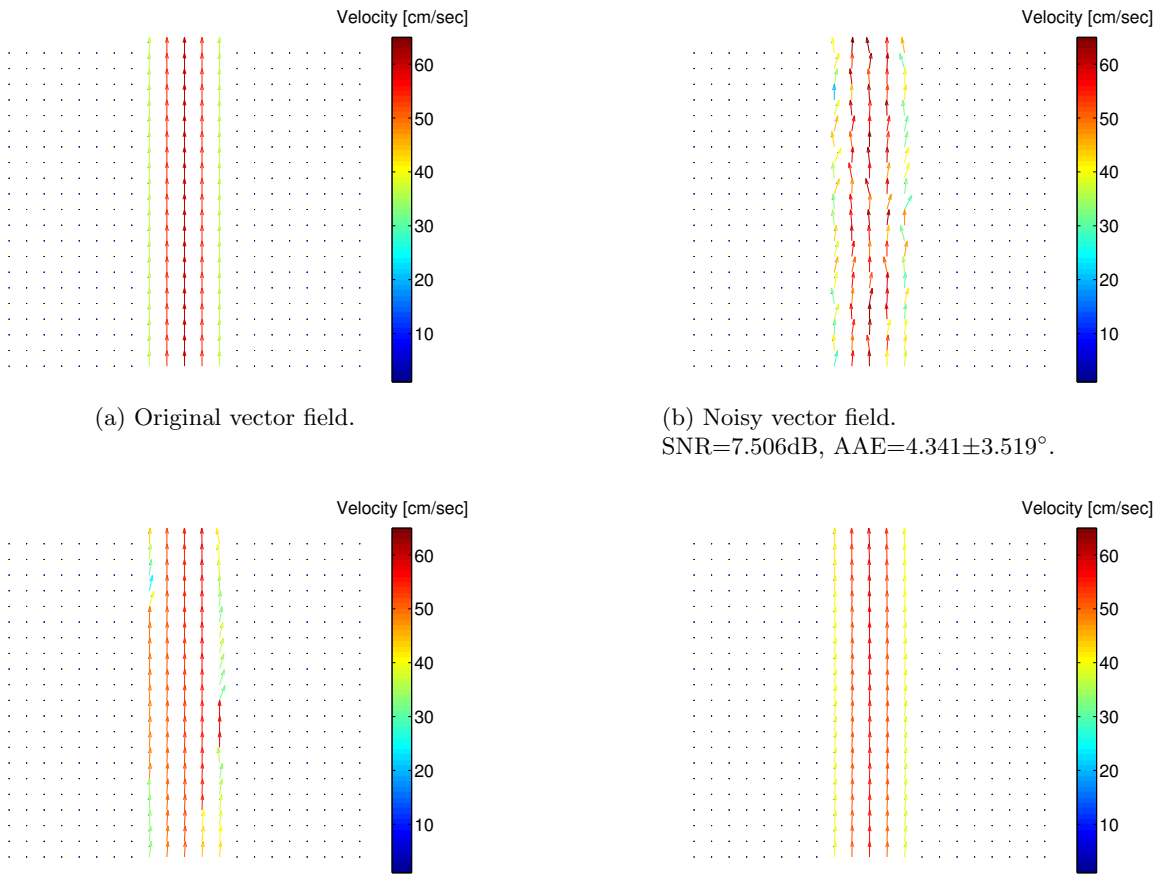
(a) Histogram. (b) Cumulative probability.

Figure 9. *Angular error analysis for Example 2.*

the simulation as well as the laminar flow characteristics, such as approximately steady flow, nearly zero velocity component in the direction normal to the flow and velocity increment toward the center of the flow. However, the second-order Taylor approximation, which better describes the laminar flow model, produced higher SNR and lower angular errors.

6.3. 4D flow. Advances in medical imaging technologies have led to new modalities such as flow sensitive magnetic resonance imaging (phase-contrast MRI) which allows the acquisition of blood flow velocities with a volumetric coverage in a time-resolved fashion, termed "4D flow MRI" or "Flow-sensitive 4D MRI" [30, 29]. The 4D flow MRI can be employed to detect and visualize temporal evolution of complex blood patterns within an acquired three-dimensional volume. The resulting flow field is an interesting object for our proposed denoising method.

We address the problem of denoising magnetic resonance imaging (MRI) vector field data. We consider a 4D flow data describing the blood flow in the carotid artery of a healthy 50 year old male volunteer. The data was acquired with a coronal slab covering the entire carotid artery. The scan was performed at a Siemens Prisma 3T clinical MRI Scanner. The sequence



(a) Original vector field.

(b) Noisy vector field.
SNR=7.506dB, AAE=4.341±3.519°.

(c) Recovery via first-order Sparsity-based over-parameterization.
SNR=14.654dB, AAE=2.128±2.539°.

(d) Recovery via second-order Sparsity-based over-parameterization.
SNR=18.406dB, AAE=1.814±1.270°.

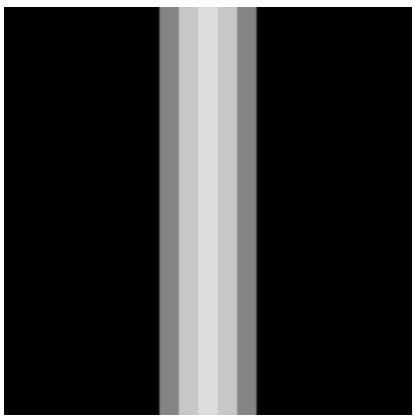
Figure 10. *Example 3 - Recovery of a quadratic polynomial vector field using different basis functions.*

was motion compensated. In order to regularize the data we assess the time point of peak carotid artery flow.

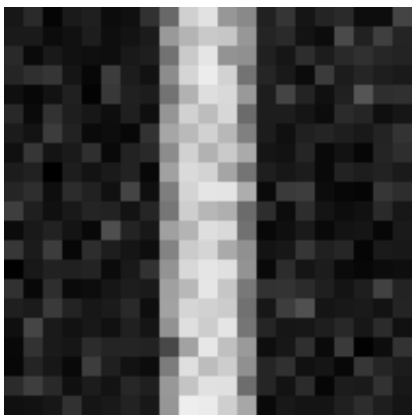
Since we wish to reconstruct several blood flow types such as plug flow and laminar flow, we use a second-order three-dimensional Taylor approximation which concludes in ten three-dimensional basis vector fields. Figures 13 and 14 present the qualitative examination of our method. The three-dimensional vector fields visualizations were generated using ParaView (Kitware Inc.).

7. Conclusions.

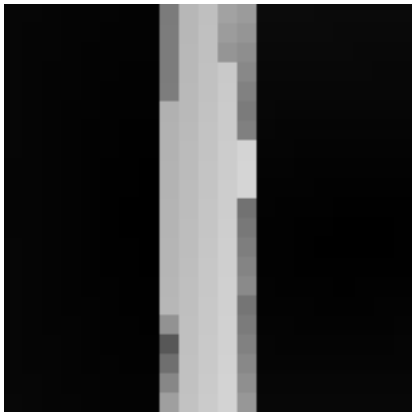
This work presents a novel framework which extends the signal denoising over-parameterization variational scheme in order to perform model-based reconstruction of vector fields. The main idea of the suggested vector field denoising method is to use basic properties of the vectors, such as magnitude and orientation, in order to effectively reconstruct the contaminated signal as a linear combination of a pre-defined basis fields set and appropri-



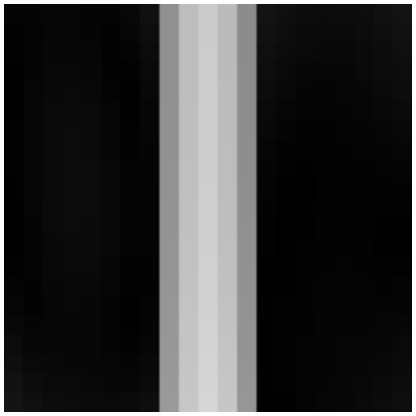
(a) Original vector field.



(b) Noisy vector field.
SNR=7.506dB, AAE=4.341±3.519°.



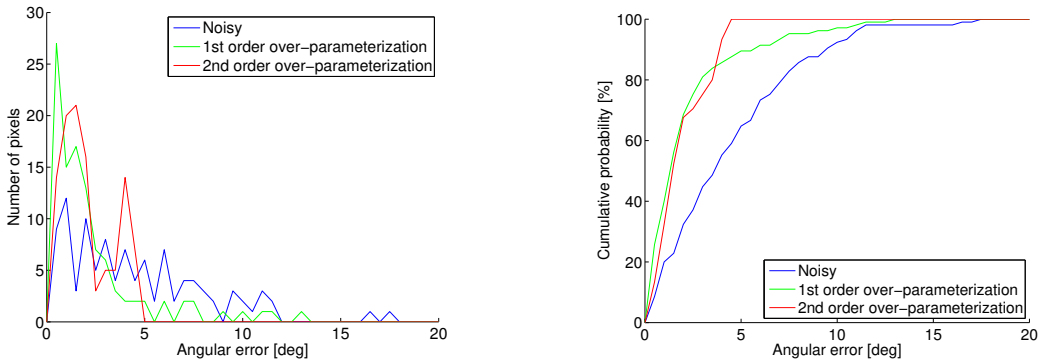
(c) Recovery via first-order Sparsity-based over-parameterization.
SNR=14.654dB, AAE=2.128±2.539°.



(d) Recovery via second-order Sparsity-based over-parameterization.
SNR=18.406dB, AAE=1.814±1.270°.

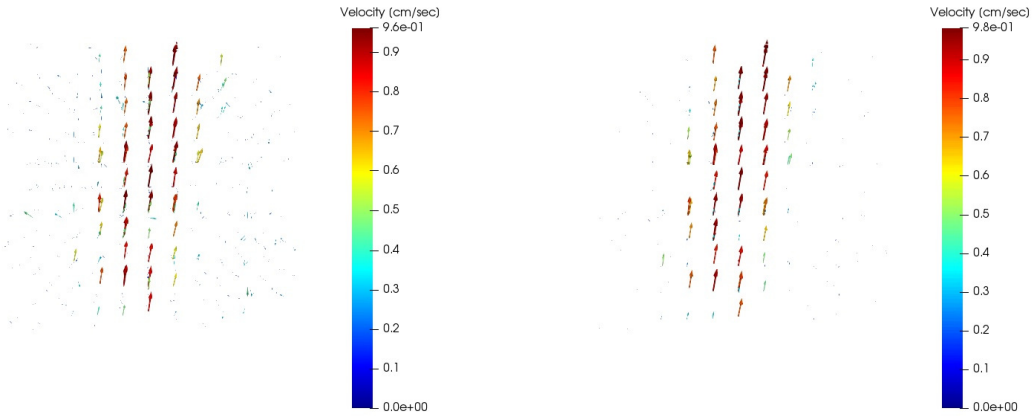
Figure 11. *Example 3 - Magnitude.*

ate coefficients. Using this representation, we utilize the over-parameterization technique to reduce the denoising problem to a simpler form of coefficient recovery. The signal estimation is done by both local fitting of the data to the selected model and global regularization of the coefficient variations. We examine a TV-based over-parameterization approach and once understating its shortcoming we propose a novel sparsity-based over-parameterization framework for vector field denoising, where the piecewise-constant coefficients are similarity-like transformations. We solve this problem by generalizing the BGAPn algorithm to support other domains, specifically the non-convex domain of similarity-like transformations. The basis field coefficients are intuitively described by matrix-valued fields and are easily regularized by reformulating the problems in terms of augmented Lagrangian functionals. Using the alternating direction method of multipliers optimization technique, we are able to separate



(a) Histogram. (b) Cumulative probability.

Figure 12. *Angular error analysis for Example 3.*



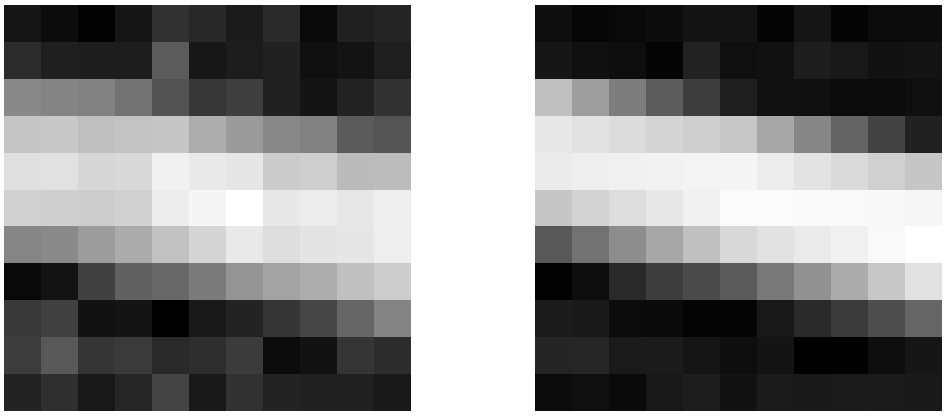
(a) Original vector field. (b) Recovery via Sparsity-based over-parameterization.

Figure 13. *Example 4 - Recovery of a real 4D flow MRI carotid artery data.*

the problem into regularization steps and projection steps, both of which can be solved in a parallel manner.

We have demonstrated the efficiency of the proposed method using a first- and second-order Taylor approximation models on two- and three-dimensional vector fields. The current sparsity-based over-parameterization framework showed good recovery results for vector fields with and without texture. The method may be further improved by averaging results produced by different sets of parameters. Another way to boost the performance of the methods is to include additional operators which take the diagonal variations into consideration. Though this work has focused on the first and second-order Taylor approximations, the extension to other models and higher orders of estimation is straightforward.

As a future work, it could be interesting to add a learning process to the scheme, tailoring the basis functions and the appropriate operators for the provided signal. The basis vector



(a) Original vector field.

(b) Recovery via Sparsity-based over-parameterization.

Figure 14. *Example 4 - Maximum intensity projections.*

fields can be used to represent the image with a new contrast, and may provide significant information, for example in radiology. In addition, further research of other applications may lead to the generalization of the proposed frameworks to other modalities and domains.

Appendix A. TV-based over-parameterization variational algorithm. In order to optimize the two variables in Equation (3.4) jointly, we use the alternating minimization technique, which obtains the general solution of the original optimization problem by alternating between minimization over each set separately, while holding the other variables fixed. In each iteration we therefore solve two minimization problems,

for $l = 1, 2, \ldots$

$$\mathbf{R}_i^l = \underset{\substack{\mathbf{R}_i:\mathbb{R}^d \to SO(n),\\ 1\le i\le m}}{\arg\min} \int_\Omega \left[\tfrac{1}{2}\left\|\mathbf{y}(\mathbf{x}) - \textstyle\sum_{i=1}^m s_i^{l-1}(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})\right\|^2 + \lambda_R \psi_R\left(\mathbf{R}\right)\right] d\mathbf{x}. \tag{A.1}$$

$$s_i^l = \underset{\substack{s_i:\mathbb{R}^d \to \mathbb{R},\\ 1\le i\le m}}{\arg\min} \int_\Omega \left[\tfrac{1}{2}\left\|\mathbf{y}(\mathbf{x}) - \textstyle\sum_{i=1}^m s_i(\mathbf{x})\mathbf{R}_i^l(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})\right\|^2 + \lambda_s \psi_s\left(\mathbf{s}\right)\right] d\mathbf{x}. \tag{A.2}$$

A.1. Optimization of the scaling coefficients. The minimization of functional (3.4) with respect to the scaling coefficients is described in problem (A.2). The objective is composed of a fitting term and a TV regularization term encouraging the coefficients to be piecewise-constant. The TV model suffers from non-linearity and non-differentiability. An efficient scheme for smoothing over the scaling coefficients can be applied by adding auxiliary fields which approximate the gradients of the data and simplify the TV regularization. According to the ADMM technique, we assign auxiliary variables $\mathbf{q}_i$, such that $\mathbf{q}_i(\mathbf{x}) = \nabla s_i(\mathbf{x}), i = 1, \ldots, m$ at each point, to form the following augmented Lagrangian problem

(A.3)

$$\min_{\substack{s_i:\mathbb{R}^d\to\mathbb{R},\\ \mathbf{q}_i:\mathbb{R}^d\to\mathbb{R}^n,\\ 1\le i\le m}} \max_{\substack{\boldsymbol{\rho}_i:\mathbb{R}^n\to\mathbb{R}^n,\\ 1\le i\le m}} L_d\left(s_1,\dots,s_m,\mathbf{q}_1,\dots,\mathbf{q}_m,\boldsymbol{\rho}_1,\dots,\boldsymbol{\rho}_m\right) =$$

$$\min_{\substack{s_i:\mathbb{R}^d\to\mathbb{R},\\ \mathbf{q}_i:\mathbb{R}^d\to\mathbb{R}^n,\\ 1\le i\le m}} \max_{\substack{\boldsymbol{\rho}_i:\mathbb{R}^d\to\mathbb{R}^n,\\ 1\le i\le m}} \int_{x\in\Omega} \begin{bmatrix} \frac{1}{2}\left\|\mathbf{y}(\mathbf{x})-\sum_{i=1}^m s_i(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})\right\|^2 \\ +\lambda_s\left\|[\mathbf{q}_1(\mathbf{x}),\mathbf{q}_2(\mathbf{x}),\dots,\mathbf{q}_m(\mathbf{x})]\right\| \\ +\frac{d}{2}\sum_{i=1}^m\left\|\mathbf{q}_i(\mathbf{x})-\nabla s_i(\mathbf{x})\right\|^2+\sum_{i=1}^m\left\langle\boldsymbol{\rho}_i(\mathbf{x}),\mathbf{q}_i(\mathbf{x})-\nabla s_i(\mathbf{x})\right\rangle \end{bmatrix} d\mathbf{x},$$

where $\boldsymbol{\rho}_i$ are the Lagrange multipliers and d is a positive constant. Hence, the method requires seeking a saddle point of functional (A.3) by using an iterative procedure.

The ADMM algorithm solves the augmented Lagrangian with respect to variables s_i and $\mathbf{q}_i$ separately and then updates the multipliers. The optimal $s_j(x)$ can be found by solving the Euler-Lagrange equation,

$$\mathbf{A}_j^T(\mathbf{x})\left(\sum_{i=1}^m \mathbf{A}_i(\mathbf{x})s_i(\mathbf{x})-\mathbf{y}(\mathbf{x})\right)+\mathrm{div}\boldsymbol{\rho}_j(\mathbf{x})+d\mathrm{div}\mathbf{q}_j(\mathbf{x})-d\Delta s_j(\mathbf{x})=0, \tag{A.4}$$

where $\mathbf{A}_i, i=1,\dots,m$ are linear operators such that $\mathbf{A}_i(\mathbf{x})s_i(\mathbf{x})=s_i(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})$.

Minimizing the augmented Lagrangian with respect to $\mathbf{q}_i$ admits a closed-form solution [49]

$$\mathbf{q}_i(\mathbf{x})=\max\left(\|\mathbf{w}_i(\mathbf{x})\|-\frac{\lambda_s}{d},0\right)\frac{\mathbf{w}_i(\mathbf{x})}{\|\mathbf{w}_i(\mathbf{x})\|},\quad \forall i=1,\dots,m, \tag{A.5}$$

where

$$\mathbf{w}_i(\mathbf{x})=\nabla s_i(\mathbf{x})-\frac{\boldsymbol{\rho}_i(\mathbf{x})}{d}.$$

Finally, $\boldsymbol{\rho}_i$ is updated according to the ADMM algorithm. An algorithmic description of the suggested scheme is presented in Algorithm A.1.

A.2. Optimization of the rotation transformations. The minimization problem with respect to the rotation transformations, which are represented by the special-orthogonal Lie group is described in problem (A.1). The data term requires that applying the rotation transformations on the basis field yield a solution that is close to the measurements, while the regularization term encourages the rotation transformations to be piecewise-constant.

An efficient scheme for smoothing maps over rotation groups includes the addition of two types of auxiliary fields, with appropriate constraints. One type approximates the data, but is forced to stay on the matrix manifold during its update, which results in turning the group constraint into a simple projection operator. Another type approximates the gradient of the data and simplifies the TV as done by known TV saddle-point solvers.

Considering the approach suggested in [40], we add auxiliary variables $\mathbf{v}_i$, such that $\mathbf{v}_i(\mathbf{x})=\mathbf{R}_i(\mathbf{x})$, $\mathbf{R}_i(\mathbf{x})\in\mathbb{R}^{n^2}$ and restrict $\mathbf{v}_i(\mathbf{x})\in SO(n)$ at each point $\forall i=1,\dots,m$. The

Algorithm A.1 ADMM algorithm for TV regularization of the scaling coefficients.

Input: noisy vector field measurements $\mathbf{y}$; constant λ_s; rotation coefficients $\{\mathbf{R}_i\}_{i=1}^m$; initial guess $\mathbf{q}_i^0, \boldsymbol{\rho}_i^0, \forall i = 1, \ldots, m$

Output: recovered scale coefficients $\{s_i\}_{i=1}^m, \forall i = 1, \ldots, m$

for $k = 1, 2, \ldots,$ until convergence **do**

Gradient update step:

$$s_i^k = \underset{\substack{s_i:\mathbb{R}^d \to \mathbb{R}, \\ 1 \le i \le m}}{\arg\min} L_d \left(s_1, \ldots, s_m, \mathbf{q}_1^{k-1}, \ldots, \mathbf{q}_m^{k-1}, \boldsymbol{\rho}_1^{k-1}, \ldots, \boldsymbol{\rho}_m^{k-1} \right), \quad \forall i = 1, \ldots, m \tag{A.6}$$

Regularization update step:

$$\mathbf{q}_i^k = \underset{\substack{\mathbf{q}_i:\mathbb{R}^d \to \mathbb{R}^n \\ 1 \le i \le m}}{\arg\min} L_d \left(s_1^k, \ldots, s_m^k, \mathbf{q}_1, \ldots, \mathbf{q}_m, \boldsymbol{\rho}_1^{k-1}, \ldots, \boldsymbol{\rho}_m^{k-1} \right), \quad \forall i = 1, \ldots, m \tag{A.7}$$

Update Lagrange multipliers:

$$\boldsymbol{\rho}_i^k(x) = \boldsymbol{\rho}^{k-1}(x) + d \left(\mathbf{q}_i^k(x) - \nabla s_i^k(x) \right), \forall i = 1, \ldots, m \tag{A.8}$$

end for

equality constraints can be enforced via augmented Lagrangian terms, which transform the minimization problem (A.1) into the following saddle-point problem:

$$\begin{aligned} &\min_{\substack{\mathbf{v}_i:\mathbb{R}^d \to SO(n), \\ \mathbf{R}_i:\mathbb{R}^d \to \mathbb{R}^{n^2}, \\ 1 \le i \le m}} \max_{\substack{\mu_i:\mathbb{R}^d \to \mathbb{R}^{n^2}, \\ 1 \le i \le m}} L_r \left(\mathbf{R}_1, \ldots, \mathbf{R}_m, \mathbf{v}_1, \ldots, \mathbf{v}_m, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \right) = \\ &\min_{\substack{\mathbf{v}_i:\mathbb{R}^d \to SO(n), \\ \mathbf{R}_i:\mathbb{R}^d \to \mathbb{R}^{n^2}, \\ 1 \le i \le m}} \max_{\substack{\mu_i:\mathbb{R}^d \to \mathbb{R}^{n^2}, \\ 1 \le i \le m}} \int_{\mathbf{x} \in \boldsymbol{\Omega}} \begin{bmatrix} \frac{1}{2} \left\| \mathbf{y}(\mathbf{x}) - \sum_{i=1}^m s_i(\mathbf{x}) \mathbf{R}_i(\mathbf{x}) \boldsymbol{\phi}_i(\mathbf{x}) \right\|^2 + \lambda_R \psi_R \left(\mathbf{R} \right) \\ + \frac{r}{2} \sum_{i=1}^m \left\| \mathbf{R}_i(\mathbf{x}) - \mathbf{v}_i(\mathbf{x}) \right\|^2 \\ + \sum_{i=1}^m \left\langle \boldsymbol{\mu}_i(x), \mathbf{R}_i(\mathbf{x}) - \mathbf{v}_i(\mathbf{x}) \right\rangle \end{bmatrix} d\mathbf{x}, \end{aligned} \tag{A.9}$$

where $\boldsymbol{\mu}_i$ are the Lagrange multipliers and r is a positive constant. This problem can be solved iteratively via the ADMM algorithm consisting of alternating minimization steps with respect to $\mathbf{R}_i$, v_i and $\boldsymbol{\mu}_i$.

Minimizing the augmented Lagrangian with respect to $\mathbf{v}_i$ yields,

$$\underset{\substack{\mathbf{v}_i:\mathbb{R}^d \to SO(n), \\ 1 \le i \le m}}{\arg\min} \int_{\mathbf{x} \in \boldsymbol{\Omega}} \left[\frac{r}{2} \sum_{i=1}^m \left\| \mathbf{R}_i(\mathbf{x}) - \mathbf{v}_i(\mathbf{x}) \right\|^2 + \sum_{i=1}^m \left\langle \boldsymbol{\mu}_i(\mathbf{x}), \mathbf{R}_i(\mathbf{x}) - \mathbf{v}_i(\mathbf{x}) \right\rangle \right] d\mathbf{x}, \tag{A.10}$$

for fixed $\mathbf{R}_i$ and $\boldsymbol{\mu}_i$. It is easy to notice that this problem can be split into m separate sub-problems, one problem for each basis field. Each sub-problem takes the form

$$\underset{\mathbf{v}_i:\mathbb{R}^d\to SO(n)}{\arg\min} \int_{\mathbf{x}\in\boldsymbol{\Omega}} \left[\frac{r}{2}\left\|\mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\right\|^2 + \langle \boldsymbol{\mu}_i(\mathbf{x}), \mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\rangle\right] d\mathbf{x}. \tag{A.11}$$

The advantage of adding auxiliary variables $\mathbf{v}_i$, is that each of these sub-problems reduces to a projection problem per point in the domain, for each basis field separately,

$$\begin{aligned} &\underset{\mathbf{v}_i:\mathbb{R}^d\to SO(n)}{\arg\min} \frac{r}{2}\left\|\mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\right\|^2 + \langle \boldsymbol{\mu}_i(\mathbf{x}), \mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\rangle \\ &= \underset{\mathbf{v}_i\in SO(n)}{\arg\min} \frac{r}{2}\left\|\mathbf{v}_i(\mathbf{x}) - \left(\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}+\mathbf{R}_i(\mathbf{x})\right)\right\|^2 = \mathrm{Proj}_{SO(n)}\left(\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}+\mathbf{R}_i\mathbf{x}\right), \end{aligned} \tag{A.12}$$

where $\mathrm{Proj}\,(\cdot)$ denotes a projector operator onto $SO(n)$ manifold. This projection finds the closest rotation matrix to the given matrix, $\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}+\mathbf{R}_i(\mathbf{x})$, where the closeness of fit is measured by the Frobenius norm. Though its not being a convex domain, the projection onto $SO(n)$ can be easily computed by means of singular value decomposition (SVD). Let $\mathbf{U}_i(\mathbf{x})\mathbf{S}_i(\mathbf{x})\mathbf{V}_i^T(\mathbf{x}) = \left(\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}+\mathbf{R}_i(\mathbf{x})\right)$ be the SVD of $\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}+\mathbf{R}_i(\mathbf{x})$, then the projection onto $SO(n)$ is

$$\mathbf{v}_i(\mathbf{x}) = \mathrm{Proj}_{SO(n)}\left(\frac{\boldsymbol{\mu}_i(\mathbf{x})}{r}-\mathbf{R}_i(\mathbf{x})\right) = \mathbf{U}_i(\mathbf{x})\mathbf{V}_i^T(\mathbf{x}). \tag{A.13}$$

An efficient optimization method for solving the augmented Lagrangian with respect to $\mathbf{R}_i$ is applied by adding auxiliary variables $\mathbf{p}_i$, such that $\mathbf{p}_i(\mathbf{x}) = \nabla\mathbf{R}_i(\mathbf{x}), \forall i = 1,\dots,m$. As in the optimization over the scaling coefficients case, the solution $\mathbf{R}_i$, is the saddle point of the augmented Lagrangian functional,

$$\begin{aligned} &\min_{\substack{\mathbf{R}_i:\mathbb{R}^d\to\mathbb{R}^{n^2},\\ \mathbf{p}_i:\mathbb{R}^d\to\mathbb{R}^{n\cdot n^2},\\ 1\le i\le m}} \max_{\substack{\nu_i:\mathbb{R}^d\to\mathbb{R}^{n\cdot n^2},\\ 1\le i\le m}} L_t\left(\mathbf{R}_1,\dots,\mathbf{R}_m,\mathbf{p}_1,\dots,\mathbf{p}_m,\boldsymbol{\nu}_1,\dots,\boldsymbol{\nu}_m\right) = \\ &\min_{\substack{\mathbf{R}_i:\mathbb{R}^d\to\mathbb{R}^{n^2},\\ p_i:\mathbb{R}^d\to\mathbb{R}^{n\cdot n^2},\\ 1\le i\le m}} \max_{\substack{\nu_i:\mathbb{R}^d\to\mathbb{R}^{n\cdot n^2}\\ 1\le i\le m}} \int_{\mathbf{x}\in\boldsymbol{\Omega}} \begin{bmatrix} \frac{1}{2}\left\|\mathbf{y}(\mathbf{x}) - \sum_{i=1}^m s_i^{k-1}(\mathbf{x})\mathbf{R}_i(\mathbf{x})\boldsymbol{\phi}_i(\mathbf{x})\right\|^2 \\ +\lambda_R\left\|[\mathbf{p}_1(\mathbf{x}),\mathbf{p}_2(\mathbf{x}),\dots,\mathbf{p}_m(\mathbf{x})]\right\| \\ +\frac{r}{2}\sum_{i=1}^m\left\|\mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\right\|^2 \\ +\sum_{i=1}^N\langle\boldsymbol{\mu}_i(\mathbf{x}),\mathbf{R}_i(\mathbf{x})-\mathbf{v}_i(\mathbf{x})\rangle \\ +\frac{t}{2}\sum_{i=1}^m\left\|\mathbf{p}_i(\mathbf{x})-\nabla\mathbf{R}_i(\mathbf{x})\right\|^2 \\ +\sum_{i=1}^m\langle\boldsymbol{\nu}_i(\mathbf{x}),\mathbf{p}_i(\mathbf{x})-\nabla\mathbf{R}_i(\mathbf{x})\rangle \end{bmatrix} d\mathbf{x}, \end{aligned} \tag{A.14}$$

where $\boldsymbol{\nu}_i, i = 1\dots,m$ are Lagrange multipliers and t is a positive constant. This sub-problem can also be solved via ADMM algorithm consisting of alternating minimization steps with respect to $\mathbf{R}_i$, $\mathbf{p}_i$ and multipliers $\boldsymbol{\nu}_i$ update. The solution scheme is summarized in Algorithm A.2.

Global convergence of the algorithm is difficult to prove due to the non-convex nature of the optimization domain. In order to ensure local convergence for this algorithm, minor modification can be made, as presented in [40].

Algorithm A.2 ADMM algorithm for rotation transformations optimization.

Input: noisy vector field measurements $\mathbf{y}$; constant λ_R; scaling coefficients $\{s_i\}_{i=1}^m$; initial guess $\mathbf{v}_i^0, \boldsymbol{\mu}_i^0, \forall i = 1, \ldots, m$

Output: recovered rotation coefficients $\{\mathbf{R}_i\}_{i=1}^m, \forall i = 1, \ldots, m$

for $k = 1, 2, \ldots,$ until convergence **do**

Regularization update step:

$$\mathbf{R}_i^k = \underset{\mathbf{R}_i:\mathbb{R}^d \to \mathbb{R}^{n^2}}{\arg\min} \; L_r\left(\mathbf{R}_1, \ldots, \mathbf{R}_m, \mathbf{v}_1^{k-1}, \ldots, \mathbf{v}_m^{k-1}, \boldsymbol{\mu}_1^{k-1}, \ldots, \boldsymbol{\mu}_m^{k-1}\right), \quad \forall i = 1, \ldots, m \tag{A.15}$$

Projection step:

$$\mathbf{v}_i^k = \underset{\mathbf{v}_i:\mathbb{R}^d \to SO(n)}{\arg\min} \; L_r\left(\mathbf{R}_1^k, \ldots, \mathbf{R}_m^k, \mathbf{v}_1, \ldots, \mathbf{v}_N, \boldsymbol{\mu}_1^{k-1}, \ldots, \boldsymbol{\mu}_m^{k-1}\right), \quad \forall i = 1, \ldots, m \tag{A.16}$$

Update Lagrange multipliers:

$$\boldsymbol{\mu}_i^k(\mathbf{x}) = \boldsymbol{\mu}_i^{k-1}(\mathbf{x}) + r\left(\mathbf{R}_i^k(\mathbf{x}) - \mathbf{v}_i^k(\mathbf{x})\right), \forall i = 1, \ldots, m \tag{A.17}$$

end for

REFERENCES

[1] L. Ambrosio and V. Tortorelli, *On the Approximation of Free Discontinuity Problems*, Preprints di matematica / Scuola Normale Superiore, Scuola Normale Superiore, 1990, https://books.google.co.il/books?id=j13JPgAACAAJ.

[2] L. Amodei and M. Benbourhim, *A vector spline approximation*, Journal of approximation theory, 67 (1991), pp. 51–79.

[3] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka-łojasiewicz inequality*, Mathematics of Operations Research, 35 (2010), pp. 438–457.

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*, SIAM, 2014.

[5] L. Bar, A. Brook, N. Sochen, and N. Kiryati, *Deblurring of color images corrupted by impulsive noise*, IEEE Transactions on Image Processing, 16 (2007), pp. 1101–1111.

[6] P. Blomgren and T. F. Chan, *Color tv: total variation methods for restoration of vector-valued images*, IEEE transactions on image processing, 7 (1998), pp. 304–309.

[7] E. Bostan, S. Lefkimmiatis, O. Vardoulis, N. Stergiopulos, and M. Unser, *Improved variational denoising of flow fields with application to phase-contrast mri data*, IEEE Signal Processing Letters, 22 (2015), pp. 762–766.

[8] E. Bostan, P. D. Tafti, and M. Unser, *A dual algorithm for l 1-regularized reconstruction of vector fields*, in Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, Ieee, 2012, pp. 1579–1582.

[9] E. Bostan, O. Vardoulis, D. Piccini, P. D. Tafti, N. Stergiopulos, and M. Unser, *Spatio-temporal regularization of flow-fields*, in Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on, IEEE, 2013, pp. 836–839.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Foundations and Trends® in Machine Learning, 3 (2011), pp. 1–122.

[11] J. R. Cebral, M. A. Castro, S. Appanaboyina, C. M. Putman, D. Millan, and A. F. Frangi, *Efficient pipeline for image-based patient-specific analysis of cerebral aneurysm hemodynamics: technique and sensitivity*, IEEE transactions on medical imaging, 24 (2005), pp. 457–467.

[12] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical imaging and vision, 20 (2004), pp. 89–97.

[13] T. Chan, S. Esedoglu, F. Park, and A. Yip, *Total variation image restoration: Overview and recent developments*, Handbook of mathematical models in computer vision, (2006), pp. 17–31.

[14] T. F. Chan, G. H. Golub, and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM journal on scientific computing, 20 (1999), pp. 1964–1977.

[15] C. Chefd'Hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, *Regularizing flows for constrained matrix-valued images*, Journal of Mathematical Imaging and Vision, 20 (2004), pp. 147–162.

[16] R. Courant, *Variational methods for the solution of problems of equilibrium and vibrations*, Bulletin of the American Mathematical Society, 49 (1943), pp. 1–23.

[17] I. Csiszár and G. Tusnády, *Information geometry and alternating minimization procedures*, Statistics and Decisions, Supplement issue, 1 (1984), pp. 205–237.

[18] F. Dodu and C. Rabut, *Irrotational or divergence-free interpolation*, Numerische Mathematik, 98 (2004), pp. 477–498.

[19] J. Eckstein and D. P. Bertsekas, *On the douglasrachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.

[20] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40.

[21] R. Giryes, M. Elad, and A. M. Bruckstein, *Sparsity based methods for overparameterized variational problems*, SIAM journal on imaging sciences, 8 (2015), pp. 2133–2159.

[22] T. Goldstein and S. Osher, *The split bregman method for l1-regularized problems*, SIAM journal on imaging sciences, 2 (2009), pp. 323–343.

[23] Y. Gur and N. Sochen, *Regularizing flows over lie groups*, Journal of Mathematical Imaging and Vision, 33 (2009), pp. 195–208.

[24] B. C. Hall, *Lie groups, Lie algebras, and representations : an elementary introduction*, Graduate texts in mathematics 222, Springer, 2nd ed. ed., 2015, http://gen.lib.rus.ec/book/index.php?md5=170209c3008e99f226bcc911c4ee7fe9.

[25] R. H. Hashemi, W. G. Bradley, and C. J. Lisanti, *MRI: The Basics*, Lippincott Williams and Wilkins, 2nd ed., 2004.

[26] M. R. Hestenes, *Multiplier and gradient methods*, Journal of optimization theory and applications, 4 (1969), pp. 303–320.

[27] B. Jung and M. Markl, *Phase-contrast mri and flow quantification*, in Magnetic Resonance Angiography: Principles and Applications, Springer New York, United States, 2012, pp. 51–64.

[28] F. R. Korosec, *Basic principles of mri and mr angiography*, in Magnetic Resonance Angiography: Principles and Applications, Springer New York, United States, 2012, pp. 3–38.

[29] M. Markl, A. Frydrychowicz, S. Kozerke, M. Hope, and O. Wieben, *4d flow mri*, Journal of Magnetic Resonance Imaging, 36 (2012), pp. 1015–1036.

[30] M. Markl, A. Harloff, T. A. Bley, M. Zaitsev, B. Jung, E. Weigang, M. Langer, J. Hennig, and A. Frydrychowicz, *Time-resolved 3d mr velocity mapping at 3t: improved navigator-gated assessment of vascular anatomy and blood flow*, Journal of magnetic resonance imaging, 25 (2007), pp. 824–831.

[31] H. A. Marquering, P. van Ooij, G. J. Streekstra, J. J. Schneiders, C. B. Majoie, A. J. Nederveen, et al., *Multiscale flow patterns within an intracranial aneurysm phantom*, IEEE Transactions on Biomedical Engineering, 58 (2011), pp. 3447–3450.

[32] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Communications on pure and applied mathematics, 42 (1989), pp. 577–685.

[33] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, *Recovery of cosparse signals with greedy analysis pursuit in the presence of noise*, in Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on, IEEE, 2011, pp. 361–364.

[34] T. Nir and A. M. Bruckstein, *On over-parameterized model based tv-denoising*, in Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on, vol. 1, IEEE, 2007, pp. 1–4.

[35] T. Nir, A. M. Bruckstein, and R. Kimmel, *Over-parameterized variational optical flow*, International Journal of Computer Vision, 76 (2008), pp. 205–216.

[36] J. Nocedal and S. Wright, *Numerical Optimization*, Springer series in operations research, Springer, 2nd ed. ed., 2006, http://gen.lib.rus.ec/book/index.php?md5=7016B74CFE6DC64C75864322EE4AA081.

[37] M. J. Powell, *A method for non linear constraints in minimization problems in optimization*, Optimization (Ed. R. Fletcher), (1969), pp. 283–298.

[38] R. T. Rockafellar, *The multiplier method of hestenes and powell applied to convex programming*, Journal of Optimization Theory and applications, 12 (1973), pp. 555–562.

[39] G. Rosman, S. Shem-Tov, D. Bitton, T. Nir, G. Adiv, R. Kimmel, A. Feuer, and A. M. Bruckstein, *Over-parameterized optical flow using a stereoscopic constraint*, in International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2011, pp. 761–772.

[40] G. Rosman, Y. Wang, X.-C. Tai, R. Kimmel, and A. M. Bruckstein, *Fast regularization of matrix-valued images*, in Efficient Algorithms for Global Optimization Methods in Computer Vision, Springer, 2014, pp. 19–43.

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

[42] G. Sapiro and D. L. Ringach, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE transactions on image processing, 5 (1996), pp. 1582–1586.

[43] J. Shah, *A common framework for curve evolution, segmentation and anisotropic diffusion*, in Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on, IEEE, 1996, pp. 136–142.

[44] S. Shem-Tov, G. Rosman, G. Adiv, R. Kimmel, and A. M. Bruckstein, *On globally optimal local modeling: from moving least squares to over-parametrization*, in Innovations for Shape Analysis, Springer, 2013, pp. 379–405.

[45] N. Sochen, R. Deriche, and L. Lopez-Perez, *Variational beltrami flows over manifolds*, in Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, vol. 1, IEEE, 2003, pp. I–861.

[46] N. Sochen, R. Kimmel, and R. Malladi, *A general framework for low level vision*, IEEE transactions on image processing, 7 (1998), pp. 310–318.

[47] P. D. Tafti, R. Delgado-Gonzalo, A. F. Stalder, and M. Unser, *Variational enhancement and denoising of flow field images*, in Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE, 2011, pp. 1061–1064.

[48] P. D. Tafti and M. Unser, *On regularized reconstruction of vector fields*, IEEE Transactions on Image Processing, 20 (2011), pp. 3163–3178.

[49] X.-C. Tai and C. Wu, *Augmented lagrangian method, dual methods and split bregman iteration for rof model*, Scale space and variational methods in computer vision, (2009), pp. 502–513.

[50] C. R. Vogel and M. E. Oman, *Iterative methods for total variation denoising*, SIAM Journal on Scientific Computing, 17 (1996), pp. 227–238.