

DIFFUSION GENERATED METHODS FOR DENOISING TARGET-VALUED IMAGES

BRAXTON OSTING* AND DONG WANG

Department of Mathematics
University of Utah
Salt Lake City, UT, USA

(Communicated by Sung Ha Kang)

ABSTRACT. We consider the inverse problem of denoising an image where each point (pixel) is an element of a target set, which we refer to as a target-valued image. The target sets considered are either (i) a closed convex set of Euclidean space or (ii) a closed subset of the sphere. The energy for the denoising problem consists of an L^2 -fidelity term which is regularized by the Dirichlet energy. A relaxation of this energy, based on the heat kernel, is introduced and the associated minimization problem is proven to be well-posed. We introduce a diffusion generated method which can be used to efficiently find minimizers of this energy. We prove results for the stability and convergence of the method for both types of target sets. The method is demonstrated on a variety of synthetic and test problems, with associated target sets given by the semi-positive definite matrices, the cube, spheres, the orthogonal matrices, and the real projective line.

1. Introduction. Let $\Omega \subset \mathbb{R}^d$ be a compact subset of Euclidean space with smooth boundary. We consider noisy *target-valued data*, $f: \Omega \rightarrow \mathbb{R}^k$, that takes values in (or near) a certain *target set*, $T \subset \mathbb{R}^k$. We assume that either

- (i) $T \subset \mathbb{R}^k$ is a closed convex set or
- (ii) $T \subset \mathbb{R}^k$ is a closed subset of the unit sphere, \mathbb{S}^{k-1} , satisfying certain additional assumptions; see Section 3.2.

Our goal will be to find a smooth map, $u: \Omega \rightarrow T$, that approximates the data, f . For $\alpha > 0$, we consider the general variational regularization of this inverse problem,

$$(1) \quad \inf_{u: \Omega \rightarrow T} E_\alpha(u), \quad \text{where } E_\alpha(u) = \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 dx + \frac{1}{2\alpha} \int_{\Omega} |u(x) - f(x)|^2 dx.$$

For a vector valued function u , the gradient in E_α should be interpreted component-wise and $|\nabla u|$ denotes the Frobenius norm of ∇u . The parameter α controls the tradeoff between the ‘smoothness of u ’ and the ‘fidelity to the data’; in the limit $\alpha \searrow 0$, we obtain $u \equiv f$ on Ω . For some target sets, T , it is difficult, either analytically or computationally, to handle the constraint that u take values in T . (This is why we didn’t specify the class of functions $u: \Omega \rightarrow T$ to take the infimum in (1).)

2010 *Mathematics Subject Classification.* 65M12, 65K10, 35K05, 49Q99.

Key words and phrases. denoising, image analysis, Merriman-Bence-Osher (MBO) diffusion generated motion, manifold-valued data, DT-MRI, line field.

The first author is supported by NSF DMS 16-19755 and DMS 17-52202.

* Corresponding author: Braxton Osting.

k	T	$L(x)$	comment	section
k	\mathbb{R}^k	0	harmonic function	
k	$T \subset \mathbb{R}^k$ convex	$\frac{1}{2} \text{dist}^2(x, T)$	convex set-valued field	§4.3
n^2	$\text{SPD}(n)$	$\frac{1}{2} \text{dist}^2(x, \text{SPD}(n))$	SPD matrix-valued field	§4.4, §4.5
1	$\{\pm 1\}$	$\frac{1}{4}(x^2 - 1)^2$	Allen-Cahn	
2	\mathbb{S}^1	$\frac{1}{4}(x ^2 - 1)^2$	Ginzburg-Landau	§4.3
k	\mathbb{S}^{k-1}	$\frac{1}{4}(x ^2 - 1)^2$	sphere-valued field	§4.1, §4.2
n^2	$O(n)$	$\frac{1}{4} \ x^t x - I_n\ _F^2$	orthogonal matrix-valued field	
k	coordinate axes, Σ_k \mathbb{RP}^1	$\frac{1}{4} \sum_{i \neq j} x_i^2 x_j^2$	Dirichlet partitions line field	§4.6, §4.7

TABLE 1. Examples of target sets, T , and penalization functions, L . We've grouped the examples by convex sets (top), subsets of the Euclidean sphere (middle), and other (bottom).

One penalization approach. Suppose, for the target set T , there exists a smooth *auxiliary function*, $L: \mathbb{R}^k \rightarrow \mathbb{R}_+$, such that $T = L^{-1}(0)$, i.e., T is the zero-level set and set of global minimizers of the non-negative function L . In this case, since $T = \arg \min_x L(x) \subset \mathbb{R}^k$, we can use the function L to *penalize* when u does not take values in T . For $\varepsilon > 0$, one may consider the relaxation of (1),

$$(2a) \quad \inf_{u \in H^1(\Omega; \mathbb{R}^k)} F_{\alpha, \varepsilon}(u),$$

where

$$(2b) \quad F_{\alpha, \varepsilon}(u) = \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 dx + \frac{1}{2\alpha} \int_{\Omega} |u(x) - f(x)|^2 dx + \frac{1}{\varepsilon^2} \int_{\Omega} L(u(x)) dx.$$

Together, the first and third terms of $F_{\alpha, \varepsilon}$ define a prior; the image is assumed to take values in T and be smooth. Clearly, for ε small, minimizing sequences must take values very near T . When there is no data present, i.e., $\alpha = \infty$, the energy in (2) simplifies to the geometric problem

$$(3a) \quad \inf_{u \in H^1(\Omega; \mathbb{R}^k)} F_{\infty, \varepsilon}(u),$$

where

$$(3b) \quad F_{\infty, \varepsilon}(u) = \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 dx + \frac{1}{\varepsilon^2} \int_{\Omega} L(u(x)) dx.$$

Energies of this general form, as well as the language “target set”, appear, e.g., in [35]. It is difficult to prove general theorems about when solutions of (2) or (3) converge to solutions of (1) (if they exist!) as $\varepsilon \searrow 0$ for general target sets, T ; generally each target set is treated on a case-by-case basis.

In what follows, we describe in more detail a few choices of target set, T , in (1), their associated auxiliary functions in (2) and (3), along with numerous applications; a summary of various choices of T is given in Table 1. Our intent is to motivate an alternative relaxation of (1), discussed below, which can be applied to all of these choices of T .

Target-valued maps and applications in imaging, inverse problems, and geometry.

Subsets of the Euclidean sphere. For $T = \{\pm 1\} = \mathbb{S}^0 \subset \mathbb{R}^1$, the auxiliary function $L(x) = \frac{1}{4}(x^2 - 1)^2$ can be used. This energy serves as a building block for a variety

of pattern formation models. In particular, from (2), we obtain

$$\varepsilon F_{\alpha,\varepsilon}(u) = \frac{\varepsilon}{2} \int_{\Omega} |\nabla u(x)|^2 dx + \frac{1}{4\varepsilon} \int_{\Omega} (u^2(x) - 1)^2 dx + \frac{\varepsilon}{2\alpha} \int_{\Omega} (u(x) - f(x))^2 dx.$$

The first two terms serve as a regularization for small ε similar to the total variation semi-norm in the Rudin-Osher-Fatemi (ROF) functional,

$$\int_{\Omega} |\nabla u| + \frac{1}{2\alpha} \int_{\Omega} (u - f)^2.$$

When used for image denoising, the total variation term serves as an ‘edge preserving regularizer’ [36].

For $T = \mathbb{S}^1 \subset \mathbb{R}^2$, the auxiliary function $L(x) = \frac{1}{4}(|x|^2 - 1)^2$ can be used. The energy in (3) corresponds to the Ginzburg-Landau equation [8] and has applications in superconductors and superfluids. In imaging, the target set $T = \mathbb{S}^1$ naturally arises when one tries to recover spatially dependent phase information, such as in Interferometric Synthetic Aperture Radar (InSAR) [34, 19]. Here, the phase difference between an interferometric pair of SAR images, obtained from slightly different camera angles, can be used to construct very accurate elevation maps. The solution of (3) with additional boundary conditions imposed can also be used to design $d = 2$ -dimensional cross fields, which have applications in, *e.g.*, computer graphics and quad mesh generation [42]. Finally, this problem is related to the simplification of vector fields for visualization [40].

For $T = \mathbb{S}^2 \subset \mathbb{R}^3$, the gradient flows of these energies are related to the heat flow of harmonic maps to \mathbb{S}^2 [12]. These equations can be obtained as simplifications of the Landau-Lifschitz equation describing non-equilibrium magnetism. This is also a simplification of the energy (the one-constant approximation) appearing in the Oseen-Frank theory for liquid crystals, where the field represents the preferred direction of molecular alignment [27, 4].

The problem in (3) with $T = O(n) \subset \{x \in \mathbb{R}^{n \times n} : \|x\|_F = 1\}$ was recently studied by the authors in [32, 47]. Here, it is natural to associate the auxiliary function $L(x) = \frac{1}{4}\|x^t x - I_n\|_F^2$, where $\|\cdot\|_F$ denotes the Frobenius norm. Since $O(1) \cong S^0$, this energy reduces to the Allen-Cahn energy when $n = 1$. Recalling that $O(n) = SO(n) \cup SO^-(n)$ and $SO(2) \cong S^1$ where $SO(n)$ (or $SO^-(n)$) denotes the group of $n \times n$ orthogonal matrix with determinate 1 (or -1). For $n = 2$, the gradient flow of this energy with initial conditions taken in $SO(2)$ reduces to the Ginzburg-Landau gradient flow [32]. This energy can be considered as a model problem for crystallography, where one considers a field that takes values in $SO(3)$ modulo the symmetry group of the crystal. This is also a model problem for the three-dimensional cross field design problem [42]. Finally, this problem is related to problems in rigid motion planning, where one tries to find a time-dependent trajectory, $u: [t_1, t_2] \rightarrow T$, where the function takes values in a set that describes admissible rigid motions, such as, *e.g.*, $T = SO(3)$ [39].

Convex sets. When the target set T is a convex set of \mathbb{R}^k , we can generally take the auxiliary function to be $L(x) = \frac{1}{2}\text{dist}^2(x, T) = \frac{1}{2} \min_{y \in T} \text{dist}^2(x, y)$. For RGB images, the image takes values in the cube, $T = [0, 1]^3 \subset \mathbb{R}^3$, as further discussed in Section 4.3.

For $T = \text{SPD}(n)$, the set of semi-positive definite (SPD) matrices, the inverse problem in (2) appears in diffusion tensor magnetic resonance imaging (MRI) [43, 24]. This application will be further discussed in Section 4.5.

Algorithm 1: A diffusion generated method for approximating minimizers of the energy in (4).

Input: Let $\tau, \lambda > 0$. Set $\Omega \in \mathbb{R}^d$, the target space as $T \in \mathbb{R}^k$, the data as $f \in L^\infty(\Omega; \mathbb{R}^k)$ and the initial guess as $u_0 = f$.

Output: A function $u \in L^\infty(\Omega; T)$ that approximately minimizes (4).
Set $s = 1$.

while not converged **do**

1. **Diffusion Step.** Extend u_{s-1} and f to $\mathbb{R}^d \setminus \Omega$ by zero. Solve the initial value problem for the free space diffusion equation until time τ :

$$\begin{aligned}\partial_t v(t, x) &= \Delta v(t, x) \\ v(0, x) &= (1 - \lambda)u_{s-1}(x) + \lambda f(x).\end{aligned}$$

Let $\tilde{u}(x) = v(\tau, x)$

2. **Projection Step.** Define $u_s \in L^\infty(\Omega; T)$ by point-wise applying Π_T to \tilde{u} ,

$$u_s(x) = \Pi_T \tilde{u}(x) \quad x \in \Omega.$$

Set $s = s + 1$.

Other target sets. For the coordinate axis, $T = \Sigma_k := \{x \in \mathbb{R}^k : \sum_{i \neq j} x_i^2 x_j^2 = 0\}$, the minimizer of (3) with an additional norm constraint gives Dirichlet partitions of Ω in the limit as $\varepsilon \rightarrow 0$; see [9, 46]. Recently, Dirichlet partitions have been used for image segmentation and data clustering [33, 52, 31].

The target set, $T = \mathbb{RP}^n$ is related to the Landau-de Gennes model, where the field describing the local orientation of a crystal is described by a Q -tensor [27, 4]. While this theory was originally used to describe nematic liquid crystals, it has also been used to describe the orientations of RNA and carbon nanotubes. Thinking of real projective space as the quotient space obtained from the n -sphere after identifying antipodal points, fields with values in $T = \mathbb{RP}^1$ are referred to as *line fields*, where a pair of antipodal directions is assigned to each point. This application will be further discussed in Section 4.7.

Further discussion of inverse problems for manifold-valued images can be found in [50, 3, 18, 20].

Results. In this paper, we derive and study an alternative relaxation of (1) than (2) based on the heat kernel. Namely, for $\tau > 0$, $\lambda \in (0, 1)$, and $f \in L^\infty(\Omega; \mathbb{R}^k)$, we consider the relaxation of (1) given by

$$(4a) \quad \inf_{u \in L^2(\Omega; T)} E_{\lambda, \tau}(u),$$

where

$$(4b) \quad E_{\lambda, \tau}(u) = -\frac{1}{2} \langle u, (e^{\Delta \tau} - I)u \rangle + \frac{\lambda}{2} \langle u - f, e^{\Delta \tau}(u - f) \rangle.$$

Here, $\langle \cdot, \cdot \rangle$, is the $L^2(\Omega; \mathbb{R}^k)$ -inner product and $e^{\Delta \tau}$ denotes the solution operator for the free space diffusion equation at time τ . If u is a vector-valued field, the diffusion operator is understood to be applied component-wise. Here, the parameter τ measures the relaxation of the problem and the parameter $\lambda = \frac{\tau}{\alpha}$ controls the data fidelity. The two terms in (4b) come from relaxing the two terms in the energy

in (1). This is explained in Section 2, together with conditions on the target set T such that (4) is well-defined, and interpretations of (4). The second term of $E_{\lambda,\tau}$ in (4b) is similar to the region-based active contour model in [25], where the idea is to use a nonlocal fidelity term to characterize the image better.

The main contribution of this paper is to derive and analyze a diffusion generated method to solve (4) for a wide class of target sets, T , including those discussed in Table 1. The proposed algorithm is given in Algorithm 1. The Algorithm consists of taking a convex combination of the previous time step and the data, diffusing until time τ , and applying a map, Π_T , point-wise to the resulting function. Here Π_T is the projection onto the target set, T , in the case that T is convex and the closest point mapping otherwise. A derivation of Algorithm 1 for the target sets considered, as well as convergence properties of the algorithm are given in Section 3.

Algorithm 1 is conceptually simple, computationally efficient, easy to implement, and applicable to a broad class of problems. Algorithm 1 can be interpreted as a splitting method for (2); see Section 3.3. However, we prefer to interpret Algorithm 1 in terms of (4) since it doesn't rely on the auxiliary function L as in (2). There are two extremes for Algorithm 1. If $\lambda = 0$, we ignore the data and find an approximate harmonic function with values in T . This is similar to the geometric problem (3) discussed above. If $\lambda = 1$, we simply dampen the highly oscillatory terms in f and apply the mapping Π_T once.

Remark 1. The argument in [23, Section 4.1] shows that the output of Algorithm 1 satisfies homogeneous Neumann boundary conditions on $\partial\Omega$. Dirichlet boundary conditions are also discussed by Laux and Yip, but we don't impose these conditions in the present work.

In Section 4, we use the proposed method to study a variety of synthetic and test numerical experiments. The target sets considered include convex sets (§4.3), $\text{SPD}(3)$ (§4.4, §4.5), \mathbb{S}^1 (§4.3), \mathbb{S}^2 (§4.1, §4.2), and \mathbb{RP}^1 (§4.6, §4.7), which have applications as described above. Generally, we find that our method gives comparable results with other manifold-valued image denoising methods, but is conceptually simpler and computationally faster.

Previous work on diffusion generated methods. The proposed method (Algorithm 1) falls into the class of diffusion generated methods (DGMs). DGMs were first introduced for $T = \{\pm 1\}$ and showed to be associated with mean curvature flow in [29, 28]. DGMs have also been generalized to generate high-order geometric motions, such as Wilmore and surface diffusion flows, in [16]. In [15], the authors used the diffusion of the distance function to generate mean curvature flow where the thresholding step was replaced by redistancing. DGMs were recently shown to be stable and generalized to multiphase mean curvature flow in [14] and applied to wetting problems in [51, 49] and foam bubble problems [44]. DGMs have been used for inverse problems for $T = \{\pm 1\}$ in [17, 45, 48]. The convergence rate of a DGM to a stationary point was proven in [32]. DGMs for $T = \mathbb{S}^1$ were introduced in [37], used for quad mesh generation in [42], and proven to be convergent in [23]. DGMs for $T = \mathbb{S}^2$ were also studied in [12]. Finally, DGMs for $T = O(n)$ was introduced and studied in [32, 47].

Outline. In Section 2, we describe properties of (4). In Section 3, we derive and study Algorithm 1 for the two types of target sets considered. In Section 4, we use

the proposed methods to study a variety of numerical experiments. We conclude in Section 5 with a discussion.

2. Properties and interpretation of the relaxed problem, (4). In this section, we motivate and derive properties of the energy, $E_{\lambda,\tau}$, given in (4), show the existence of solutions to the variational problem in (4), and give two interpretations of $E_{\lambda,\tau}$.

2.1. Motivation and properties of $E_{\lambda,\tau}$. For $u_0 \in L^1(\Omega; \mathbb{R}^k)$, we write $e^{\Delta\tau}u_0$ to denote the solution to the free space heat equation with initial condition u_0 at time τ ,

$$(e^{\Delta\tau}u_0)(x) = u(x, \tau) = (4\pi\tau)^{-d/2} \int_{\Omega} e^{-|x-y|^2/4\tau} u_0(y) dy, \quad x \in \Omega.$$

Let $\langle u, v \rangle = \int_{\Omega} u(x) \cdot v(x) dx$ denote the $L^2(\Omega; \mathbb{R}^k)$ inner product. Here \cdot is interpreted as the dot product in \mathbb{R}^k or the Frobenius inner product if u and v are matrix valued fields.

We write

$$(5) \quad \|\nabla u\|^2 = \langle u, -\Delta u \rangle = -\frac{1}{\tau} \langle u, (e^{\Delta\tau} - I)u \rangle + o(1) \text{ as } \tau \searrow 0.$$

Also, we have

$$(6) \quad \|u - f\|^2 = \langle u - f, u - f \rangle = \langle u - f, e^{\Delta\tau}(u - f) \rangle + o(1) \text{ as } \tau \searrow 0.$$

Then, using (5) and (6), we approximate the energy in the inverse problem (1) by

$$E_{\alpha}(u) = -\frac{1}{2\tau} \langle u, (e^{\Delta\tau} - I)u \rangle + \frac{1}{2\alpha} \langle u - f, e^{\Delta\tau}(u - f) \rangle + o(1) \text{ as } \tau \searrow 0.$$

Defining $\lambda = \frac{\tau}{\alpha}$, we obtain

$$E_{\alpha}(u) = \tau^{-1} E_{\lambda,\tau}(u) + o(1) \text{ as } \tau \searrow 0,$$

where $E_{\lambda,\tau}$ is given in (4b).

Lemma 2.1. *Assume $\lambda \in (0, 1)$, $\tau > 0$, and $f \in L^{\infty}(\Omega; \mathbb{R}^k)$. Then the following properties hold for the functional $E_{\lambda,\tau}$ defined in (4b).*

- (i) $E_{\lambda,\tau}$ is non-negative on $L^2(\Omega; \mathbb{R}^k)$.
- (ii) $E_{\lambda,\tau}$ is continuous with respect to the strong topology on $L^2(\Omega; \mathbb{R}^k)$.
- (iii) The Fréchet derivative of $E_{\lambda,\tau}: L^2(\Omega; \mathbb{R}^k) \rightarrow \mathbb{R}$ with respect to $u \in L^2(\Omega; \mathbb{R}^k)$ is

$$\frac{\delta E_{\lambda,\tau}(u)}{\delta u} = -(e^{\Delta\tau} - I)u + \lambda e^{\Delta\tau}(u - f).$$

- (iv) The Fréchet derivative, $\frac{\delta E_{\lambda,\tau}}{\delta u}: L^2(\Omega; \mathbb{R}^k) \rightarrow L^2(\Omega; \mathbb{R}^k)$ is Lipschitz continuous with a Lipschitz constant $L = 2 - \lambda$.

- (v) $E_{\lambda,\tau}$ is strongly convex on $L^2(\Omega; \mathbb{R}^k)$ with constant λ .

- (vi) We have the bound

$$E_{\lambda,\tau}(u) \geq \frac{\lambda}{2} (\|u\| - \|e^{\Delta\tau}f\|)^2 + \frac{\lambda}{2} \left(\left\| e^{\Delta\tau/2}f \right\|^2 - \|e^{\Delta\tau}f\|^2 \right),$$

so the sublevel sets of $E_{\lambda,\tau}$ are bounded in $L^2(\Omega; \mathbb{R}^k)$.

- (vii) $E_{\lambda,\tau}$ is sequentially weakly lower semi-continuous on $L^2(\Omega; \mathbb{R}^k)$.

A proof of Lemma 2.1 is given in Appendix A.

2.2. Existence. We define $L^p(\Omega; T)$ to be the subset of $L^p(\Omega; \mathbb{R}^k)$ consisting of maps $\Omega \rightarrow \mathbb{R}^k$ with image essentially in $T \subset \mathbb{R}^k$, i.e., functions $u \in L^p(\Omega; \mathbb{R}^k)$ such that $\text{ess.im}(u) \subset T$, where

$$\text{ess.im}(u) = \{y \in \mathbb{R}^k \mid \text{for all } \varepsilon > 0: \text{meas}(\{x \in \Omega: |u(x) - y| < \varepsilon\}) > 0\}.$$

Theorem 2.2. *Let $T \subset \mathbb{R}^k$ be a closed subset and assume $\lambda \in (0, 1)$, $\tau > 0$, and $f \in L^\infty(\Omega; \mathbb{R}^k)$. Then there exists a solution $u^* \in L^2(\Omega; T)$ that attains the infimum in (4).*

Proof. We use the direct method in the calculus of variations to establish existence. By Lemma 2.1(i), the functional is non-negative on $L^2(\Omega; T)$. Let $(u_j)_{j \in \mathbb{N}} \subset L^2(\Omega; T)$ be a minimizing sequence, i.e., $E_{\lambda, \tau}(u_j) \rightarrow E_{\lambda, \tau}^* = \inf_{u \in L^2(\Omega; T)} E_{\lambda, \tau}(u)$ as $j \rightarrow \infty$. By Lemma 2.1(vi), the minimizing sequence is bounded in $L^2(\Omega; T)$. By, e.g., [2, Theorem 2.4.3], there exists a subsequence, which we continue to denote by $(u_j)_{j \in \mathbb{N}}$ and $u^* \in L^2(\Omega; T)$, such that $u_j \rightharpoonup u^*$. By the sequential weak lower semi-continuity of $E_{\lambda, \tau}$, (see Lemma 2.1(vii)),

$$E_{\lambda, \tau}^* = \lim_{j \rightarrow \infty} E_{\lambda, \tau}(u_j) \geq E_{\lambda, \tau}(u^*),$$

which shows that u^* attains the infimum. \square

2.3. Fourier interpretation of the relaxed energy. In this section, we use the Fourier transform to describe the sense in which the minimizer in (4) achieves a balance between smoothness and fidelity to the data, f . Recall that the solution to the diffusion equation with initial data, $u(x, t) = u_0(x)$ can be expressed using the Fourier transform, $u(x, t) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-\tau|k|^2} \hat{u}_0(k) e^{ik \cdot x} dk$, where $\hat{u}_0(k) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} u_0(y) e^{-ik \cdot y} dy$. It follows that the energy in (4b) can be written

$$E_{\lambda, \tau}(u) = \int_{\mathbb{R}^d} \frac{1}{2} \left(1 - e^{-\tau|k|^2}\right) |\hat{u}(k)|^2 + \frac{\lambda}{2} e^{-\tau|k|^2} |\hat{u}_0(k) - \hat{f}(k)|^2 dk.$$

This transformation shows that any function u with small energy has the following two properties:

- the Fourier transform of u , given by $\hat{u}(k)$, should be small for $|k| \gg \tau^{-1/2}$.
- the Fourier transform of the residual, $u - f$, should be small for $|k| \ll \tau^{-1/2}$.

2.4. Perimeter interpretation of the relaxed energy. In this section, for convenience, we discuss the target set $T = \{0, 1\}$, which can be obtained by shifting and rescaling $T = \{\pm 1\}$. In [14], Esedoglu and Otto use k indicator functions $\{u_i\}_{i=1}^k$ of k domains $\{D_i\}_{i=1}^k$ to implicitly represent each domain and the interfaces, γ_{ij} , between D_i and D_j . When $\tau \ll 1$, the area of γ_{ij} can be approximated by

$$(7) \quad |\gamma_{ij}| \approx \frac{1}{\sqrt{\tau}} \int u_1 G_\tau * u_2 dx, \quad \text{where } G_\tau(x) = \frac{1}{(4\pi\tau)^{d/2}} \exp\left(-\frac{|x|^2}{4\tau}\right)$$

is the Gaussian kernel; see also [1, 30]. This is equivalent to the regularity term in (5) because $\int u_1 u_2 dx = 0$. The expression in (7) was shown to Γ -converge to the area of γ_{ij} when $\tau \searrow 0$ in [1, 30, 14]. In [14], based on this approximation, Esedoglu and Otto successfully generalized the original MBO method to a general threshold dynamics method to model the multiphase mean curvature flow by using a relaxation and linearization procedure. This procedure provides a proof of unconditional stability and consistency of the algorithm. In [21], the algorithm was

rigorously proved to converge to multiphase mean curvature flow with an angle constraint at the multiple junction when $\tau \searrow 0$. A convergence proof for $T = \mathbb{S}^1$ is given in [22].

3. Derivation and properties of the diffusion generated algorithm. In the following two subsections, we separately derive a diffusion generated method when the target set is (i) a convex set or (ii) a closed subset of the unit sphere. Both derivations lead to Algorithm 1. In Section 3.3, we give an energy splitting interpretation of Algorithm 1.

3.1. The target set is a closed convex set. When the target set, T , is closed and convex, we directly use the gradient projection algorithm (see, *e.g.*, [7]) for a fixed time step size, 1, to find the solution of (4). That is, if u_n is the solution at the n -th iteration ($n \in \mathbb{N}$), we define the $n+1$ -th iteration, u_{n+1} , by

$$u_{n+1} = \tilde{\Pi}_T \left(u_n - \frac{\delta E_{\lambda,\tau}(u_n)}{\delta u} \right),$$

where $\tilde{\Pi}_T(v)$ is the projection of $v \in L^2(\Omega; \mathbb{R}^k)$ into $L^2(\Omega; T)$, *i.e.*,

$$(8) \quad \tilde{\Pi}_T(v) := \arg \min_{u \in L^2(\Omega; T)} \|u - v\|^2.$$

Here,

$$\frac{\delta E_{\lambda,\tau}(u_n)}{\delta u} = -(e^{\Delta\tau} - I) u_n + \lambda e^{\Delta\tau} (u_n - f)$$

is the variation of $E_{\lambda,\tau}(u)$ with respect to u at $u = u_n$; see Lemma 2.1(iii). Direct calculation gives

$$u_{n+1} = \tilde{\Pi}_T(e^{\Delta\tau}((1-\lambda)u_n + \lambda f)).$$

Here, we take a convex combination of the data and current iterate, solve the diffusion equation until time τ , and project into the target set, T . Since $w := e^{\Delta\tau}((1-\lambda)u_n + \lambda f)$ is a $C^\infty(\Omega; \mathbb{R}^k)$ function, the minimization problem (8) defining $\tilde{\Pi}_T(w)$ is solved by the pointwise projection into the convex target set T , *i.e.*,

$$(9) \quad u_{n+1}(x) = \Pi_T(w(x)) \quad x \in \Omega,$$

where $\Pi_T(w) := \arg \min_{u \in T} \|u - w\|^2$ is the projection of $w \in \mathbb{R}^k$ into the convex target set T . This can be seen by writing

$$\begin{aligned} \|u - w\|^2 &= \int_{\Omega} |u(x) - w(x)|^2 dx \\ &= \int_{\Omega} |u(x) - \Pi_T(w(x)) + \Pi_T(w(x)) - w(x)|^2 dx \\ &\geq \int_{\Omega} |u(x) - \Pi_T(w(x))|^2 dx + \int_{\Omega} |\Pi_T(w(x)) - w(x)|^2 dx \\ &\geq \|\Pi_T(w) - w\|^2, \end{aligned}$$

where the first inequality follows from the first-order characterization of the projection operator, Π_T [13, Ch. II, Prop. 2.1] and the second inequality is an equality when $u(x) = \Pi_T(w(x))$ pointwise, as in (9). The algorithm is summarized in Algorithm 1.

The following theorem gives a convergence result for Algorithm 1 for T closed and convex.

Theorem 3.1. *Let T be a closed and convex set and assume $\tau > 0$, $\lambda \in (0, 1)$, and $f \in L^\infty(\Omega; \mathbb{R}^k)$. The sequence generated by (9) for any initial condition $u_0 \in L^2(\Omega; \mathbb{R}^k)$ strongly converges in $L^2(\Omega; T)$ to the unique minimum of (4), i.e., $\|u_n - u^*\| \rightarrow 0$. Furthermore, the sequence converges at a geometric rate $1 - \lambda$, i.e.,*

$$\|u_n - u^*\| \leq (1 - \lambda)^n \|u_0 - u^*\|, \quad \forall n \in \mathbb{N}.$$

Proof. By Theorem 2.2, there exists an optimal solution $u^* \in L^2(\Omega; T)$ to (4), necessarily satisfying

$$\left\langle \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*}, v - u^* \right\rangle \geq 0, \quad \forall v \in L^2(\Omega; T).$$

The uniqueness of this solution follows from the convexity of $L^2(\Omega; T)$ and strong convexity of $E_{\lambda,\tau}$; see Lemma 2.1(v). Adding and subtracting u^* , we can rewrite this as

$$\left\langle u^* - \left(u^* - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right), v - u^* \right\rangle \geq 0, \quad \forall v \in L^2(\Omega; T).$$

By [13, p.40], we conclude that

$$u^* = \Pi_T \left(u^* - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right), \quad \text{a.e. } x \in \Omega.$$

Using the definition of the method in (9), we obtain

$$\|u_{n+1} - u^*\|^2 = \left\| \Pi_T \left(u_n - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} \right) - \Pi_T \left(u^* - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right) \right\|^2.$$

By the non-expansiveness of the projection (see, e.g., [2, Prop. 3.3.2]), it follows that

$$\begin{aligned} (10) \quad & \|u_{n+1} - u^*\|^2 \\ & \leq \left\| u_n - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - u^* + \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2 \\ & = \|u_n - u^*\|^2 - 2 \left\langle u_n - u^*, \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\rangle \\ & \quad + \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2 \end{aligned}$$

Now adding the two inequalities

$$\begin{aligned} E_{\lambda,\tau}(u) & \geq E_{\lambda,\tau}(u^*) + \left\langle \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*}, u - u^* \right\rangle + \frac{L}{2} \|u - u^*\|^2 \\ E_{\lambda,\tau}(u^*) & \geq E_{\lambda,\tau}(u) + \left\langle \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u}, u^* - u \right\rangle + \frac{L}{2} \|u - u^*\|^2 \end{aligned}$$

we obtain

$$\left\langle \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*}, u - u^* \right\rangle \geq L \|u - u^*\|^2.$$

Combining this with the inequality

$$\left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\| \leq L \|u - u^*\|,$$

where L is the Lipschitz constant computed in Lemma 2.1(iv), we obtain the inequality

$$\left\langle \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*}, u - u^* \right\rangle \geq \frac{1}{L} \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2.$$

Using this inequality in (10), we obtain

$$(11) \quad \left(\frac{2}{L} - 1 \right) \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2 \leq \|u_n - u^*\|^2 - \|u_{n+1} - u^*\|^2.$$

In particular, we have that $\|u_n - u^*\| \leq \|u_0 - u^*\|$ for all n , implying that $\{\|u_n\|\}$ is bounded.

On one hand, by summing (11) from $n = 0$ to N and taking the \limsup as $N \rightarrow \infty$, we obtain

$$\begin{aligned} \left(\frac{2}{L} - 1 \right) \limsup_{N \rightarrow \infty} \sum_{n=0}^N \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2 &\leq \|u_0 - u^*\|^2 - \limsup_{N \rightarrow \infty} \|u_N - u^*\|^2 \\ &\leq \|u_0 - u^*\|^2 < \infty. \end{aligned}$$

We conclude that

$$\lim_{n \rightarrow \infty} \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} = \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*}.$$

On the other hand, from (11), we also have

$$\|u_{n+1} - u^*\|^2 \leq \|u_n - u^*\|^2 - \left(\frac{2}{L} - 1 \right) \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2.$$

By summing these relations over $n = M, \dots, N$ for arbitrary M and N with $M < N$, taking the \limsup as $N \rightarrow \infty$, and taking the \liminf as $M \rightarrow \infty$, we obtain

$$\begin{aligned} \limsup_{N \rightarrow \infty} \|u_{N+1} - u^*\|^2 &\leq \liminf_{M \rightarrow \infty} \|u_M - u^*\|^2 - \left(\frac{2}{L} - 1 \right) \liminf_{M \rightarrow \infty} \sum_{n=M}^{\infty} \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\|^2 \\ &= \liminf_{M \rightarrow \infty} \|u_M - u^*\|^2. \end{aligned}$$

Hence, we are led to the conclusion that the sequence $\{\|u_n - u^*\|\}$ is convergent. Thus, $\{u_n\}_{n \in \mathbb{N}}$ weakly converges in $L^2(\Omega; T)$ to a $\tilde{u} \in L^2(\Omega; T)$, i.e., $\lim_{n \rightarrow \infty} \langle u_n - \tilde{u}, v \rangle = 0$, $\forall v \in L^2(\Omega; T)$. To see that $\tilde{u} = u^*$, prove strong convergence, and obtain the convergence rate, we use the non-expansiveness of the projection to obtain

$$\begin{aligned} \|u_{n+1} - u^*\| &= \left\| \Pi_T \left(u_n - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} \right) - \Pi_T \left(u^* - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right) \right\| \\ &\leq \left\| u_n - \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u_n} - u^* + \frac{\delta E_{\lambda,\tau}(u)}{\delta u} \Big|_{u=u^*} \right\| \\ &= \|(1 - \lambda) e^{\Delta\tau}(u_n - u^*)\| \\ &\leq (1 - \lambda) \|u_n - u^*\|. \end{aligned}$$

The desired statement then follows. \square

3.2. The target set is a closed subset of the sphere. We consider a target set, T , satisfying the following properties:

1. T is a closed subset of the sphere, \mathbb{S}^{k-1} , i.e., $T \subseteq \{x: |x| = 1\} \subset \mathbb{R}^k$.
2. There exists a measure zero set, $\mathcal{N} \subset \mathbb{R}^k$, such that for every point in $\mathbb{R}^k \setminus \mathcal{N}$, we can define the closest point map (see, for example, [38, 26]), $\Pi_T: \mathbb{R}^k \setminus \mathcal{N} \rightarrow T$, which takes points to one of their closest points in T ,

$$\Pi_T x = \arg \min_{y \in T} |x - y|^2.$$

3. We define the closed convex set $\mathcal{B} = \text{conv}(T)$ to be the convex hull of T .

Example. For $T = \mathbb{S}^{k-1}$, we define $\mathcal{N} = \{0\}$, $\Pi_T x = \frac{x}{|x|}$, and $\mathcal{B} = \{x: |x| \leq 1\}$.

Example. For $T = O(n) \subset \mathbb{R}^{n \times n}$, \mathcal{N} is the set of singular $n \times n$ matrices, and

$$\Pi_T A = U V^t,$$

where $A = U \Sigma V^t$ is the singular value decomposition of A . We have $\mathcal{B} = \{A \in \mathbb{R}^{n \times n}: \|A\|_s \leq 1\}$ where $\|\cdot\|_s$ is the spectral norm. See [32] for further details.

Since $\Omega \subset \mathbb{R}^d$ is compact, $L^\infty(\Omega; T) \subset L^2(\Omega; T)$. If $T \subset \mathbb{R}^k$ is compact, then the following Lemma shows that the converse holds.

Lemma 3.2. *If $T \subset \mathbb{R}^k$ is compact, then $L^p(\Omega; T) \subset L^\infty(\Omega; T)$ for every $p \geq 1$.*

A proof of Lemma 3.2 is given in Appendix A.

Since $\langle u, u \rangle = 1$, (4b) can be rewritten

$$\begin{aligned} E_{\lambda, \tau}(u) &= \frac{1}{2} - \frac{1}{2} \langle u, e^{\Delta \tau} u \rangle + \frac{\lambda}{2} \langle u - f, e^{\Delta \tau} (u - f) \rangle \\ &= -\frac{1}{2}(1 - \lambda) \langle u, e^{\Delta \tau} u \rangle - \lambda \langle u, e^{\Delta \tau} f \rangle + \left(\frac{1}{2} + \frac{\lambda}{2} \langle f, e^{\Delta \tau} f \rangle \right). \end{aligned}$$

Ignoring the constant term and multiplying by -1 , the relaxed problem (4a) becomes

$$(12a) \quad \max_{u \in L^\infty(\Omega; T)} \tilde{E}_{\lambda, \tau}(u),$$

where

$$(12b) \quad \tilde{E}_{\lambda, \tau}(u) := \frac{1}{2}(1 - \lambda) \langle u, e^{\Delta \tau} u \rangle + \lambda \langle u, e^{\Delta \tau} f \rangle.$$

The existence to a solution in (12a) follows from Theorem 2.2.

The following Lemma follows from calculations similar to those in the proof of Lemma 2.1.

Lemma 3.3. *Assume $\lambda \in (0, 1)$, $\tau > 0$, and $f \in L^\infty(\Omega; \mathbb{R}^k)$. Then the following properties hold for the functional $\tilde{E}_{\lambda, \tau}$ defined in (12b).*

(i) *The variation of $\tilde{E}_{\lambda, \tau}: L^2(\Omega; \mathbb{R}^k) \rightarrow \mathbb{R}$ at u in the direction v is*

$$(13) \quad L_{\lambda, \tau}^u(v) := \left\langle v, \frac{\delta \tilde{E}_{\lambda, \tau}(u)}{\delta u} \right\rangle = \langle v, (1 - \lambda)e^{\Delta \tau} u + \lambda e^{\Delta \tau} f \rangle.$$

(ii) *$\tilde{E}_{\lambda, \tau}(u)$ is convex on $L^2(\Omega; \mathbb{R}^k)$.*

Using the fact that the maximizer of a convex functional is attained by an extremal point, we have that the maximum in (12a) is attained by an extremal point of $L^\infty(\Omega; \mathcal{B})$, i.e., $L^\infty(\Omega; T)$, as stated in the following Lemma.

Lemma 3.4. *The optimization problem in (12a) is equivalent to*

$$(14) \quad \max_{u \in L^\infty(\Omega; \mathcal{B})} \tilde{E}_{\lambda, \tau}(u).$$

The sequential linear programming approach to solving (14) is to consider a sequence of functions $\{u_n\}_{n=0}^\infty$ which satisfies

$$u_{n+1} = \arg \max_{u \in L^\infty(\Omega; \mathcal{B})} L_{\lambda, \tau}^{u_n}(u), \quad u_0 \in L^\infty(\Omega; T) \text{ is given}$$

where the linear functional $L_{\lambda, \tau}^{u_n}$ is given in (13).

Lemma 3.5. *If $e^{\Delta\tau}((1 - \lambda)u_n + \lambda f)(x) \notin \mathcal{N}$ a.e. $x \in \Omega$, then the maximizer of the linear functional $L_{\lambda, \tau}^{u_n}(u)$ over $L^\infty(\Omega; T)$ is*

$$u_{n+1} = \Pi_T(e^{\Delta\tau}((1 - \lambda)u_n + \lambda f)).$$

Proof. Writing $w = e^{\Delta\tau}((1 - \lambda)u_n + \lambda f)$, we have $L_{\lambda, \tau}^{u_n}(u) = \langle u, w \rangle$, it follows that

$$\arg \max_{u \in L^\infty(\Omega; T)} \langle u, w \rangle = \arg \min_{u \in L^2(\Omega; T)} \|u - w\|^2.$$

Since $w \in C^\infty$ and using the definition of Π_T , the result follows. \square

The iterates in Lemma 3.5 are equivalent to those in Algorithm 1. The following theorem gives the stability of Algorithm 1. The proof can be adapted directly from [32, Prop. 4.3].

Theorem 3.6. *Assume T is a closed subset of \mathbb{S}^{k-1} such that the closest point mapping, Π_T is defined on $\mathbb{R}^k \setminus \mathcal{N}$. Let $\lambda \in (0, 1)$, $\tau > 0$, and $f \in L^\infty(\Omega; \mathbb{R}^k)$. For any $n \in \mathbb{N}$, if $e^{\Delta\tau}((1 - \lambda)u_n + \lambda f)(x) \notin \mathcal{N}$ a.e. $x \in \Omega$, then the functional $E_{\lambda, \tau}$, defined in (4a), is non-increasing, i.e., $E_{\lambda, \tau}(u_{n+1}) \leq E_{\lambda, \tau}(u_n)$.*

Remark 2. Regarding the assumption in Theorem 3.6, in all numerical experiments conducted (see Section 4), we did not observe that $e^{\Delta\tau}((1 - \lambda)u_n + \lambda f)(x) \in \mathcal{N}$ for any $x \in \Omega$ or $n \in \mathbb{N}$. If this condition did occur, a random closest point could be assigned by Π_T .

Remark 3. For the special case that $T = \mathbb{S}^0 = \{\pm 1\}$, a proof of convergence for a discrete version (i.e., Ω is discretized) of Algorithm 1 can be adapted from [32, Prop. 4.4].

3.3. Energy splitting interpretation of Algorithm 1. In this section, we interpret Algorithm 1 as a splitting method for (2).

Define the proximal operator for the functional $E: L^2(\Omega; \mathbb{R}^k) \rightarrow \mathbb{R}$, $\mathbf{Prox}_E: L^2(\Omega; \mathbb{R}^k) \rightarrow L^2(\Omega; \mathbb{R}^k)$, by

$$\mathbf{Prox}_E(v) := \arg \min_{u \in L^2(\Omega; \mathbb{R}^k)} E(u) + \frac{1}{2} \|u - v\|^2,$$

and, for $\beta > 0$, similarly define the proximal operator of the scaled functional $\beta E: L^2(\Omega; \mathbb{R}^k) \rightarrow \mathbb{R}$, $\mathbf{Prox}_{\beta E}: L^2(\Omega; \mathbb{R}^k) \rightarrow L^2(\Omega; \mathbb{R}^k)$, by

$$\mathbf{Prox}_{\beta E}(v) := \arg \min_{u \in L^2(\Omega; \mathbb{R}^k)} E(u) + \frac{1}{2\beta} \|u - v\|^2.$$

To simplify notation, we rewrite (2) by

$$(15) \quad \inf_{u \in H^1(\Omega; \mathbb{R}^k)} \mathcal{E}_1(u) + \mathcal{E}_2(u) + \mathcal{E}_3(u)$$

where $\mathcal{E}_1(u) = \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 dx$, $\mathcal{E}_2(u) = \frac{1}{2\alpha} \int_{\Omega} |u(x) - f(x)|^2 dx$, and $\mathcal{E}_3(u) = \frac{1}{\varepsilon^2} \int_{\Omega} L(u(x)) dx$.

Now, we introduce an iterative method by formally splitting the minimization in (15) as follows. Let u_n be the solution at the n -th iteration of Algorithm 1. We first consider \mathcal{E}_2 and solve

$$(16) \quad u_n^* = \mathbf{Prox}_{\tau \mathcal{E}_2}(u_n).$$

We then evolve u_n^* by the gradient flow of \mathcal{E}_1 until time $t = \tau$,

$$(17) \quad \begin{cases} v_t = \Delta v, \\ v|_{t=0} = u_n^* \end{cases}$$

to obtain $u_n^{**} = v(\tau)$. Finally, we consider \mathcal{E}_3 and set

$$(18) \quad u_{n+1} = \mathbf{Prox}_{\tau \mathcal{E}_3}(u_n^{**}).$$

We can directly solve (16), by setting the first variation to zero,

$$\frac{\delta(\mathcal{E}_2(u) + \frac{1}{2\tau} \|u - u_n\|^2)}{\delta u} = \frac{1}{\alpha}(u - f) + \frac{1}{\tau}(u - u_n) = 0,$$

and solving for u to obtain

$$u_n^* = (1 - \tilde{\lambda})u_n + \tilde{\lambda}f, \quad \text{where } \tilde{\lambda} = \frac{\tau}{\alpha + \tau} = \frac{\lambda}{1 + \lambda}.$$

This together with (17) gives the diffusion step in Algorithm 1 for a modified convex combination parameter.

As $\varepsilon \searrow 0$, the solution to (18) is obtained by taking a function $u \in L^2(\Omega; T)$ such that

$$u = \arg \min_{u \in L^2(\Omega; T)} \|u - u_n^{**}\|^2,$$

which is precisely the point-wise projection,

$$u_{n+1} = \Pi_T(u_n^{**}).$$

This is the projection step in Algorithm 1.

4. Computational examples. In this Section, we demonstrate the diffusion generated method in Algorithm 1, developed in Section 3, on several synthetic and test numerical experiments. Several of the numerical experiments considered are from [3], which can provide a comparison. For these examples, we both cite the section number and include details of the examples for completeness.

The data in Sections 4.1 and 4.2 is periodic, so we solve the heat diffusion equation in Algorithm 1 with a periodic boundary condition instead of the free-space heat diffusion equation. For all other examples, we solve the free-space heat diffusion equation. All methods were implemented in MATLAB and results reported below were obtained on a laptop with a 2.7GHz Intel Core i5 processor and 8GB of RAM.

4.1. Example: a synthetic \mathbb{S}^2 -valued one-dimensional image. Following [3, §5.1], we consider the lemniscate of Bernoulli, given by

$$\gamma(t) = \frac{\pi/2}{\sin^2(t) + 1} (\cos(t), \cos(1)\sin(t)), \quad t \in [0, 2\pi].$$

For $p \in \mathbb{S}^2$ and $\xi \in T_p\mathbb{S}^2$, let $\Gamma_{p,\xi}$ be the unique geodesic starting from $\Gamma_{p,\xi}(0) = p$ with $\dot{\Gamma}_{p,\xi}(0) = \xi$. Then we define $\exp_p: T_p\mathbb{S}^2 \rightarrow \mathbb{S}^2$ by $\exp_p(\xi) = \Gamma_{p,\xi}(1)$. Then, one spherical lemniscate curve can be obtained by

$$\gamma_S(t) = \exp_p(\gamma(t))$$

with $p = (0, 0, 1)$. We discretize the curve in the parameter space of t at $t_i := \frac{2\pi i}{511}, i = 0, \dots, 511$ to get an \mathbb{S}^2 -valued signal, which we denote by $\{f_{o,i}\}_{i=0}^{511}$. We add noise to the data by taking $f_i = \exp_{f_{o,i}}(\eta_i)$ with $\eta_i = (\eta_{i1}, \eta_{i2})$, where η_{i1} and η_{i2} are independent Gaussian noises with standard deviation of $\frac{\pi}{30}$. In Figure 1(a), the blue markers indicate the original data and the red markers indicate the noisy data.

In this example, we take the target set, T , to be $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$. Then, we have $\mathcal{N} = \{x \in \mathbb{R}^2 : x = 0\}$ and $\Pi_T(x) = \frac{x}{|x|}$ if $x \neq 0$. Applying Algorithm 1, we get the denoised results shown in Figures 1(b)–1(d) with a fixed $\tau = 10^{-3}$ and $\lambda = 0.05, 0.1$, and 0.15 , respectively. Since the original image is periodic, we solve the diffusion equation in Algorithm 1 with periodic boundary conditions.

We use the Hausdorff distance between the denoised curve and the original curve as a quantity to measure the quality of our results. In Figures 1(b)–1(d), the Hausdorff distances are 0.0721, 0.0501 and 0.0504, which indicates that the denoised results very closely match the original data and that the results are relatively insensitive to the parameter λ . All of these simulations were completed within 0.01 seconds.

4.2. Example: a synthetic \mathbb{S}^2 -valued image. Again, following [3, §5.1], we define an \mathbb{S}^2 -valued vector-field by

$$G(t, s) = R_{x(t)+y(s)} S_{x(t)-y(s)} e_3, \quad t \in [0, 8\pi], \quad s \in [0, 8\pi]$$

where

$$x(t) = t + \frac{\pi}{2} \left\lfloor \frac{t}{2\pi} \right\rfloor, \quad y(t) = t + \frac{\pi}{2} \left\lfloor \frac{t}{2\pi} \right\rfloor,$$

$$R_\theta := \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad S_\theta := \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}.$$

We sample both dimensions, t and s , with $n = 64$ points to obtain a discrete vector field $f_o \subset \mathbb{S}^2$ which is given in Figure 2(a). Similar to the noise in 4.1, we add the Gaussian noise in the tangential plane of each point with standard deviation $\frac{4\pi}{45}$; the noisy data is displayed in Figure 2(b).

In this example, we take the target set, T , to be $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$. Then, we have $\mathcal{N} = \{x \in \mathbb{R}^3 : x = 0\}$ and $\Pi_T(x) = \frac{x}{|x|}$ if $x \neq 0$. Since the original image is periodic, we solve the diffusion equation in Algorithm 1 with periodic boundary conditions.

In this example, we use the L^2 error between the denoised vector field and the original vector field to measure the quality of the denoised results. The results of

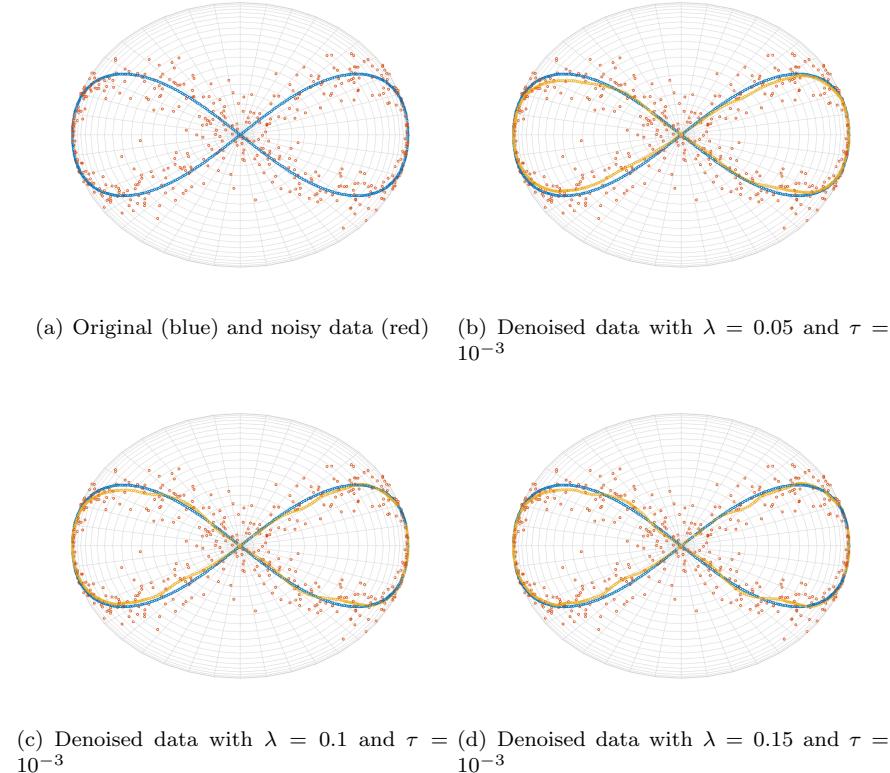


FIGURE 1. Results of denoising an obstructed lemniscate of Bernoulli on the sphere, S^2 , with $\lambda = 0.05, 0.1$ and 0.15 , respectively. The Hausdorff distances from the denoised curve to the original curve are $0.721, 0.501$ and 0.504 , respectively. In this simulation, τ is fixed as 10^{-3} . See Section 4.1.

Algorithm 1 with $\tau = 10^{-3}$ and $\lambda = 0.05, 0.1, 0.15$, and 0.2 are displayed in Figure 2(c)–(f) and the L^2 errors are $0.1544, 0.1058, 0.0794$, and 0.0668 , respectively. The numerical results show that Algorithm 1 is relatively insensitive to the value of the parameter λ and it is applicable to recover the original data image in Figure 2(a). It is also observed that the denoised image is slightly smoothed and the L^2 error increases when decreasing the value of λ . Again, the algorithm performs very efficiently; all simulations can be done within 0.2 seconds.

To further show the potential of Algorithm 1, we apply the algorithm to a discontinuous image with large jumps. We define an S^2 -valued vector-field by

$$G(t, s) = \mathbf{1}_D \vec{v}_1 + (1 - \mathbf{1}_D) \vec{v}_2, \quad t \in [0, 2\pi], s \in [0, 2\pi]$$

where D is the white region as in Figure 3(a), $\mathbf{1}_D$ is the indicator function of the white region, $\vec{v}_1 = \frac{1}{\sqrt{2}}[\frac{\sqrt{3}}{2}, \frac{1}{2}, 1]$, and $\vec{v}_2 = \frac{1}{\sqrt{2}}[-\frac{\sqrt{3}}{2}, \frac{1}{2}, 1]$. Again, we sample both dimensions, t and s , with $n = 64$ points to obtain a discrete vector field $f_o \subset S^2$ which is given in Figure 3(b) and add the Gaussian noise in the tangential plane of each point with standard deviation $\frac{4\pi}{45}$ to obtain the noisy data which is displayed in

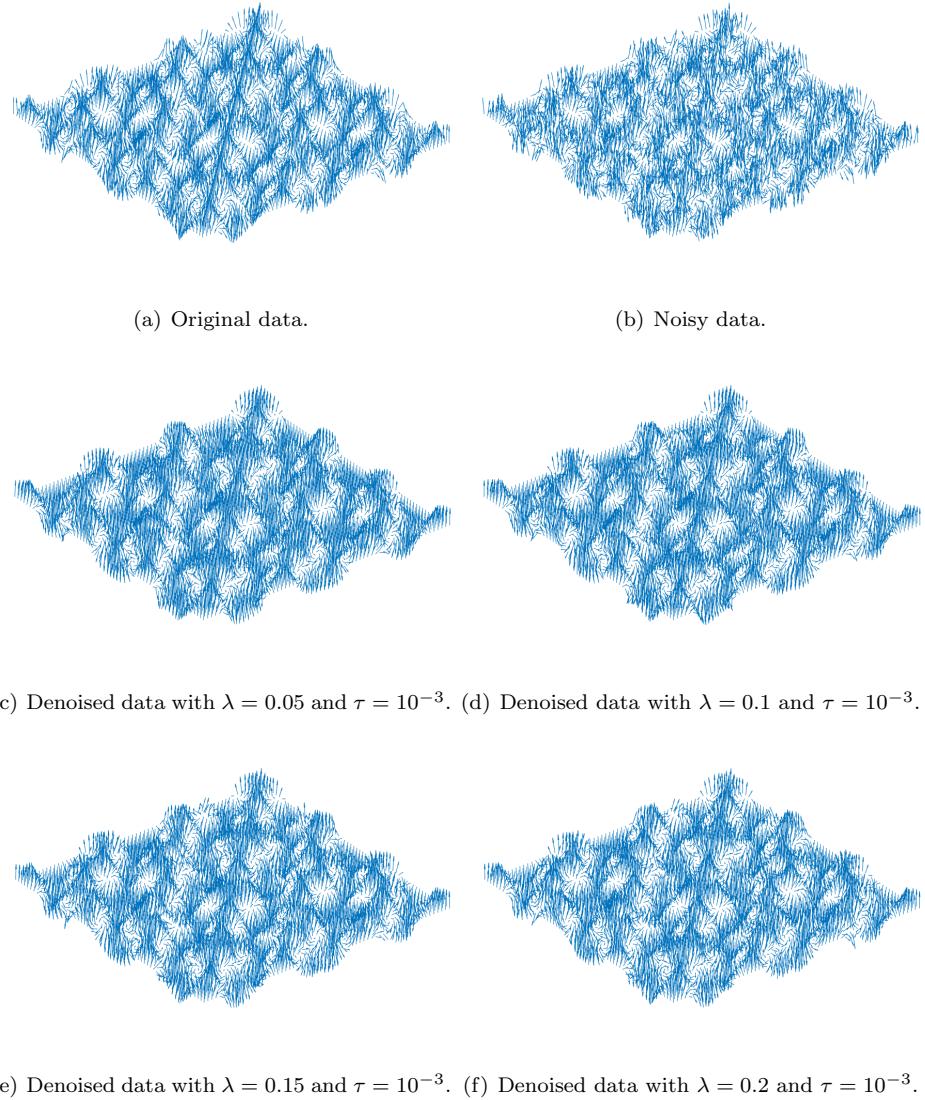


FIGURE 2. Results of the reconstruction of a noisy \mathbb{S}^2 -valued image with $\lambda = 0.05, 0.1, 0.15$, and 0.2 , respectively. In all the simulations, τ is fixed as 10^{-3} . See Section 4.2.

Figure 3(c). The reconstructed image of Algorithm 1 is displayed in Figure 3(f). The results from two TV-type methods are displayed in Figure 3(d,e). They are obtained by using the code from Manifold-Valued Image Restoration Toolbox (MVIRT) which is based on [3]. We tried different parameters (*i.e.*, α and β in [3]) and Figure 3(d,e) are the best results obtained. We observe that our result is also competitive to the results from TV-type methods. In particular, the discontinuity is also well resolved even if it is not clear in the noisy data.

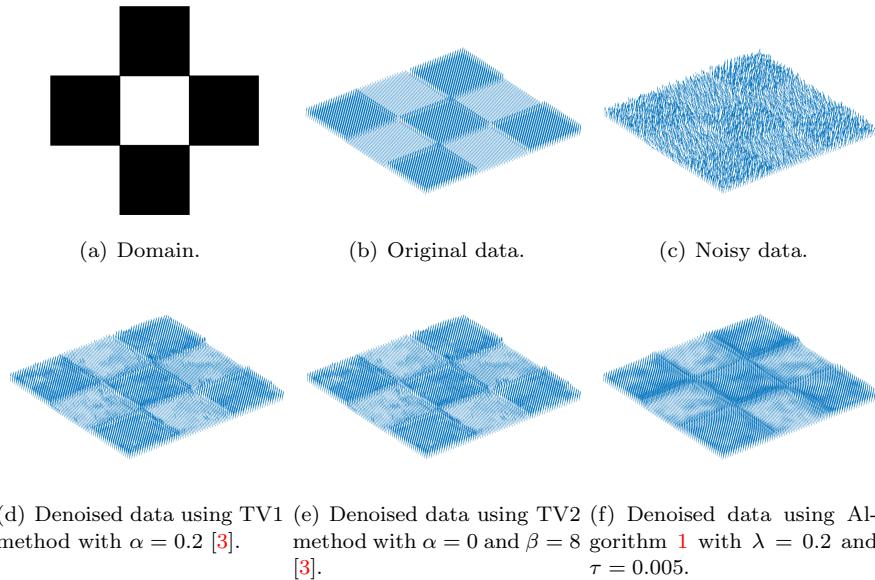


FIGURE 3. Results of the reconstruction of a noisy \mathbb{S}^2 -valued image with different methods. See Section 4.2.

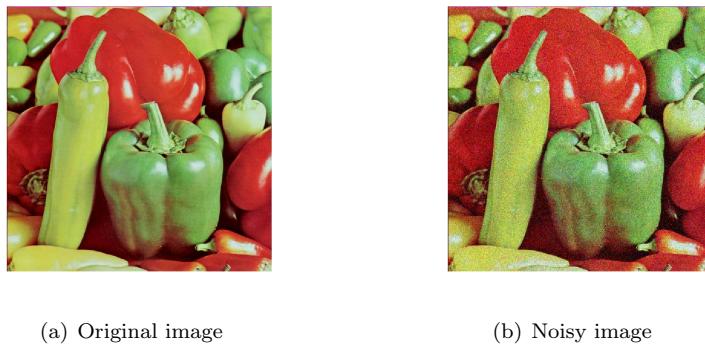


FIGURE 4. The original ‘peppers’ image and one corrupted with noise. See Section 4.3.

4.3. Example: the ‘peppers’ image. Following [3, §5.1], we consider the denoising of the ‘peppers’ image, shown in Figure 4(a). It is distorted with Gaussian noise in each of the red, green, and blue (RGB) channels with standard deviation 0.1, as shown in Figure 4(b).

We first consider the image represented in RGB channels. Here, the target set is the unit cube, $T = [0, 1]^3 \subset \mathbb{R}^3$. Since this is a convex set, we let Π_T be the projection and apply Algorithm 1 to denoise Figure 4(b). Figure 5 displays the denoised result for different values of the parameter λ .

We next consider the image represented in hue, saturation, and value (HSV) channels. Here, the color space, $(H, S, V) \in S^1 \times [0, 1] \times [0, 1]$, consists of a S^1 -valued hue component, H , and two $[0, 1]$ -valued components: saturation, S , and value, V . For this product target space, T , we define the mapping Π_T to map onto each component individually. The results of applying Algorithm 1 to denoising Figure 4(b) in the HSV channels are displayed in Figure 6 for different values of the parameter λ .

We use a relatively large value of λ compared to other numerical experiments since the original image is non-smooth. We observe that the denoised images are slightly blurred, but the results are robust to changes in the parameter λ . We use the peak signal-to-noise ratio (PSNR) to evaluate the quality of the denoised image. The PSNR for the RGB channel is computed by Matlab's `psnr` function and the PSNR for the HSV channel is the PSNR for the corresponding RGB channel. The PSNR obtained in [3] for both RGB and HSV channels ranges between 28.16 and 31.24, which are slightly better than the values obtained using this much simpler method.



(a) $\lambda = 0.85, \tau = 10^{-4}$, (b) $\lambda = 0.9, \tau = 10^{-4}$, PSNR= 28.3754. (c) $\lambda = 0.95, \tau = 10^{-4}$, PSNR= 28.4118.

FIGURE 5. Denoising in the RGB channel on the noisy ‘peppers’ image (see Figure 4(b)) with $\tau = 10^{-4}$ and $\lambda = 0.85, 0.9$, and 0.95. The PSNR listed below each image indicates the quality of the result. See Section 4.3.

4.4. Example: an SPD(3)-valued image. Following [3, §5.2], we consider the reconstruction of a synthetic symmetric positive definite (SPD(3)) matrix-valued image. Here, we consider the positive semidefinite matrices as a cone in Euclidean space.

An SPD(3)-valued image is constructed by sampling

$$G(s, t) := A(s, t) \operatorname{diag}[v(s, t)] A(s, t)^t, \quad \text{where } s, t \in [0, 1].$$

Here, we define

$$A(s, t) = R_{x_2, x_3}(\pi s) R_{x_1, x_2}(|2\pi s - \pi|) R_{x_1, x_2}(|\pi(t - s - |t - s|) - \pi|),$$

$$v(s, t) = \left(1 + \delta_{x+y, 1}, 1 + s + t + \frac{3}{2} \delta_{s, \frac{1}{2}}, 4 - s - t + \frac{3}{2} \delta_{t, \frac{1}{2}} \right)^T$$



(a) $\lambda = 0.85$, $\tau = 10^{-4}$, (b) $\lambda = 0.9$, $\tau = 10^{-4}$, PSNR= 26.4092. (c) $\lambda = 0.95$, $\tau = 10^{-4}$, PSNR= 26.4179.

FIGURE 6. Denoising in the HSV channel on the noisy ‘peppers’ image (see Figure 4(b)) with $\tau = 10^{-4}$ and $\lambda = 0.85, 0.9$, and 0.95 . The PSNR are listed below each image to indicate the quality of the result. See Section 4.3.

where $R_{x_i, x_j}(t)$ is the rotation matrix in the x_i, x_j -plane with angle t and $\delta_{a,b} = \begin{cases} 1 & \text{if } a > b \\ 0 & \text{else} \end{cases}$.

We discretize the parameter space (s, t) with 25×25 grid points to obtain a 25×25 matrix-valued image $f = \{f_{i,j}\}_{i,j=1}^{25} \subset \text{SPD}(3)$. The SPD matrix is visualized in Figure 7(a) by the corresponding ellipsoid at each pixel location. The noisy data in Figure 7(b) is generated by adding noise as follows. We first compute the Cholesky factorization of f , $f = R^T R$ and then define the noisy data, \tilde{f} , by $\tilde{f} = A A^T$, where $A = R + \sqrt{0.03}B$ and B is a 3×3 upper triangular matrix with each element being a random number from the standard normal distribution. Since f is symmetric positive definite, R is well defined in the Cholesky factorization. We note that adding noise in such way implies that \tilde{f} is symmetric positive definite.

In this example, we take the target set to be the group of 3×3 symmetric positive definite matrices, $\text{SPD}(3)$, and solve the free space heat diffusion equation in Algorithm 1. Then, we use the mapping

$$\Pi_T(A) = U \Sigma_+ V^T$$

where $U \Sigma V^T$ is the singular value decomposition of A and $\Sigma_+(i, j) = \max(\Sigma(i, j), 0)$.

Figure 8 displays the results of reconstruction with $\tau = 10^{-3}$ and $\lambda = 0.05, 0.1$, and 0.15 , respectively. The denoised images are very close to the original data. Algorithm 1 is relatively insensitive to the value of the parameter λ . Figure 8(a) is slightly smoothed when λ is relatively small and the regularity term dominates.

4.5. Example: DT-MRI. Following [3, §5.2], we attempt to remove naturally occurring noise from the image. We apply Algorithm 1 to a dataset from the Camino project [11] of a diffusion tensor magnetic resonance image (DT-MRI) of the human head. From the complete data set of $f: \Omega \rightarrow \{\mathcal{P}(3)\}^{112 \times 112 \times 50}$ where $\Omega = [112] \times [112] \times [50]$, we take the 28-th traversal plane, $\Omega_0 = [112] \times [112] \times \{28\} \subset \Omega$ as the dataset for reconstruction (See Figure 9(a)) and zoom-in on the subset $\Omega_1 = \{28, \dots, 87\} \times \{24, \dots, 73\} \times \{28\} \subset \Omega_0$ in Figure 9(b).

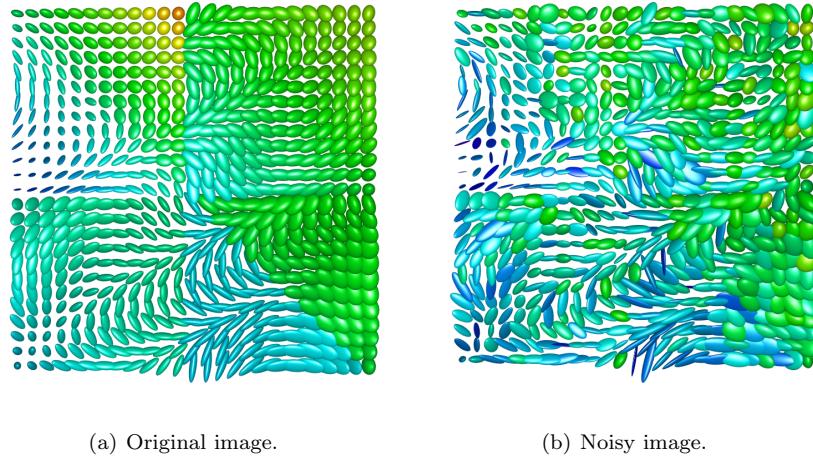


FIGURE 7. The original and noisy SPD-valued image. See Section 4.4.

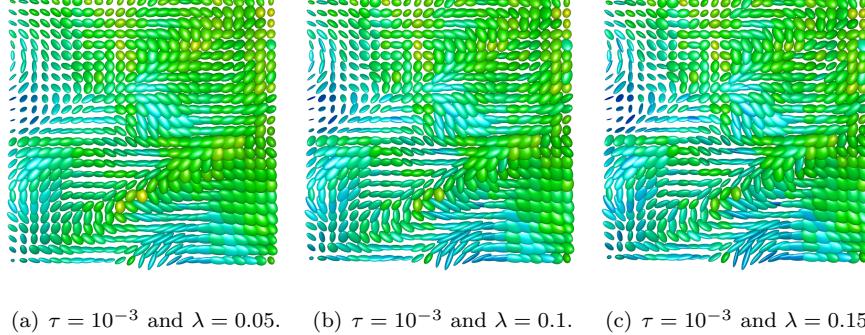
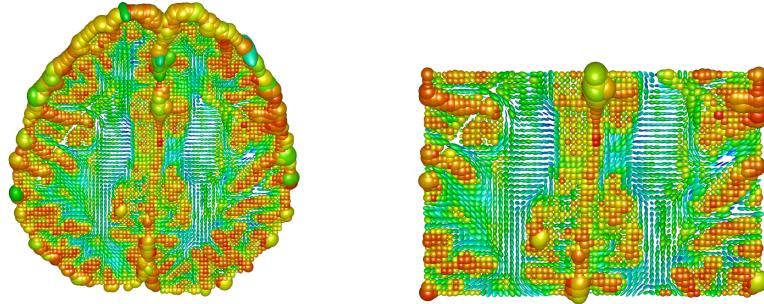


FIGURE 8. The result of reconstruction in Figure 7(b) with $\tau = 10^{-3}$ and $\lambda = 0.05, 0.1$, and 0.15 . See Section 4.4.

As in the example in Section 4.4, we take the target set to be the group of 3×3 symmetric positive definite matrices, $\text{SPD}(3)$, and solve the free space heat diffusion equation in Algorithm 1. Figure 10 displays the reconstructions obtained using Algorithm 1 of the data in Figure 9 with $\tau = 10^{-4}$ and $\lambda = 0.1, 0.2$, and 0.3 . The first row displays the reconstructed data in Ω_0 and the second row displays the subset Ω_1 of the corresponding reconstructed data. The results are very similar to those in [3].

4.6. Example: an \mathbb{RP}^1 -valued image. In this section, we denoise a real projective line (\mathbb{RP}^1)-valued image. It is not clear that real projective spaces, \mathbb{RP}^n , for $n > 1$ can be cast within the framework of the proposed methods. However, due to the topological equivalence of \mathbb{RP}^1 with the circle, \mathbb{S}^1 , we can study real projection line-valued images as follows.

We recall that \mathbb{RP}^{n-1} can be viewed as identifying antipodal points of the unit n -sphere, $\mathbb{S}^{n-1} \subset \mathbb{R}^k$. Thus, identifying \mathbb{R}^2 with \mathbb{C} , we can uniquely represent each



(a) Slice 28 of the original Camino DT-MRI data, Ω_0 .
(b) Subset $\Omega_1 \subset \Omega_0$.

FIGURE 9. Slice 28 of the original Camino DT-MRI data and the subset Ω_1 . See Section 4.5.

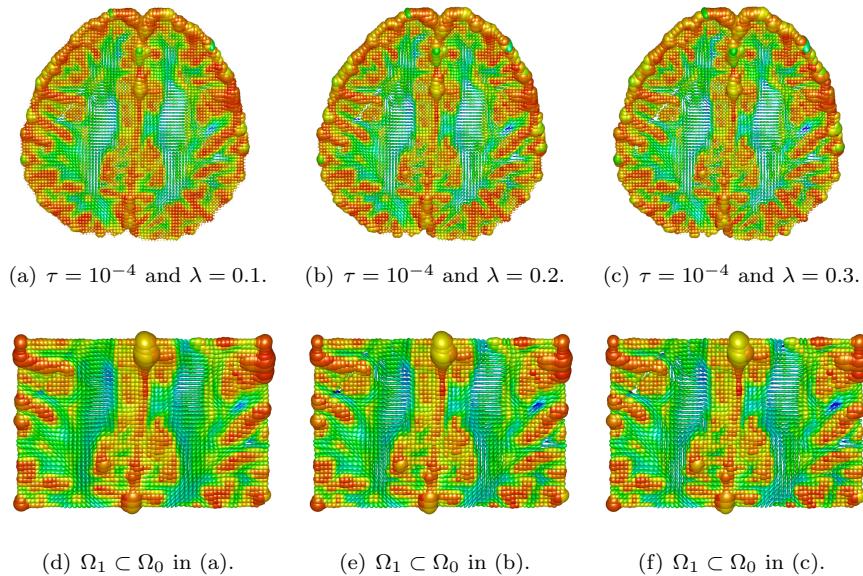


FIGURE 10. Results of the reconstruction in Figure 9(a) with $\tau = 10^{-4}$ and $\lambda = 0.1, 0.2$, and 0.3 . See Section 4.5.

element $\{x, -x\} \in \mathbb{RP}^1$, by its squared value, $x^2 \in \mathbb{S}^1 \subset \mathbb{C}$. We then denoise this representation map using Algorithm 1 and the mapping $\Pi_T = \frac{x}{|x|}$. Finally, by taking the two square roots of the denoised representation map, we obtain a denoised \mathbb{RP}^1 -valued image. A similar strategy is employed in [42] for cross-valued fields.

We define a synthetic \mathbb{RP}^1 -valued image as follows. First define

$$z(x, y) = \left(\Re(\sqrt{ix - y}), \Im(\sqrt{ix - y}) \right)$$

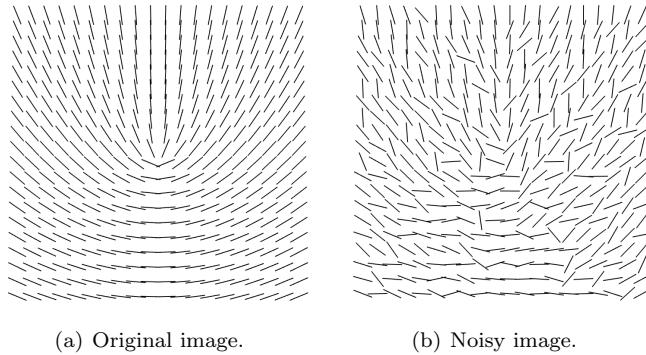


FIGURE 11. The original and noisy synthetic line-field image. See Section 4.6.

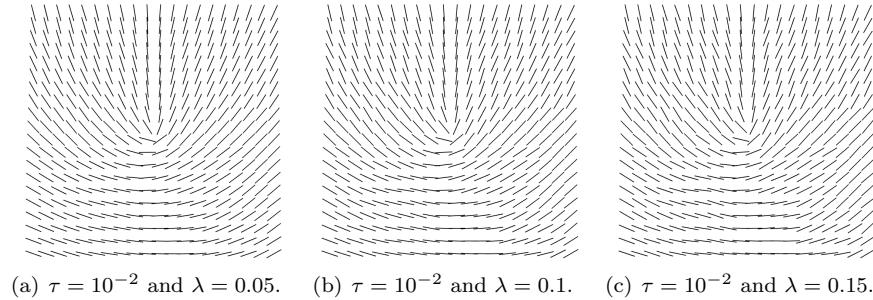


FIGURE 12. Denoising results for the line-field in Figure 11(b) with $\tau = 10^{-2}$ and $\lambda = 0.05, 0.1$, and 0.15 . See Section 4.6.

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts. Now we define the line field,

$$f(x, y) = \left\{ \frac{[z(x, y)]_+}{|z(x, y)|}, \frac{[z(x, y)]_-}{|z(x, y)|} \right\} \in \mathbb{RP}^1, \quad \text{where } x, y \in [-1, 1].$$

Here, $[z]_+ = (|z_1|, z_2) \in S^1 \cap \{z: z_1 > 0\}$ and $[z]_- = -[z]_+$. We discretize the parameter space (x, y) by 20×20 grid points to give a discretized line field, $\{f_{i,j}\}_{i,j=1}^{20} \subset \mathbb{RP}^1$, which is displayed in Figure 11(a). We then add Gaussian noise with standard deviation 0.3 to (x_i, y_i) point- and component-wise to have $(\tilde{x}_i, \tilde{y}_i)$. Noisy data is then generated by $\tilde{f}_{i,j} = f(\tilde{x}_i, \tilde{y}_j)$ and plotted in Figure 11(b).

Proceeding as described above, we view the noisy line field as taking values in \mathbb{C} and point-wise square the values to obtain the representation field, $\tilde{f}^2 \subset \mathbb{S}^1$. We apply Algorithm 1 to \tilde{f}^2 with target set, $T = \mathbb{S}^1$. We then take the point-wise square root of the resulting field to obtain the denoised line field. The results are displayed in Figure 12 with $\tau = 10^{-2}$ and $\lambda = 0.05, 0.1$, and 0.15 . We observe that the reconstructed images are very close to the original data and the index $+1/2$ singularity is well-preserved.

4.7. Example: a fingerprint image. In this section, using the methods described in Section 4.6 for \mathbb{RP}^1 , we analyze an image of a fingerprint; see Figures 13(a) and (d). From the fingerprint, based on the image intensity of the fingerprint image, we

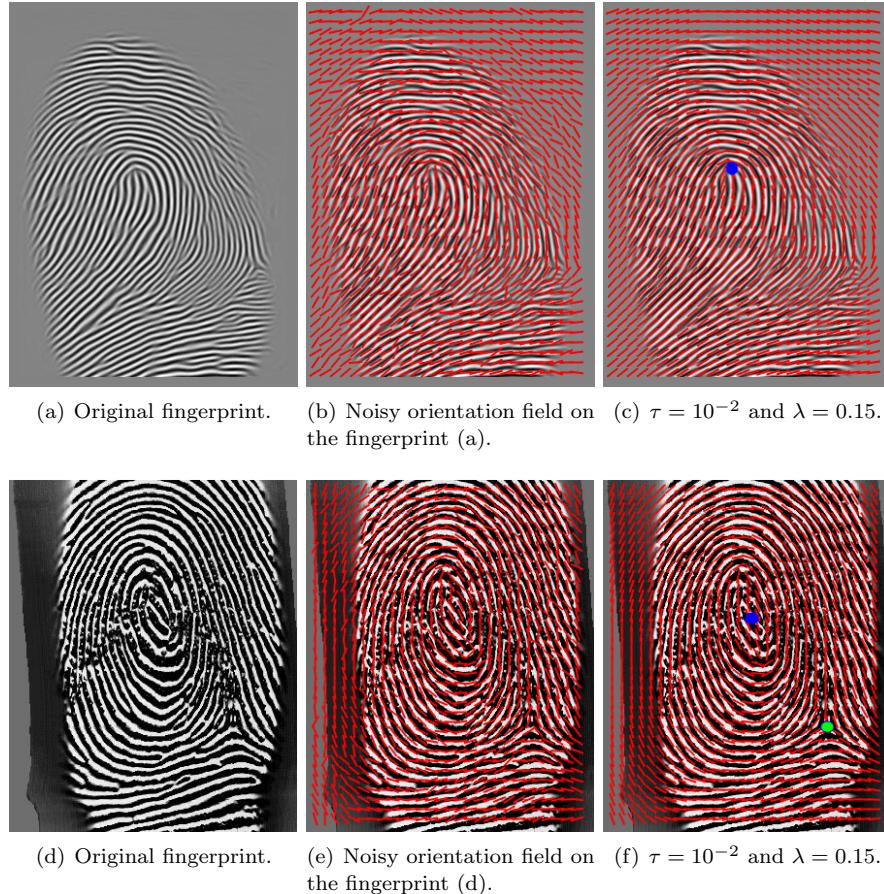


FIGURE 13. Denoising results for two fingerprint images with $\tau = 10^{-2}$ and $\lambda = 0.15$. See Section 4.7.

calculate the gradient of the intensity to roughly extract a line field, $\{f_{ij}\} \subset \mathbb{RP}^1$; see Figures 13(b) and (e). We apply Algorithm 1 on the ‘squared field’ $\{f_{ij}^2\}$ and display the denoised results in Figures 13(c) and 13(f). From the denoised line field we observe that the line field is tangential to the fingerprint, in particular, the singularities are well resolved. In Figure 13(c), the blue dot indicates a singularity with index 1/2. In Figure 13(f), there are two singularities: the one indicated by the blue dot has index 1 and the one indicated by the green dot has index -1/2.

In both experiments, we choose $\tau = 10^{-2}$ and $\lambda = 0.15$. We extracted the line field at 38×28 points for 13(a) and 35×25 points for 13(d). Both simulations were done in 5×10^{-3} seconds, which demonstrates the efficiency of the proposed algorithm.

5. Discussion. In this paper, we introduced and analyzed a nonlocal energy for denoising target-valued images. We derived a diffusion generated method to minimize the energy and performed a variety of numerical experiments to show that the method is efficient, stable, and applicable to a wide variety of target-sets. There are a variety of interesting future directions for this work.

The closest comparison for our numerical results can be found in [50, 3]. The models developed in these papers are nonsmooth variational models which include total variation or second order differences in the regularization term. While we expect that these methods preserve edges better than the proposed method, the results in the numerical experiments are visually very similar. However, due to the simplicity of viewing the target set in an ambient Euclidean space, our methods should be faster. As with any inverse problem, the ‘best’ method depends on the structure of the image and the noise as well as the size of the data. More work should be done to understand the statistical framework for which these methods are consistent and robust estimators.

In this method, we have taken the domain, Ω , to be a Euclidean set. It would be very interesting to consider the case when Ω is a graph and the energy (4) is formulated using the analogous graph operators [6, 41, 10, 5].

In this work, we have only looked at the denoising problem for target-valued images. Other image analysis tasks for target-valued images, including inpainting, segmentation, and registration, could be handled using similar techniques.

Appendix A. Proof of Lemmas 2.1 and 3.2.

Proof of Lemma 2.1. (i) Using that for $u_0 \in L^2(\Omega; \mathbb{R}^k)$, we have that $\|e^{\Delta\tau} u_0\|_{L^2(\Omega; \mathbb{R}^k)} \leq \|u_0\|_{L^2(\Omega; \mathbb{R}^k)}$. We compute

$$\begin{aligned} E_{\lambda, \tau}(u) &= -\frac{1}{2} \langle u, (e^{\Delta\tau} - I)u \rangle + \frac{\lambda}{2} \langle u - f, e^{\Delta\tau}(u - f) \rangle \\ &= \frac{1}{2} \|u\|^2 - \frac{1}{2} \|e^{\Delta\tau/2} u\|^2 + \frac{\lambda}{2} \|e^{\Delta\tau/2}(u - f)\|^2 \\ &\geq 0 \end{aligned}$$

(ii) Let $u, v \in L^2(\Omega; \mathbb{R}^k)$. Observe that

$$E_{\lambda, \tau}(u) = \frac{1}{2} \langle u, (I - (1 - \lambda)e^{\Delta\tau}) u \rangle - \lambda \langle e^{\Delta\tau} f, u \rangle + \frac{\lambda}{2} \langle f, e^{\Delta\tau} f \rangle.$$

Using $\|e^{\Delta\tau} u_0\|_{L^2(\Omega; \mathbb{R}^k)} \leq \|u_0\|_{L^2(\Omega; \mathbb{R}^k)}$ again, we compute

$$\begin{aligned} |E_{\lambda, \tau}(u) - E_{\lambda, \tau}(v)| &= \left| \frac{1}{2} \langle (I - (1 - \lambda)e^{\Delta\tau})(u + v), u - v \rangle - \lambda \langle e^{\Delta\tau} f, u - v \rangle \right| \\ &\leq (\|u + v\| + \lambda \|e^{\Delta\tau} f\|) \|u - v\|. \end{aligned}$$

Let $\varepsilon > 0$, $\delta = \min \left\{ 1, \frac{\varepsilon}{1+2\|v\|+\lambda\|e^{\Delta\tau} f\|} \right\}$, and $\|u - v\| \leq \delta$. Then

$$|E_{\lambda, \tau}(u) - E_{\lambda, \tau}(v)| \leq (\|u - v\| + 2\|v\| + \lambda \|e^{\Delta\tau} f\|) \|u - v\| \leq \varepsilon.$$

(iii) From the definition, we directly compute

$$\delta E_{\lambda, \tau}(u) = \frac{d(E_{\lambda, \tau}(u + s\delta u))}{ds} \Big|_{s=0} = \langle \delta u, -(e^{\Delta\tau} - I)u + \lambda e^{\Delta\tau}(u - f) \rangle.$$

(iv) For $\lambda \in (0, 1)$, we compute

$$\begin{aligned} \left\| \frac{\delta E_{\lambda,\tau}(u)}{\delta u} - \frac{\delta E_{\lambda,\tau}(u)}{\delta v} \right\| &= \| - (e^{\Delta\tau} - I)(u - v) + \lambda e^{\Delta\tau}(u - v) \| \\ &\leq (1 - \lambda) \| e^{\Delta\tau}(u - v) \| + \| u - v \| \\ &\leq (1 - \lambda) \| u - v \| + \| u - v \| \\ &= L \| u - v \|. \end{aligned}$$

(v) From direct calculation, we have for any $u, v \in L^2(\Omega; T)$ and any $\gamma \in (0, 1)$,

$$\begin{aligned} \gamma E_{\lambda,\tau}(u) + (1 - \gamma) E_{\lambda,\tau}(v) - E_{\lambda,\tau}(\gamma u + (1 - \gamma)v) \\ = -\frac{1}{2}\gamma(1 - \gamma) \langle u - v, ((1 - \lambda)e^{\Delta\tau} - I)(u - v) \rangle \\ \geq -\frac{1}{2}\gamma(1 - \gamma)(1 - \lambda - 1) \| u - v \|^2 \\ = \frac{\lambda}{2}\gamma(1 - \gamma) \| u - v \|^2. \end{aligned}$$

This implies that $E_{\lambda,\tau}$ is strongly convex with constant λ .

(vi) Using the semi-group property and maximum principle of the operator $e^{\Delta\tau}$, we have

$$\langle u, (I - e^{\Delta\tau})u \rangle = \|u\|^2 - \|e^{\Delta\tau/2}u\|^2 \geq \|u\|^2 - \|e^{\Delta\tau/2}\|^2 \|u\|^2 \geq 0.$$

We then compute

$$\begin{aligned} E_{\lambda,\tau}(u) &= -\frac{1}{2} \langle u, (e^{\Delta\tau} - I)u \rangle + \frac{\lambda}{2} \langle u - f, e^{\Delta\tau}(u - f) \rangle \\ &= \frac{1}{2}(1 - \lambda) \langle u, (I - e^{\Delta\tau})u \rangle + \frac{\lambda}{2} \|u\|^2 - \lambda \langle e^{\Delta\tau}f, u \rangle + \frac{\lambda}{2} \|e^{\Delta\tau/2}f\|^2 \\ &\geq \frac{\lambda}{2} \|u\|^2 - \lambda \|e^{\Delta\tau}f\| \|u\| + \frac{\lambda}{2} \|e^{\Delta\tau/2}f\|^2 \\ &= \frac{\lambda}{2} (\|u\| - \|e^{\Delta\tau}f\|)^2 + \frac{\lambda}{2} \left(\|e^{\Delta\tau/2}f\|^2 - \|e^{\Delta\tau}f\|^2 \right). \end{aligned}$$

(vii) We have shown that $E_{\lambda,\tau}$ is convex and continuous, hence lower semi-continuous. By, e.g., [2, Theorem 3.3.3], $E_{\lambda,\tau}$ is weakly lower semi-continuous. \square

Proof of Lemma 3.2. If T is compact, then there exists $M > 0$ such that $|z| \leq M$ for every $z \in T$. Let $f \in L^p(\Omega; T)$ for $p \geq 1$. Then $\text{ess.im}(f) \subset T$ implies that for all $z \in \mathbb{R}^k \setminus T$, there exists $\varepsilon > 0$ such that

$$\mu(\{x \in \Omega : |f(x) - z| \leq \varepsilon\}) = 0.$$

In particular, we have that $\mu(\{x \in \Omega : |f(x)| > M\}) = 0$ which shows that

$$\inf\{a \in \mathbb{R} : \mu(\{x \in \Omega : |f(x)| > a\}) = 0\} \leq M,$$

which implies that $f \in L^\infty(\Omega; T)$. \square

Acknowledgments. We would like to thank the referees for their helpful comments.

REFERENCES

- [1] G. Alberti and G. Bellettini, *A non-local anisotropic model for phase transitions: Asymptotic behaviour of rescaled energies*, *European Journal of Applied Mathematics*, **9** (1998), 261–284.
- [2] H. Attouch, G. Buttazzo and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*, Society for Industrial and Applied Mathematics, 2006, URL <http://dx.doi.org/10.1137/1.9780898718782>.
- [3] M. Bačák, R. Bergmann, G. Steidl and A. Weinmann, *A second order nonsmooth variational model for restoring manifold-valued images*, *SIAM Journal on Scientific Computing*, **38** (2016), A567–A597.
- [4] J. M. Ball, *Mathematics and liquid crystals*, *Molecular Crystals and Liquid Crystals*, **647** (2017), 1–27.
- [5] R. Bergmann and D. Tenbrinck, *A graph framework for manifold-valued data*, *SIAM Journal on Imaging Sciences*, **11** (2018), 325–360.
- [6] A. L. Bertozzi and A. Flenner, *Diffuse interface models on graphs for classification of high dimensional data*, *SIAM Review*, **58** (2016), 293–328.
- [7] D. P. Bertsekas, *Convex Optimization Algorithms*, Athena Scientific, 2015.
- [8] F. Bethuel, H. Brezis and F. Hélein, *Ginzburg-Landau Vortices*, vol. 13 of Progress in nonlinear differential equations and their applications, Birkhäuser Boston, 1994.
- [9] L. A. Cafferelli and F. H. Lin, *An optimal partition problem for eigenvalues*, *Journal of Scientific Computing*, **31** (2007), 5–18.
- [10] L. Calatroni, Y. van Gennip, C.-B. Schönlieb, H. M. Rowland and A. Flenner, *Graph clustering, variational image segmentation methods and hough transform scale detection for object measurement in images*, *Journal of Mathematical Imaging and Vision*, **57** (2016), 269–291.
- [11] P. Cook, Y. Bai, S. Nedjati-Gilani, K. Seunarine, M. Hall, G. Parker and D. C. Alexander, Camino: Open-source diffusion-mri reconstruction and processing, <http://camino.cs.ucl.ac.uk/>, 2006.
- [12] W. E and X.-P. Wang, *Numerical methods for the Landau-Lifshitz equation*, *SIAM Journal on Numerical Analysis*, **38** (2000), 1647–1665.
- [13] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Society for Industrial and Applied Mathematics, 1999.
- [14] S. Esedoglu and F. Otto, *Threshold dynamics for networks with arbitrary surface tensions*, *Communications on Pure and Applied Mathematics*, **68** (2015), 808–864.
- [15] S. Esedoglu, S. Ruuth and R. Tsai, *Diffusion generated motion using signed distance functions*, *Journal of Computational Physics*, **229** (2010), 1017–1042.
- [16] S. Esedoglu, S. J. Ruuth and R. Tsai, *Threshold dynamics for high order geometric motions*, *Interfaces and Free Boundaries*, **10** (2008), 263–282.
- [17] S. Esedoglu and Y.-H. R. Tsai, *Threshold dynamics for the piecewise constant Mumford–Shah functional*, *Journal of Computational Physics*, **211** (2006), 367–384.
- [18] P. Grohs, M. Sprecher and T. Yu, *Scattered manifold-valued data approximation*, *Numerische Mathematik*, **135** (2016), 987–1010.
- [19] B. M. Kampes, *Radar Interferometry*, Springer Netherlands, 2006.
- [20] F. Laus, M. Nikolova, J. Persch and G. Steidl, *A nonlocal denoising algorithm for manifold-valued images using second order statistics*, *SIAM Journal on Imaging Sciences*, **10** (2017), 416–448.
- [21] T. Laux and F. Otto, *Convergence of the thresholding scheme for multi-phase mean-curvature flow*, *Calculus of Variations and Partial Differential Equations*, **55** (2016), Art. 129, 74 pp.
- [22] T. Laux and D. Swartz, *Convergence of thresholding schemes incorporating bulk effects*, *Interfaces and Free Boundaries*, **19** (2017), 273–304.
- [23] T. Laux and N. K. Yip, *Analysis of diffusion generated motion for mean curvature flow in codimension two: A gradient-flow approach*, *Archive for Rational Mechanics and Analysis*, **232** (2018), 1113–1163.
- [24] C. Lenglet, J. Campbell, M. Descoteaux, G. Haro, P. Savadjiev, D. Wassermann, A. Anwander, R. Deriche, G. Pike, G. Sapiro, K. Siddiqi and P. Thompson, *Mathematical methods for diffusion MRI processing*, *NeuroImage*, **45** (2009), S111–S122.
- [25] C. Li, C.-Y. Kao, J. C. Gore and Z. Ding, *Minimization of region-scalable fitting energy for image segmentation*, *IEEE Transactions on Image Processing*, **17** (2008), 1940–1949.

- [26] C. B. Macdonald, B. Merriman and S. J. Ruuth, **Simple computation of reaction-diffusion processes on point clouds**, *Proceedings of the National Academy of Sciences*, **110** (2013), 9209–9214.
- [27] A. Majumdar and A. Zarnescu, **Landau-de Gennes theory of nematic liquid crystals: The Oseen-Frank limit and beyond**, *Archive for Rational Mechanics and Analysis*, **196** (2009), 227–280.
- [28] B. Merriman, J. K. Bence and S. Osher, Diffusion generated motion by mean curvature, UCLA CAM Report 92-18, 1992, URL <ftp://ftp.math.ucla.edu/pub/camreport/cam92-18.pdf>.
- [29] B. Merriman, J. Bence and S. Osher, Diffusion generated motion by mean curvature, *AMS Selected Letters, Crystal Grower's Workshop*, 73–83.
- [30] M. Miranda, D. Pallara, F. Paronetto and M. Preunkert, **Short-time heat flow and functions of bounded variation in r^n** , *Annales-Faculte des Sciences Toulouse Mathematiques*, **16** (2007), 125–145.
- [31] B. Osting and T. H. Reeb, **Consistency of Dirichlet partitions**, *SIAM Journal on Mathematical Analysis*, **49** (2017), 4251–4274.
- [32] B. Osting and D. Wang, **A diffusion generated method for orthogonal matrix-valued fields**, *Mathematics of Computation*, **89** (2020), 515–550.
- [33] B. Osting, C. D. White and E. Oudet, **Minimal Dirichlet energy partitions for graphs**, *SIAM Journal on Scientific Computing*, **36** (2014), A1635–A1651.
- [34] F. Rocca, C. Prati and A. Ferretti, An overview of SAR interferometry, in *Proceedings of the 3rd ERS Symposium on Space at the Service of our Environment*, 1997, URL <http://earth.esa.int/workshops/ers97/program-details/speeches/rocca-et-al/>.
- [35] J. Rubinstein, P. Sternberg and J. B. Keller, **Reaction-diffusion processes and evolution to harmonic maps**, *SIAM Journal on Applied Mathematics*, **49** (1989), 1722–1733.
- [36] L. I. Rudin, S. Osher and E. Fatemi, **Nonlinear total variation based noise removal algorithms**, *Physica D: Nonlinear Phenomena*, **60** (1992), 259–268.
- [37] S. J. Ruuth, B. Merriman, J. Xin and S. Osher, **Diffusion-generated motion by mean curvature for filaments**, *Journal of Nonlinear Science*, **11** (2001), 473–493.
- [38] S. J. Ruuth and B. Merriman, **A simple embedding method for solving partial differential equations on surfaces**, *Journal of Computational Physics*, **227** (2008), 1943–1961.
- [39] L. Sciaivocco and B. Siciliano, **Modelling and Control of Robot Manipulators**, Springer London, 2000.
- [40] P. Skraba, B. Wang, G. Chen and P. Rosen, **Robustness-based simplification of 2d steady and unsteady vector fields**, *IEEE Transactions on Visualization and Computer Graphics*, **21** (2015), 930–944.
- [41] Y. van Gennip, N. Guillen, B. Osting and A. L. Bertozzi, **Mean curvature, threshold dynamics, and phase field theory on finite graphs**, *Milan Journal of Mathematics*, **82** (2014), 3–65.
- [42] R. VierTEL and B. Osting, **An approach to quad meshing based on harmonic cross valued maps and the Ginzburg-Landau theory**, *SIAM Journal on Scientific Computing*, **41** (2019), A452–A479.
- [43] B. A. Wandell, Clarifying human white matter, *Annual Review of Neuroscience*, **39** (2016), 103–128.
- [44] D. Wang, A. Cherkaev and B. Osting, **Dynamics and stationary configurations of heterogeneous foams**, *PLOS ONE*, **14** (2019), e0215836.
- [45] D. Wang, H. Li, X. Wei and X.-P. Wang, **An efficient iterative thresholding method for image segmentation**, *Journal of Computational Physics*, **350** (2017), 657–667.
- [46] D. Wang and B. Osting, **A diffusion generated method for computing Dirichlet partitions**, *Journal of Computational and Applied Mathematics*, **351** (2019), 302–316.
- [47] D. Wang, B. Osting and X.-P. Wang, **Interface dynamics for an Allen-Cahn-type equation governing a matrix-valued field**, *SIAM Multiscale Modelling and Simulation*, **17** (2019), 1252–1273.
- [48] D. Wang and X.-P. Wang, The iterative convolution-thresholding method (ictm) for image segmentation, arXiv preprint, arXiv1904.10917.
- [49] D. Wang, X.-P. Wang and X. Xu, **An improved threshold dynamics method for wetting dynamics**, *Journal of Computational Physics*, **392** (2019), 291–310.
- [50] A. Weinmann, L. Demaret and M. Storath, **Total variation regularization for manifold-valued data**, *SIAM Journal on Imaging Sciences*, **7** (2014), 2226–2257.
- [51] X. Xu, D. Wang and X.-P. Wang, **An efficient threshold dynamics method for wetting on rough surfaces**, *Journal of Computational Physics*, **330** (2017), 510–528.

- [52] D. Zosso and B. Osting, [A minimal surface criterion for graph partitioning](#), *AIMS Inverse Problems and Imaging*, **10** (2016), 1149–1180.

Received November 2018; revised October 2019.

E-mail address: osting@math.utah.edu

E-mail address: dwang@math.utah.edu