Mean curvature, threshold dynamics, and phase field theory on finite graphs

Yves van Gennip, Nestor Guillen, Braxton Osting, Andrea L. Bertozzi

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA 90095, USA
{yvgennip, nestor, braxton, bertozzi}@math.ucla.edu

June 28, 2013

Abstract

In the continuum, close connections exist between mean curvature flow, the Allen-Cahn (AC) partial differential equation, and the Merriman-Bence-Osher (MBO) threshold dynamics scheme. Graph analogues of these processes have recently seen a rise in popularity as relaxations of NP-complete combinatorial problems, which demands deeper theoretical underpinnings of the graph processes. The aim of this paper is to introduce these graph processes in the light of their continuum counterparts, provide some background, prove the first results connecting them, illustrate these processes with examples and identify open questions for future study.

We derive a graph curvature from the graph cut function, the natural graph counterpart of total variation (perimeter). This derivation and the resulting curvature definition differ from those in earlier literature, where the continuum mean curvature is simply discretized, and bears many similarities to the continuum nonlocal curvature or nonlocal means formulation. This new graph curvature is not only relevant for graph MBO dynamics, but also appears in the variational formulation of a discrete time graph mean curvature flow.

We prove estimates showing that the dynamics are trivial for both MBO and AC evolutions if the parameters (the time-step and diffuse interface scale, respectively) are sufficiently small (a phenomenon known as "freezing" or "pinning") and also that the dynamics for MBO are nontrivial if the time step is large enough. These bounds are in terms of graph quantities such as the spectrum of the graph Laplacian and the graph curvature. Adapting a Lyapunov functional for the continuum MBO scheme to graphs, we prove that the graph MBO scheme converges to a stationary state in a finite number of iterations. Variations on this scheme have recently become popular in the literature as ways to minimize (continuum) nonlocal total variation.

Keywords: spectral graph theory, Allen-Cahn equation, Ginzburg-Landau functional, Merriman-Bence-Osher threshold dynamics, graph cut function, total variation, mean curvature flow, nonlocal mean curvature, gamma convergence, graph coarea formula.

MSC2010: 34B45, 35R02, 53C44, 53A10, 49K15, 49Q05, 35K05.

Contents

1 Introduction

Motion by mean curvature and flows involving mean curvature in general appear in many important *continuum* models, including models coming from materials science [MS64, Tay92], fluid dynamics [HS98], and combustion [XY11, Pet00]. All such models involve a front propagating with a velocity depending on the mean curvature of the front. Recently, there has been an increasing interest in using ideas from continuum PDEs (related to mean curvature) in *discrete* applications such as image analysis, machine learning and data clustering [BF12, vGB12, MKB12, GCMB+13, HLPB13].

This paper initiates a systematic study of the definition of mean curvature for vertex sets of an arbitrary graph $G = (V, E)$. We examine the effectiveness of the algorithms in the recent papers mentioned above and on how they may be improved. The graphs considered are arbitrary graphs and are not necessarily obtained as the discretization of a continuum problem, so our perspective is only parallel to one that is purely motivated by numerical analysis. In particular, we do not assume an embedding of the graph in a low dimensional space.

Of course, the various definitions of curvature in the (usual) continuum setting (see Appendix A for a brief review and some references) motivate and inform our approach to defining the curvature of a vertex set $S \subset V$ using the discrete total variation norm and the discrete divergence of a "normal" edge flow. Since they are closely related to questions of mean curvature, the Allen-Cahn equation and the MBO scheme for arbitrary graphs $G = (V, E)$ arise naturally in the present investigation. Theoretical and numerical examples are used to highlight possible connections between all these concepts, leading to a number of open questions given in Section 7.

Graphs Laplacians, Allen-Cahn, and MBO. Graph Laplacians are the central objects of study in spectral graph theory [Chu97]. These graph operators share many properties with their continuum counterparts. The Allen-Cahn equation on the graph V is defined in terms of the graph Laplacian, Δ, and any (typically bistable quartic) potential, W. One considers a phase field, $u\colon V \times \mathbb{R}_+ \to \mathbb{R}$, solving the differential equation,

$$\dot{u} = -\Delta u - \frac{1}{\varepsilon} W'(u).$$

This nonlinear equation has received greater attention recently, spurred by some of the applications mentioned above. The graph Allen-Cahn equation was introduced in the context of data classification in [BF12] and, in a number of examples, was shown to be both accurate and efficient. As is well known, the continuum Allen-Cahn equation is the L^2 gradient flow associated to the Ginzburg-Landau functional. This is also true in the graph setting. In [vGB12] it was shown that the graph Ginzburg-Landau functional Γ-converges to the graph cut objective functional on graphs, if the characteristic length scale ε goes to zero. Moreover, a relationship between the graph cut functional and the continuum total variation functional was given. At the same time, the continuum Allen-Cahn solution is known to converge to mean curvature flow, when $\varepsilon \to 0$ [BSS93]. Furthermore, the mean curvature is directly related to the first variation of the total variation functional. In this paper we therefore study the graph Allen-Cahn equation and make connections between it and a graph cut derived 'graph curvature'.

The third ingredient in this paper is the threshold dynamics or Merriman-Bence-Osher (MBO) algorithm on graphs. Its continuum counterpart was introduced in [MBO92, MBO93] and consists of iteratively solving the heat equation for a short time, τ, and thresholding the result to an indicator function. It is known that, for short diffusion times τ, this approximates mean curvature flow [Eva93, BG95]. In this paper we therefore also study the connections between the graph MBO scheme, the graph AC equation, and graph curvature.

In a recent series of papers [ELB08, TEL08, DELT10, DEL11, TEL11, DEL12a, EDL12, DEL12c, DEL12b] Elmoataz et al. study partial differential equations and front propagation on graphs, mainly from a numerical point of view. In these papers the 1-Laplacian on a graph is used as curvature, which differs from our approach. We use the anisotropic graph total variation instead of the isotropic total variation (see Section 2 for definitions), since [vGB12] suggests that is the natural total variation on graphs.

A common obstacle, when transferring results and intuitions from the continuum to graphs, is the (implicit) lower bound on the accessible length scales on a graph. We show in Theorem 5.3 and Theorem 4.2, that if ε or τ is too small, then the Allen-Cahn equation exhibits "freezing" (or "pinning"), or the MBO scheme is stationary, respectively. Hence the interesting dynamics happen at small but positive, ε or τ, rather than in the limits as $\varepsilon \to 0$ or $\tau \to 0$. Related is the lack of a chain rule for derivatives on graphs[1], which can be traced back to the absence of 'infinitesimal' length scales on a graph. As a consequence, the level set approach, which has proven very successful in describing continuum mean curvature flow, is not independent of the level set function on a graph.

New results. The finite spectral radius of the Laplacian is used to derive explicit bounds on the parameters for both threshold dynamics (MBO) and the graph Allen-Cahn equation that guarantee "freezing" or "pinning" of the initial phase, a phenomenon that has been observed in numerical simulations and is well known for discretizations of the continuum processes [MBO92][MBO94, Section 4.4].

In the opposite direction, an argument based on the comparison principle is used to obtain a lower bound for the MBO time step that guarantees that a specific node of the phase changes in a single MBO iteration. This bound is given in terms of a new notion of mean curvature for general graphs, and as such, it is a "local" quantity (as opposed to one coming from spectral data). Such local bounds may be of use in developing adaptive time stepping schemes that complement the (spectral) adaptive schemes, such as those developed for discretizations of the continuum mean curvature flow [Ruu96, ZD09]. In this sense, introducing the graph mean curvature and highlighting its connection with subjects in continuum PDE (MBO, Ginzburg-Landau, and nonlocal mean curvature) and graph theory (graph cuts, connectivity) are the main contributions of this work.

The results in Sections 4 and 5 and the numerical evidence and explicit examples in Section 6 suggest several open questions about the graph MBO scheme, the graph Allen-Cahn equation, and graph mean curvature, which are discussed in Section 7. These are interesting questions for future work.

Outline. In Section 2 the relevant graph based calculus is introduced, setting the notation for the rest of the paper. In particular, the graph Laplacian and its basic properties are discussed. Section 3 discusses curvature and mean curvature flow on a graph. Sections 4 and 5 discuss the MBO scheme and Allen-Cahn equation on graphs, respectively, and sufficient conditions are given on the parameters to guarantee freezing or pinning of the initial conditions. Section 6 explores the graph processes and concepts introduced in these previous three sections through theoretical and computational examples. Finally, we conclude in Section 7, with a discussion and a few open questions based on the new estimates and examples from previous sections. In Appendix A, we make some remarks on the continuum mean curvature flow. In Appendix B, we derive the graph co-area formula. Appendix C discusses some similarities between the graph Laplacian, the graph 1-Laplacian, and the graph curvature.

[1] That is, if $u \in \mathcal{V}$ and $f\colon \mathcal{V} \to \mathcal{V}$, then $\nabla f(u) \neq L_f \nabla u$ for any linear operator L_f.

2 Setup

We work on a finite[2], undirected graph $G = (V, E)$ with vertex set[3] $V = \{i\}_{i=1}^n$ and edge set $E \subset V^2$. The graph is weighted; each edge $(i, j) \in E$, incident on nodes i and j, is assigned a weight $\omega_{ij} > 0$. Since the graph is undirected, (i, j) is identified with (j, i) and $\omega_{ij} = \omega_{ji}$. To simplify notation we extend ω_{ij} to be zero for all $(i, j) \in V^2$ which do not correspond to an edge. The degree of node i is $d_i := \sum_{j\in V} \omega_{ij}$. Denote the maximal and minimal degrees by $d_+ := \max_{i\in V} d_i$ and $d_- := \min_{i\in V} d_i$. We assume that G has no isolated nodes, *i.e.*, $d_- > 0$. For each $i \in V$ we then have a non-empty set of neighbors $\mathcal{N}_i := \{j \in V : \omega_{ij} > 0\}$. We also assume G has no self-loops, *i.e.*, $\omega_{ii} = 0$. In particular $i \notin \mathcal{N}_i$.

Let $\mathcal{V}$ be the space of all functions $V \to \mathbb{R}$ and $\mathcal{E}$ the space of all skew-symmetric[4] functions $E \to \mathbb{R}$. Again to simplify notation, we extend each $\varphi \in \mathcal{E}$ to a function $\varphi\colon V^2 \to \mathbb{R}$ by setting $\varphi_{ij} = 0$ if $(i, j) \notin E$. As justified in earlier work [vGB12] we introduce the following inner products and operators for parameters $q \in [1/2, 1]$ and $r \in [0, 1]$:

$$\langle u, v\rangle_{\mathcal{V}} := \sum_{i\in V} u_i v_i d_i^r, \quad \langle \varphi, \phi\rangle_{\mathcal{E}} := \frac{1}{2}\sum_{i,j\in V} \varphi_{ij}\phi_{ij}\omega_{ij}^{2q-1}, \quad (\varphi \cdot \phi)_i := \frac{1}{2}\sum_{j\in V} \varphi_{ij}\phi_{ij}\omega_{ij}^{2q-1},$$
$$(\nabla u)_{ij} := \omega_{ij}^{1-q}(u_j - u_i), \qquad (\operatorname{div}\varphi)_i := \frac{1}{2d_i^r}\sum_{j\in V} \omega_{ij}^q(\varphi_{ji} - \varphi_{ij}).$$

Note that in the sum in $\langle \varphi, \phi\rangle_{\mathcal{E}}$ the indices i and j both run over all nodes. The edges (i, j) and (j, i) are counted separately, hence the correction factor $\frac{1}{2}$. Note that the powers $2q - 1$ and $1 - q$ in the $\mathcal{E}$ inner product and 'dot product' and in the gradient, are zero for the admissible choices $q = \frac{1}{2}$ and $q = 1$ respectively. In these cases we define $\omega_{ij}^0 = 0$ whenever $\omega_{ij} = 0$, so as not to make the inner product, 'dot product', or gradient, nonlocal on the graph. The inner products on $\mathcal{V}$ and $\mathcal{E}$[5] are analogous to a weighted L^2 inner products in the continuum case, while the 'dot product' inner product $(\varphi \cdot \phi)_i$ is analogous to a weighted dot product between vector(field)s (without the integration of the L^2 inner product). A direct computation shows that div and ∇ are adjoints with respect to $\langle\cdot,\cdot\rangle_{\mathcal{V}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{E}}$, namely for $u \in \mathcal{V}$ and $\phi \in \mathcal{E}$ we have

$$\langle \nabla u, \phi\rangle_{\mathcal{E}} = \langle u, \operatorname{div}\phi\rangle_{\mathcal{V}}.$$

The characteristic function of a node set $S \subset V$ is $\chi_S \in \mathcal{V}$, defined via $(\chi_S)_i := \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S. \end{cases}$

This leads to the following associated norms, Laplacians, set measures, and total variation functionals:

- Inner product norms, $\|u\|_{\mathcal{V}} := \sqrt{\langle u, u\rangle_{\mathcal{V}}}$ and $\|\varphi\|_{\mathcal{E}} := \sqrt{\langle \varphi, \varphi\rangle_{\mathcal{E}}}$.
- Maximum norms[6] , $\|u\|_{\mathcal{V},\infty} := \max\{|u_i|\colon i \in V\}$ and $\|\varphi\|_{\mathcal{E},\infty} := \max\{|\varphi_{ij}|\colon i, j \in V\}$.

[2] In this paper, we are working with a fixed graph G with a finite number of nodes. In no sense are we considering a sequence of graphs or taking a "continuum limit".

[3] We will use the terms "vertex" and "node" interchangeably.

[4] The necessity of skew-symmetry may not be obvious at this point, but it is a common requirement for consistency of certain concepts in discrete calculus, see *e.g.*, [FT04, GO09, GP10, CMOP11].

[5] Note that $\langle\cdot,\cdot\rangle_{\mathcal{E}}$ is indeed an inner product on the space of (skew-symmetric) functions $E \to \mathbb{R}$, but not for the space of functions $V^2 \to \mathbb{R}$, because for those functions the 'inner product' can be zero for nontrivial functions.

[6] To justify these definitions and convince ourselves that there should be no ω_{ij} or d_i included in the maximum norms we define $\|\varphi\|_{\mathcal{E},p}^p := \frac{1}{2}\sum_{i,j\in V} \varphi_{ij}^p \omega_{ij}^{2q-1}$. Adapting the proofs in the continuum case, *e.g.*, [Ada75, Theorems

- The norm corresponding to the dot product $|\varphi|_i := \sqrt{(\varphi \cdot \varphi)_i}$. Note that $|\phi| \in \mathcal{V}$.
- The Dirichlet energy
$$\frac{1}{2}\|\nabla u\|_{\mathcal{E}}^2 = \frac{1}{4}\sum_{i,j\in V} \omega_{ij}(u_i - u_j)^2.$$
- The graph Laplacian $\Delta := \operatorname{div} \circ \nabla \colon \mathcal{V} \to \mathcal{V}$. So
$$(\Delta u)_i := d_i^{-r} \sum_{j\in V} \omega_{ij}(u_i - u_j).$$
It is worth noting that the sign convention for the graph Laplacian is opposite to that used for the continuum Laplacian in most of the PDE literature (in particular, the graph Laplacian Δ is a positive semidefinite operator).
When $r = 0$, Δ is referred to as the unnormalized weighted graph Laplacian. When $r = 1$, Δ is referred to as the asymmetric normalized graph Laplacian or random walk graph Laplacian. Another Laplacian, often encountered in the spectral graph theory literature, is the symmetric normalized graph Laplacian. This one falls outside the scope of the current setup and will not be considered in this paper. For general references on the graph Laplacian, consult [Moh91, Chu97, vL07, BLS07].
- For $S \subset V$, the set measures
$$\begin{aligned} \operatorname{vol} S &= \sum_{i\in S} d_i^r = \|\chi_S\|_{\mathcal{V}}^2, \\ |S| &= \text{number of vertices in } S. \end{aligned}$$
Note that $|S|$ is just a special case of vol S, for $r = 0$ (recall we assume $d_- > 0$).
- The isotropic and anisotropic total variation $\mathrm{TV} \colon \mathcal{V} \to \mathbb{R}$ and $\mathrm{TV}_a^q \colon \mathcal{V} \to \mathbb{R}$ respectively:
$$\begin{aligned} \mathrm{TV}(u) &:= \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} \colon \varphi \in \mathcal{E},\ \max_{i\in V} |\varphi|_i \le 1\} \\ &= \sum_{i\in V} |\nabla u|_i = \frac{\sqrt{2}}{2}\sum_{i\in V}\sqrt{\sum_{j\in V}\omega_{ij}(u_i-u_j)^2} \\ \mathrm{TV}_{\mathrm{a}}^q(u) &:= \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} \colon \varphi \in \mathcal{E},\ \|\varphi\|_{\mathcal{E},\infty} \le 1\} \\ &= \langle \nabla u, \operatorname{sgn}(\nabla u)\rangle_{\mathcal{E}} = \frac{1}{2}\sum_{i,j\in V} \omega_{ij}^q |u_i - u_j|. \end{aligned}$$
Here, the signum function is understood to act element-wise on the elements of ∇u. We note that the maximum in the definition of TV is achieved by taking
$$\varphi_{ij} = \varphi_{ij}^{TV} := \begin{cases} \frac{(\nabla u)_{ij}}{|\nabla u|_i} & |\nabla u|_i \neq 0 \\ 0 & |\nabla u|_i = 0. \end{cases}$$

2.3 and 2.8], to the graph situation we can prove a Hölder inequality $\|\varphi\phi\|_{\mathcal{E},1} \le \|\varphi\|_{\mathcal{E},p}\|\phi\|_{\mathcal{E},p'}$ for $1 < p, p' < \infty$ such that $\frac{1}{p} + \frac{1}{p'} = 1$, an embedding theorem of the form $\|\varphi\|_{\mathcal{E},p} \le \left(\frac{1}{2}\sum_{i,j\in V}\omega_{ij}^{2q-1}\right)^{\frac{1}{p}-\frac{1}{s}} \|\varphi\|_{\mathcal{E},s}$ for $1 \le p \le s \le \infty$, and the limit $\lim_{p\to\infty}\|\varphi\|_{\mathcal{E},p} = \|\varphi\|_{\mathcal{E},\infty}$. A similar result holds for the norms on $\mathcal{V}$.

The maximum in the definition of $\mathrm{TV}_{\mathrm{a}}^q$ is achieved by

$$\varphi = \varphi^a := \operatorname{sgn}(\nabla u). \tag{1}$$

Note that the values φ^{TV} and φ^a take on the set $\{\nabla u = 0\}$ are irrelevant for achieving the maximum, hence these functions are not uniquely determined. The quantity $\operatorname{div} \frac{\nabla u}{|\nabla u|}$ is often referred to as the 1-Laplacian of u.

The anisotropic total variation of the indicator function[7] for the set $S \subset V$, denoted χ_S, is given by

$$\mathrm{TV}_{\mathrm{a}}^q(\chi_S) = \sum_{i \in S, j \in S^c} \omega_{ij}^q. \tag{2}$$

Thus, the total variation of a set S is equivalent to the graph cut between S and $S^c := V \setminus S$ which is used in graph theory and spectral clustering [SM00]. For future reference it is useful to note that

$$\mathrm{TV}_{\mathrm{a}}^1(\chi_S) = \langle \nabla \chi_S, \nabla \chi_S \rangle_{\mathcal{E}} = \langle \chi_S, \Delta \chi_S \rangle_{\mathcal{V}}. \tag{3}$$

Lemma 2.1. *The norms $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{V},\infty}$ are equivalent, with optimal constants given by*

$$d_-^{\frac{r}{2}} \|u\|_{\mathcal{V},\infty} \leq \|u\|_{\mathcal{V}} \leq \sqrt{\operatorname{vol} V} \|u\|_{\mathcal{V},\infty}.$$

Proof. We compute $\|u\|_{\mathcal{V}}^2 = \sum_{i \in V} d_i^r u_i^2 \leq \max_{i \in V} u_i^2 \sum_{i \in V} d_i^r = (\operatorname{vol} V) \|u\|_{\mathcal{V},\infty}^2$, which is saturated if $u = \chi_V$. Also, $\|u\|_{\mathcal{V}}^2 = \sum_{i \in V} d_i^r u_i^2 \geq d_-^r \max_{i \in V} u_i^2 = d_-^r \|u\|_{\mathcal{V},\infty}^2$. If $j \in V$ is such that $d_j = d_-$, this bound is attained for $u = \chi_{\{j\}}$. □

Next we recall the definitions of node set boundaries and (signed) graph distance.

Definition 2.2. *For $j \in \mathcal{N}_i$, we define $d_{ij}^G := \omega_{ij}^{q-1}$, and we set $d^G(i,i) := 0$. A path on V is a sequence $\gamma = \{i_1, i_2, ..., i_m\}$ for some $m \in \mathbb{N}$ such that $i_{k+1} \in \mathcal{N}_{i_k}$ for each $k \in \{1, \ldots, m-1\}$. Given a path $\gamma = \{i_1, ..., i_m\}$, its length is defined as*

$$|\gamma| := \sum_{i=1}^{m-1} d_{i_k i_{k+1}}^G.$$

Then, the graph distance between arbitrary $i, j \in V$ is given by

$$d_{ij}^G := \min_{\gamma} |\gamma|$$

where the minimum is taken over all paths γ with $i_1 = i, i_N = j$. In other words, d_{ij}^G is the minimal distance to go from node i to node j, traveling only via existing edges, where each edge represents a distance ω_{ij}^{q-1}. For a given set $S \subset V$, we define the graph distance to S at each node as the minimal graph distance to a node in S:

$$d_i^S := \min_{j \in S} d_{ij}^G.$$

As argued in, for example, [MOS12, Section 3.1, Example 2], d^S is the solution $u \in \mathcal{V}$ to the equation

$$\begin{cases} \min_{j \in \mathcal{N}_i} (\nabla u)_{ij} + 1 = 0 & \text{if } i \in V \setminus S, \\ u_i = 0 & \text{if } i \in S. \end{cases} \tag{4}$$

[7] For χ_V, the indicator function of the full node set, we also write the constant function 1.

Definition 2.3. *The boundary of* $S \subset V$ *is*[8]

$$\partial S := \{i \in S \colon \exists j \in V \ \textit{s.t.}\ (\nabla \chi_S)_{ij} < 0\}.$$

Note that $\partial S \subset S$. Alternative definitions appear in the literature in which $\partial S \subset S^c$.

2.1 Basic spectral properties of the graph Laplacian, Δ

In this section, we collect a number of spectral properties of the graph Laplacian $\Delta \colon \mathcal{V} \to \mathcal{V}$. Further discussion and details for the special cases $r = 0$ and $r = 1$ can be found in [Moh91, Chu97, vL07, BLS07], from which our presentation follows.

Note that $\Delta \colon \mathcal{V} \to \mathcal{V}$ is a self-adjoint operator in the $\mathcal{V}$ norm. For $u \in \mathcal{V} \setminus \{0\}$, the *Rayleigh quotient* $R \colon \mathcal{V} \to \mathbb{R}$ is defined as

$$R(u) := \frac{\langle u, \Delta u \rangle_{\mathcal{V}}}{\|u\|_{\mathcal{V}}} = \frac{\|\nabla u\|_{\mathcal{E}}^2}{\|u\|_{\mathcal{V}}^2}.$$

The eigenvalues of the graph Laplacian, Δ, are then defined via the variational formulation,

$$\lambda_k = \min_{\substack{F_k \subset \mathcal{V} \\ \text{subspace of dim } k}} \max_{u \in F_k \setminus \{0\}} R(u). \tag{5}$$

The minimum in (5), is attained when F_k is spanned by the first k eigenfunctions, *i.e.*, the eigenfunctions corresponding to the k smallest eigenvalues, counting multiplicities. In particular, there are n non-negative real eigenvalues (counted with multiplicity), denoted $\{\lambda_k\}_{k=1}^n$. If we denote the span of the first $k-1$ eigenfunctions by $\hat{F}_{k-1} = \mathit{span}(\{u_i\}_{i=1}^{k-1})$, then (5) can be rewritten

$$\lambda_k = \min_{\substack{u \in \mathcal{V} \setminus \{0\} \\ u \perp_{\mathcal{V}} \hat{F}_{k-1}}} R(u). \tag{6}$$

where $u \perp_{\mathcal{V}} \hat{F}_{k-1}$ indicates that u is orthogonal (in the sense of $\langle \cdot, \cdot \rangle_{\mathcal{V}}$) to u_i, for $i \in \{1, \ldots, k-1\}$. Taking variations of the Rayleigh quotient with respect to u, we find that (λ_k, u_k) satisfies (6) if and only if $u \perp_{\mathcal{V}} \hat{F}_{k-1}$ and, for all $v \in \mathcal{V}$,

$$\langle \Delta u_k, v \rangle_{\mathcal{V}} = \lambda_k \langle u_k, v \rangle_{\mathcal{V}}. \tag{7}$$

Finally, unwinding the definitions, we find that (7) is equivalent to the matrix eigenvalue problem

$$Lx = \lambda x \qquad \text{where } L = D^{-r}[D - A], \quad x \in \mathbb{R}^n, \tag{8}$$

where $A_{ij} = \omega_{ij}$ and $D_{ii} = d_i$ is a diagonal matrix[9]. Recall that the *spectral radius* ρ of Δ is defined as the maximum of the absolute values of the eigenvalues of Δ,

$$\rho(\Delta) := \max_{1 \le i \le n} \lambda_i = \lambda_n = \sup_{u \in \mathcal{V} \setminus \{0\}} R(u).$$

[8] Similarly, by changing the "strictly less than" inequalities into "strictly larger than" inequalities the boundary $\partial(S^c)$ of the set S^c is defined. The reduced boundary of S can be defined as the following subset of ∂S (compare with the continuum case in [AFP00, Definition 3.54]):

$$\partial^* S := \{i \in S \colon \exists! j \in V \colon (\nabla \chi_S)_{ij} < 0\},$$

and again similarly for $\partial^* S^c$.

[9] Note that here we use the fact that there are no isolated vertices, *i.e.*, $d_i > 0$ for all $i \in V$.

Lemma 2.4 (Spectral properties of the graph Laplacian, Δ)**.** *The following properties are satisfied:*

(a) The smallest eigenvalue is $\lambda_1 = 0$. The multiplicity of $\lambda_1 = 0$ is the number of connected components of the graph and the associated eigenspace is spanned by set of indicator vectors for each connected component. If there is only one connected component, λ_1 is simple and the first (unnormalized) eigenfunction is $u_1 = 1 = \chi_V$.

(b) The operator norm of Δ, $\|\Delta\|_{\mathcal{V}} := \sup_{u\neq 0} \frac{\|\Delta u\|_{\mathcal{V}}}{\|u\|_{\mathcal{V}}}$, and the spectral radius are equal, $\|\Delta\|_{\mathcal{V}} = \rho(\Delta)$. This implies that, for all $u \in \mathcal{V}$,

$$\|\Delta u\|_{\mathcal{V}} \leq \rho(\Delta)\ \|u\|_{\mathcal{V}}.$$

(c) The trace satisfies $\operatorname{tr}\ \Delta = \sum_{k=1}^n \lambda_k = \sum_{i\in V} d_i^{1-r}$. Consequently,

$$\lambda_2 \leq \frac{1}{n-1}\sum_{i\in V} d_i^{1-r} \leq \frac{n\ d_+^{1-r}}{n-1} \qquad \text{and} \qquad \lambda_n \geq \frac{1}{n-1}\sum_{i\in V} d_i^{1-r} \geq \frac{n\ d_-^{1-r}}{n-1} \tag{9}$$

(d) If G is not a complete graph then

$$\lambda_2 \leq \min_{(i,j)\notin E} \frac{d_i d_j^{2r} + d_i^{2r} d_j}{d_i^r d_j^{2r} + d_i^{2r} d_j^r}.$$

(e) The second eigenvalue satisfies

$$\lambda_2 \leq \min_{S\subset V} \frac{(\operatorname{vol} V) TV_a^1(\chi_S)}{(\operatorname{vol} S)(\operatorname{vol} S^c)} \leq 2 \min_{S\subset V} \frac{TV_a^1(\chi_S)}{\min(\operatorname{vol} S, \operatorname{vol} S^c)}.$$

(f) The spectral radius of Δ satisfies $\rho \leq 2\ d_+^{1-r}$.

Proof. (a) These properties follow directly from (6).

(b) Noting that Δ is a self-adjoint operator, a proof can be found in, for example, [RS79, Thm. VI.6].

(c) Because the trace of the operator Δ is equal to the trace of its matrix representation, we have $\operatorname{tr}\ \Delta = \operatorname{tr}\ L = \sum_{k=1}^n \lambda_k$. Since we assume there are no self-loops in the graph, $\operatorname{tr}\ D^{-r}A = 0$, hence $\operatorname{tr}\ L = \operatorname{tr}\ D^{1-r}$. Equation (9) follows from the fact that $\lambda_1 = 0$ and the maximum (minimum) of a set is greater (less) than or equal to the mean of the set.

(d) If G is not a complete graph, then there exists an $(a, b) \notin E$. We define the test function $v \in \mathcal{V}$,

$$v_i = \begin{cases} d_b^r & i = a \\ -d_a^r & i = b \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\langle v, 1\rangle_{\mathcal{V}} = 0$. The desired upper bound then follows from (6).

(e) For $S \in V$, define the test function $v \in \mathcal{V}$,

$$v_i = \begin{cases} \operatorname{vol} S^c & i \in S \\ -\operatorname{vol} S & i \in S^c. \end{cases}$$

Then $\langle v, 1\rangle_{\mathcal{V}} = 0$ and $\|v\|_{\mathcal{V}}^2 = (\text{vol } S^c)(\text{vol } S)\text{vol } V$. Using (2), we compute $\|\nabla v\|_{\mathcal{E}}^2 = (\text{vol } V)^2 \text{TV}_{\text{a}}^1(\chi_S)$. The first inequality then follows from (6). For the second inequality,

$$\frac{(\text{vol } V)TV_a^1(\chi_S)}{(\text{vol } S)(\text{vol } S^c)} = \frac{TV_a^1(\chi_S)}{\text{vol } S} + \frac{TV_a^1(\chi_S)}{\text{vol } S^c} \leq 2\frac{TV_a^1(\chi_S)}{\min(\text{vol } S, \text{vol } S^c)}.$$

(f) Using the identity $(a-b)^2 \leq 2(a^2+b^2)$, we compute

$$\begin{aligned}\rho(\Delta) &= \sup_{u \in \mathcal{V}\setminus\{0\}} R(u) = \sup_{u \in \mathcal{V}\setminus\{0\}} \frac{1}{2}\frac{\sum_{ij} w_{ij}(u_i - u_j)^2}{\sum_i d_i^r u_i^2} \\ &\leq \sup_{u \in \mathcal{V}\setminus\{0\}} \frac{\sum_{ij} w_{ij}(u_i^2 + u_j^2)}{\sum_i d_i^r u_i^2} \\ &= \sup_{u \in \mathcal{V}\setminus\{0\}} 2\frac{\sum_i d_i u_i^2}{\sum_i d_i^r u_i^2}\end{aligned}$$

If $j \in V$ is such that $d_j = d_+$, then the supremum is attained by the vector $u = \chi_{\{j\}}$ and the result follows. □

The following lemma states properties of the diffusion operator $e^{-\Delta t}\colon \mathcal{V} \to \mathcal{V}$.

Lemma 2.5 (Diffusion on a graph)**.** *Let $u(t) := e^{-\Delta t}u_0$ for $t \geq 0$ denote the evolution of $u_0 \in \mathcal{V}$ by the diffusion operator. The following properties hold.*

(a) The mass,

$$M(u) := \langle u, \chi_V\rangle_{\mathcal{V}} = \sum_{i \in V} u_i d_i^r, \tag{10}$$

is conserved, i.e., $M(u(t)) = M(u_0)$ for all $t \geq 0$.

(b) $\frac{d}{dt}\|u\|_{\mathcal{V}}^2 = -2\|\nabla u\|_{\mathcal{E}}^2 \leq 0$. In particular, $\|e^{-\Delta t}u_0\|_{\mathcal{V}} \leq \|u_0\|_{\mathcal{V}}$.

(c) Let the mass, M, be defined as in (10), λ_2 be the second eigenvalue of the graph Laplacian, and $\varepsilon > 0$. Assume the graph is connected. If $\tau > \frac{1}{\lambda_2}\log\left(\varepsilon^{-1}\ d_-^{-\frac{r}{2}}\ \|u_0 - (\text{vol } V)^{-1}M\|_{\mathcal{V}}\right)$, then

$$\|u(t) - (\text{vol } V)^{-1}M\|_{\mathcal{V},\infty} \leq \varepsilon, \qquad \text{for all } t > \tau.$$

(d) (Comparison Principle) If, for all $j \in V$, $(u_0)_j \leq (v_0)_j$, then $(e^{-t\Delta}u_0)_j \leq (e^{-t\Delta}v_0)_j$, for all $j \in V$ and $t \geq 0$. In particular, $\|e^{-t\Delta}u_0\|_{\mathcal{V},\infty} \leq \|u_0\|_{\mathcal{V},\infty}$.

If V is connected, the strong comparison principle holds: If, for all $j \in V$, $(u_0)_j \leq (v_0)_j$, and for some $j_0 \in V$, $(u_0)_{j_0} < (v_0)_{j_0}$, then, for all $k \in V$ and $t > 0$, $(e^{-t\Delta}u_0)_k < (e^{-t\Delta}v_0)_k$.

Proof. (a) We compute $\frac{d}{dt}M(u) = \langle \dot{u}, \chi_V\rangle_{\mathcal{V}} = -\langle \Delta u, \chi_V\rangle_{\mathcal{V}} = -\langle \nabla u, \nabla\chi_V\rangle_{\mathcal{E}} = 0$.

(b) We compute $\frac{1}{2}\frac{d}{dt}\|u\|_{\mathcal{V}}^2 = \langle u, \dot{u}\rangle_{\mathcal{V}} = -\langle u, \Delta u\rangle_{\mathcal{V}} = -\langle \nabla u, \nabla u\rangle_{\mathcal{E}} = -\|\nabla u\|_{\mathcal{E}}^2$.

(c) If $\{(\lambda_j, v_j)\}_{j=1}^n$ denote the eigenpairs of the graph Laplacian with $\mathcal{V}$-normalized eigenvectors, then the spectral decomposition of u is given by

$$u(t) = \sum_{j=1}^n e^{-\lambda_j t}\ \langle u_0, v_j\rangle_{\mathcal{V}}\ v_j. \tag{11}$$

Recalling from Lemma 2.4 that $\lambda_1 = 0$ and $v_1 = (\text{vol } V)^{-\frac{1}{2}}\chi_V$ and using (11), we compute

$$\|u - (\text{vol } V)^{-1}M\|_{\mathcal{V}} = \|\sum_{j>1} e^{-\lambda_j t}\langle u_0, v_j\rangle_{\mathcal{V}}\ v_j\|_{\mathcal{V}} \leq e^{-\lambda_2 t}\|u_0 - (\text{vol } V)^{-1}M\|_{\mathcal{V}}.$$

But by Lemma 2.1, this implies

$$\|u - (\text{vol } V)^{-1}M\|_{\mathcal{V},\infty} \leq d_-^{-\frac{r}{2}} e^{-\lambda_2 t}\|u_0 - (\text{vol } V)^{-1}M\|_{\mathcal{V}}.$$

The result now follows, since by Lemma 2.4, $\lambda_2 > 0$ for a connected graph.

(d) If $u_0 \equiv v_0$, then $u(t) \equiv v(t)$, for all $t > 0$, by the uniqueness theorem for ordinary differential equations. In this case there is nothing more to prove. Moreover, by repeating the argument on each connected component we may assume without loss of generality that the entire graph is connected.

Let u_0, v_0 be such that, for all $j \in V$, we have $(u_0)_j \leq (v_0)_j$, and for some $j_0 \in V$, $(u_0)_{j_0} < (v_0)_{j_0}$. We will show that in this case $u_j(t) < v_j(t)$, for every $j \in V$ and all $t > 0$, which proves the strong comparison principle and in particular the comparison principle.

Arguing by contradiction, suppose that $v_j(t) \geq u_j(t)$ for some t and some j. Let t_0 be the last time $v(t)$ lies everywhere above $u(t)$, that is

$$t_0 := \sup\{\ t \geq 0 : \forall\ s \in (0,t) \text{ we have } \forall j\ u_j(s) \leq v_j(s)\}.$$

By our assumption we have that $0 \leq t_0 < \infty$. Then, from the definition of t_0 there is some $k \in V$ such that $u_k(t_0) = v_k(t_0)$ and $\dot{u}_k(t_0) \geq \dot{v}_k(t_0)$. Moreover, again due to the definition of t_0, $u_j(t_0) \leq v_j(t_0)$, for all $j \in V$. This shows that if $u_i(t_0) < v_i(t_0)$ for some neighbor i of k, then $(-\Delta u(t_0))_k > (-\Delta v(t_0))_k$ and

$$0 = \dot{u}_k(t_0) + (\Delta u(t_0))_k < \dot{v}_k(t_0) + (\Delta v(t_0))_k = 0,$$

which is a contradiction. We conclude that $u(t_0) = v(t_0)$ at all neighbors of k, and by iterating the above argument we get in fact that $u(t_0) = v(t_0)$ at all nodes of V, since we are dealing with the case where V is connected. By the uniqueness theorem for ordinary differential equations we conclude that $u_0 = v_0$ at all nodes, which gives a contradiction, and the strong comparison principle is proved. □

Remark 2.6. The dependence of the convergence of $u(t)$ to the steady state $(\text{vol } V)^{-1}M\chi_V$ on the second Laplacian eigenvalue, λ_2, in Lemma 2.5(c), is related to the rate of convergence of a Markov process on a graph to the uniform distribution [SBXD04]. Due to this property, λ_2^{-1} is sometimes referred to as the *mixing-time* for a graph. The second eigenvalue λ_2 is also referred to as the *algebraic connectivity* or *Fiedler value* for a graph [Fie73], and plays an important role in many applications. The robustness of a network to node/edge failures is highly dependent on the algebraic connectivity of the graph. In the "chip-firing game" of Björner, Lovász, and Shor, the algebraic connectivity dictates the length of a terminating game [BLS91]. The algebraic connectivity is also related to the informativeness of a least-squares ranking on a graph [OBO13]. Consequently, algebraic connectivity is a measure of performance for the convergence rate in sensor networks, data fusion, load balancing, and consensus problems [OSFM07].

2.2 Relation between graph Laplacians and balanced graph cuts

In spectral graph theory there are some well known connections between the various graph Laplacians and different normalizations of the graph cut $\text{TV}_{\text{a}}^1(\chi_S) = \sum_{i\in S, j\in S^c} \omega_{ij}$ from (2). For

example, when $r = 0$ (and hence vol $S = |S|$), the quantity $\frac{TV_a^1(\chi_S)}{\min(\text{vol } S, \text{vol } S^c)}$ in Lemma 2.4(d) is the *Cheeger cut* and its minimum over $S \subset V$ is the *Cheeger constant*, *e.g.*, [SM00, SB10]. Let $S_1, \ldots, S_k$ be a partition of all the nodes V, for a given integer k and define the balanced graph cut

$$C_r(S_1, \ldots, S_k) := \sum_{i=1}^{k} \frac{\mathrm{TV}_{\mathrm{a}}^1(\chi_{S_i})}{\text{vol } S_i}.$$

We use the subscript r to remind us that vol S_i depends on r. For $r = 0$ this quantity is known in the literature as the *ratio cut*[10], for $r = 1$ as the *normalized cut* [SM00, vL07]. These quantities are introduced, because minimization of the graph cut, without a balancing term in the denominator, often leads to a partition with many singleton sets, which is typically unwanted in the application at hand. Minimization of this balanced cut over all partitions of V is an NP-complete problem [WW93, SM00], but a relaxation of this problem can be defined using the graph Laplacian. For example, in [vL07, Section 5] it is shown that

$$C_r(S_1, \ldots, S_k) = \mathrm{Tr}(H^T L H),$$

where H is an n by k matrix with elements

$$h_{ij} := \begin{cases} (\text{vol } S_j)^{-\frac{1}{2}} & \text{if } i \in S_j, \\ 0 & \text{else.} \end{cases} \tag{12}$$

Note that H is orthonormal in the $\mathcal{V}$ inner product, *i.e.*, $H^T D^r H = I$, where I is the k by k identity matrix. Hence the minimization of C_r over all partitions, is equivalent to minimizing $\mathrm{Tr}(H^T L H)$ over all n by k $\mathcal{V}$-orthonormal matrices of the form as in (12). The relaxation of this NP-complete minimization problem is now formulated by dropping the condition (12) and minimizing over all n by k $\mathcal{V}$-orthonormal matrices. The problem then becomes an eigenvalue problem and the first k eigenvectors of L are expected to be approximations to the indicator functions of the optimal partition $S_1, \ldots, S_k$. There is often no guarantee of the quality of this approximation though [ST96, GM98, KVV00, KVV04, ST07, vL07].

In order to turn the approximations of the indicator functions into true indicator functions a method like thresholding or k-means clustering is often used [NJW02]. For partitioning the nodes into two subsets ($k = 2$), the potential term in the Allen-Cahn equation (discussed in Section 5) can be interpreted as a nonlinear extension of the graph Laplacian eigenvalue problem which forces the approximate solutions to be close to indicator functions. In this light, it is interesting that the graph Ginzburg-Landau functional (with possibly a mass constraint), of which the Allen-Cahn equation is a gradient flow, Γ-converges to the graph cut functional [vGB12].

3 Curvature and mean curvature flow on graphs

3.1 Graph curvature

In the continuum case, the mean curvature is given by (minus) the divergence of the normal vector field on the boundary of the set (see Appendix A.2 and (46)). This normal vector field achieves the supremum in the definition of total variation of a characteristic function (under sufficient smoothness conditions). Hence, we define the normal of a vertex set by using φ^a from (1) which achieves the supremum of the anisotropic total variation.

[10] Confusingly, the ratio cut is sometimes also called *average cut*, and the Cheeger cut $\frac{\mathrm{TV}_{\mathrm{a}}^1(\chi_S)}{\min(|S|,|S^c|)}$ is sometimes called ratio cut.

Definition 3.1. *The normal of the vertex set $S \subset V$ is*

$$\nu_{ij}^S := sgn\big((\nabla \chi_S)_{ij}\big) = \begin{cases} 1 & \text{if } \omega_{ij} > 0, i \in S^c, \text{ and } j \in S, \\ -1 & \text{if } \omega_{ij} > 0, j \in S^c, \text{ and } i \in S, \\ 0 & \text{else.} \end{cases} \tag{13}$$

As in the continuum case, we define the curvature of a set as the divergence of the normal.

Definition 3.2. *The curvature of the vertex set $S \subset V$ at node $i \in V$ is*

$$(\kappa_S^{q,r})_i := (div\, \nu^S)_i = d_i^{-r} \begin{cases} \sum_{j \in S^c} \omega_{ij}^q & \text{if } i \in S, \\ -\sum_{j \in S} \omega_{ij}^q & \text{if } i \in S^c. \end{cases} \tag{14}$$

Recall from (1), that ν_{ij}^S is not uniquely determined on $\{(i,j) \in E \colon (\nabla \chi_S)_{ij} = 0\}$ (*i.e.*, away from the boundary $\partial S \cup \partial(S^c)$, in the sense of Definition 2.3) and hence the value 0 in (13) is a choice corresponding to the extension of the normal field away from the boundary. This ambiguity is irrelevant when $\operatorname{div} \nu^S$ is coupled to the characteristic function χ_S via the $\mathcal{V}$-inner product, as in

$$\mathrm{TV}_{\mathrm{a}}^q(\chi_S) = \langle \kappa_S^{q,r}, \chi_S \rangle_{\mathcal{V}}, \tag{15}$$

but care should be taken when trying to interpret the normal or the curvature outside this setting.

Note that for $q = 1$ and $S \subset V$, $|(\kappa_S^{1,r})_i| \leq d_i^{1-r}$ for all $i \in V$. Also, $\langle \kappa_S^{q,r}, \chi_V \rangle_{\mathcal{V}} = \langle \nu, \operatorname{grad} \chi_V \rangle_{\mathcal{V}} = 0$. The curvature, $\kappa_S^{q,r}$, has the property that it vanishes away from the boundary $\partial S \cup \partial(S^c)$. In particular, we see that the above mentioned ambiguity is also irrelevant for pairings with χ_{S^c} since

$$\langle \kappa_S^{q,r}, \chi_{S^c} \rangle_{\mathcal{V}} = \langle \kappa_S^{q,r}, \chi_V \rangle_{\mathcal{V}} - \langle \kappa_S^{q,r}, \chi_S \rangle_{\mathcal{V}} = -\mathrm{TV}_{\mathrm{a}}^q(\chi_S). \tag{16}$$

Let $S, \hat{S} \subset V$ be two node sets, then (15) implies

$$\begin{aligned} \mathrm{TV}_{\mathrm{a}}^q(\chi_{\hat{S}}) - \mathrm{TV}_{\mathrm{a}}^q(\chi_S) &= \langle \kappa_{\hat{S}}^{q,r}, \chi_{\hat{S}} \rangle_{\mathcal{V}} - \langle \kappa_S^{q,r}, \chi_S \rangle_{\mathcal{V}} \\ &= \langle \kappa_{\hat{S}}^{q,r} + \kappa_S^{q,r}, \chi_{\hat{S}} - \chi_S \rangle_{\mathcal{V}} + \langle \kappa_{\hat{S}}^{q,r}, \chi_S \rangle_{\mathcal{V}} - \langle \kappa_S^{q,r}, \chi_{\hat{S}} \rangle_{\mathcal{V}}. \end{aligned}$$

For the last two terms, we compute

$$\begin{aligned} \langle \kappa_{\hat{S}}^{q,r}, \chi_S \rangle_{\mathcal{V}} - \langle \kappa_S^{q,r}, \chi_{\hat{S}} \rangle_{\mathcal{V}} &= \left[\sum_{i \in S \cap \hat{S}} \sum_{j \in \hat{S}^c} - \sum_{i \in S \setminus \hat{S}} \sum_{j \in \hat{S}} - \sum_{i \in \hat{S} \cap S} \sum_{j \in S^c} + \sum_{i \in \hat{S} \setminus S} \sum_{j \in S} \right] \omega_{ij}^q \\ &= \left[- \sum_{i \in S \setminus \hat{S}} \sum_{j \in \hat{S}} + \sum_{i \in S} \sum_{j \in \hat{S}} - \sum_{i \in S} \sum_{j \in \hat{S} \cap S} - \sum_{i \in \hat{S} \cap S} \left(\sum_{j \in S^c} - \sum_{j \in \hat{S}^c} \right) \right] \omega_{ij}^q \\ &= \sum_{i \in \hat{S} \cap S} \left[\sum_{j \in \hat{S}} - \sum_{j \in S} - \sum_{j \in S^c} + \sum_{j \in \hat{S}^c} \right] \omega_{ij}^q = \sum_{i \in \hat{S} \cap S} \left[\sum_{j \in V} - \sum_{j \in V} \right] \omega_{ij}^q = 0, \end{aligned}$$

and thus

$$\mathrm{TV}_{\mathrm{a}}^q(\chi_{\hat{S}}) - \mathrm{TV}_{\mathrm{a}}^q(\chi_S) = \langle \kappa_{\hat{S}}^{q,r} + \kappa_S^{q,r}, \chi_{\hat{S}} - \chi_S \rangle_{\mathcal{V}}. \tag{17}$$

In particular, if $\hat{S} = S \setminus \{n\}$ for a node $n \in S$, then

$$\mathrm{TV}^q_{\mathrm{a}}(\chi_{S\setminus\{n\}}) - \mathrm{TV}^q_{\mathrm{a}}(\chi_S) = -\langle \kappa^{q,r}_{S\setminus\{n\}} + \kappa^{q,r}_S, \chi_{\{n\}}\rangle_{\mathcal{V}} = \sum_{i\in S} \omega^q_{in} - \omega^q_{nn} - \sum_{j\in S^c} \omega^q_{nj}$$
$$= \left(\sum_{j\in S} - \sum_{j\in S^c}\right) \omega^q_{nj}.$$

Because we assume there are no self-loops, in the final equality $\omega_{nn} = 0$. A similar computation shows for $n \in S^c$

$$\mathrm{TV}^q_{\mathrm{a}}(\chi_{S\cup\{n\}}) - \mathrm{TV}^q_{\mathrm{a}}(\chi_S) = \left(\sum_{j\in S^c} - \sum_{j\in S}\right) \omega^q_{nj} - \omega^q_{nn} = \left(\sum_{j\in S^c} - \sum_{j\in S}\right) \omega^q_{nj}.$$

The preceding discussion implies the following: if $\Omega \subset V$ is such that S minimizes $\mathrm{TV}^q_{\mathrm{a}}(\chi_S)$ among all sets $S' \subset V$ such that $S \Delta S' \subset \Omega$, then we have that

$$\begin{cases} \left(\sum_{j\in S} - \sum_{j\in S^c}\right) \omega^q_{nj} \le 0 & \text{if } n \in S^c \cap \Omega, \\ \left(\sum_{j\in S} - \sum_{j\in S^c}\right) \omega^q_{nj} \ge 0 & \text{if } n \in S \cap \Omega, \end{cases} \tag{18}$$

(compare with the nonlocal mean curvature, (23)). Here, Ω is a set where S and S' are forced to agree (similar to enforcing a boundary condition in the continuum case). The two inequality conditions in (18) are opposite for the two sides of the interface between S and S^c. This strengthens the heuristic idea that the 'real' interface, where there would be an equality condition, is lost due to the lower bound on the accessible length scales on a graph.

Remark 3.3. It is interesting to make some connections between the graph total variation TV_{a} and graph curvature $\kappa^{1,r}_S$ on the one hand, and the local clustering coefficient [WS98] on the other. Assume G is unweighted (and undirected, as per usual in this paper), then the clustering coefficient C_i of node i, is the number of triangles node i is part of, divided by the number of possible triangles in the neighborhood of i, *i.e.*, $C_i = \frac{2|T_i|}{d_i(d_i-1)}$, where

$$T_i := \{\{i,j,k\} : (i,j),(i,k),(k,i) \in E\}.$$

(A version for weighted graphs was introduced in [BBPSV04, Formula 5].) Using (2) and (16), we can rewrite, for $r = 1$,

$$C_i = \frac{1}{d_i(d_i-1)} \sum_{j,h\in\mathcal{N}_i} \omega_{jh} = \frac{1}{d_i(d_i-1)} \left(\sum_{j\in V}\sum_{h\in\mathcal{N}_i} \omega_{jh} - \sum_{j\in\mathcal{N}_i^c}\sum_{h\in\mathcal{N}_i} \omega_{jh}\right)$$
$$= \frac{1}{d_i(d_i-1)} \left(\sum_{h\in\mathcal{N}_i} d_h - \mathrm{TV}_{\mathrm{a}}(\chi_{\mathcal{N}_i})\right) = \frac{1}{d_i(d_i-1)} \left(\mathrm{vol}\ \mathcal{N}_i - \mathrm{TV}_{\mathrm{a}}(\chi_{\mathcal{N}_i})\right)$$
$$= \frac{1}{d_i(d_i-1)} \langle \chi_{\mathcal{N}_i}, \chi_{\mathcal{N}_i} + \kappa^{1,1}_{\mathcal{N}_i^c}\rangle_{\mathcal{V}}.$$

As for the continuum case, we can arrive at the graph curvature, $\kappa^{q,r}_S$, in several ways. We discuss some of them below. However, the analogy with the continuum curvature becomes even more apparent if instead of the standard mean curvature, we consider the nonlocal mean curvature (see Section 3.2).

If we (formally) compute the first variation of the continuum $\mathrm{TV}(u)$ over all functions $u \in BV$ of *bounded variation* (and not restricting ourselves to characteristic functions only), we find $\operatorname{div} \frac{\nabla u}{|\nabla u|}$, the curvature of the level sets of u. In (28) and Appendix C we follow a similar procedure on the graph and again find $\operatorname{div} \varphi^a$, with φ^a from (1).

Similarly, in the continuum case, an alternative definition of the mean curvature is $\operatorname{div}\left(\frac{\nabla \chi_S}{|\nabla \chi_S|}\right)$ which is a Radon measure defined on the boundary. However, $|\nabla \chi_S| = 1$ on the boundary of S, so that the mean curvature is simply $\Delta \chi_S$. As long as S is a rectifiable set, this computation can be made rigorous in the context of BV functions (see [EG92, Chapter 5]). Computing the analogous quantity on a graph, we find

$$(\Delta \chi_S)_i = d_i^{-r} \begin{cases} \sum\limits_{j \in S^c} \omega_{ij} & \text{if } i \in S, \\ -\sum\limits_{j \in S} \omega_{ij} & \text{if } i \in S^c. \end{cases} \tag{19}$$

This is equal to $\kappa_S^{1,r}$ (compare with (3)). The choice $q = 1$ is a natural one, because it corresponds to the Γ-limit of the graph Ginzburg-Landau functional (GL_ε) [vGB12], whose definition is given in Section 5.

In the continuum case the mean curvature $\kappa(x)$ at the point $x \in \partial S$ in the boundary of a set $S \in \mathbb{R}^d$ satisfies the property that, if S is smooth enough, then for any given ball $B_\delta(x)$ of radius δ and center $x \in \partial S$,

$$|B_\delta(x) \cap S| - \tfrac{1}{2}|B_\delta(x)| = \kappa \delta^2 |B_\delta(x)| + o(\delta^2 |B_\delta(x)|).$$

Note that if S were a half space, then the expression on the right is zero for all δ, since ∂S would separate each $B_\delta(x)$ in sets of equal volume. Thus κ measures how much ∂S deviates from cutting $B_\delta(x)$ in sets of equal volume. The analogous computation on a graph, replacing the ball by the set $\overline{\mathcal{N}_i} := \mathcal{N}_i \cup \{i\}$ of neighbors of node i together with node $\{i\}$, gives, for $S \subset V$,

$$\operatorname{vol}\,(\overline{\mathcal{N}_i} \cap S) - \tfrac{1}{2}\operatorname{vol}\,\overline{\mathcal{N}_i} = \sum_{j \in \overline{\mathcal{N}_i} \cap S} d_j^r - \tfrac{1}{2} \sum_{j \in \overline{\mathcal{N}_i}} d_j^r = \tfrac{1}{2} \left(\sum_{j \in \overline{\mathcal{N}_i} \cap S} - \sum_{j \in \overline{\mathcal{N}_i} \cap S^c} \right) d_j^r.$$

Note that if $r = 0$ and G is an unweighted graph, such that $\omega_{ij} \in \{0, 1\}$, then

$$\operatorname{vol}\,(\overline{\mathcal{N}_i} \cap S) - \tfrac{1}{2}\operatorname{vol}\,\overline{\mathcal{N}_i} = \tfrac{1}{2} \left(\sum_{j \in \overline{\mathcal{N}_i} \cap S} - \sum_{j \in \overline{\mathcal{N}_i} \cap S^c} \right) \omega_{ij}^q = \tfrac{1}{2} \left(\sum_{j \in S} - \sum_{j \in S^c} \right) \omega_{ij}^q,$$

in whose right hand side we recognize (18). In the last equality we used that $\omega_{ij} = 0$, if $j \notin \mathcal{N}_i$.

The equality in (19) has an interesting consequence on the level of Γ-convergence of functionals. For more information about the theory of Γ-convergence, we refer to [DM93, Bra02]. The first result of Γ-convergence for the Ginzburg-Landau functional goes back to work of Modica and Mortola [MM77, Mod87].

Theorem 3.4. *Let $g \in C(\mathbb{R}^n)$, $\varepsilon > 0$, $\tilde{W} \in C^2(\mathbb{R})$ a nonnegative double well potential with wells at 0 and 1, and consider the functionals $f_\varepsilon, f_0 : \mathcal{V} \to \mathbb{R}$, defined by*

$$f_\varepsilon(u) := \sum_{i \in V} g\big((\Delta u)_i\big) + \frac{1}{\varepsilon} \sum_{i \in V} \tilde{W}(u_i),$$

$$f_0(u) := \begin{cases} \sum_{i \in V} g\big((\kappa_S^{1,r})_i\big) & \text{if } u = \chi_S \text{ for some } S \subset V, \\ +\infty & \text{else.} \end{cases}$$

Then $f_\varepsilon \xrightarrow{\Gamma} f_0$ *as* $\varepsilon \to 0$ *(using any of the equivalent metrics on* $\mathbb{R}^n$*).*

Furthermore, if the double well potential $\tilde{W}$ *satisfies a coercivity condition —i.e., there exists a* $c > 0$ *such that for large* $|u|$, $\tilde{W}(u) \leq c(u^2 - 1)$*— then compactness holds, in the following sense: Let* $\{\varepsilon_n\}_{n=1}^{\infty} \subset \mathbb{R}_+$ *be such that* $\varepsilon_n \to 0$ *as* $n \to \infty$*, and let* $\{u_n\}_{n=1}^{\infty}$ *be a sequence such that there exists a* $C > 0$ *such that for all* $n \in \mathbb{N}$ $f_{\varepsilon_n}(u_n) < C$*. Then there exists a subsequence* $\{u_{n'}\}_{n?=1}^{\infty} \subset \{u_n\}_{n=1}^{\infty}$ *and a* u_∞ *of the form* $u_\infty = \chi_S$*, for some* $S \subset V$*, such that* $u_{n'} \to u_\infty$ *as* $n \to \infty$*.*

Proof. The key point in the proof of the Γ-convergence is to note that f_ε is a continuous perturbation of the functional $w_\varepsilon : \mathcal{V} \to \mathbb{R}$,

$$w_\varepsilon(u) := \frac{1}{\varepsilon} \sum_{i \in V} \tilde{W}(u_i).$$

By [vGB12, Lemma 3.3][11] $w_\varepsilon \xrightarrow{\Gamma} w_0$ as $\varepsilon \to 0$, where

$$w_0(u) := \begin{cases} 0 & \text{if } u = \chi_S \text{ for some } S \subset V, \\ +\infty & \text{else.} \end{cases}$$

By a well known property of Γ-convergence [DM93, Proposition 6.21], the Γ-limit is preserved under continuous perturbations. Then using the fact, shown above in (19), that $\Delta u = \kappa_S^{1,r}$ if $u = \chi_S$, completes the proof of Γ-convergence.

The compactness result is a direct adaptation of the proof of [vGB12, Theorem 3.2] to the current functionals f_ε. □

Remark 3.5. Note that in Theorem 3.4 above, we can also use the double well potential W with wells at ± 1, instead of $\tilde{W}$. In that case, the limit functional w_0 in the proof takes finite values only for functions of the form $u = \chi_S - \chi_{S^c}$, for some $S \subset V$. Because $\Delta(\chi_S + \chi_{S^c}) = \Delta\chi_V = 0$, we have $\Delta(\chi_S - \chi_{S^c}) = 2\kappa_S^{1,r}$ and hence the limit functional f_0 takes the form

$$f_0(u) := \begin{cases} \sum_{i \in V} g\big(2(\kappa_S^{1,r})_i\big) & \text{if } u = \chi_S \text{ for some } S \subset V, \\ +\infty & \text{else.} \end{cases}$$

We end this subsection with another similarity between the graph based objects we introduced and their continuum counterparts. The gradient of the graph distance $d^{\partial S}$, from Definition 2.2, agrees with the normal ν, from (13), on the boundary of S induced by the graph distance, in the sense of the following lemma. This again corresponds to what we expect based on the continuum case. We define the signed distance to ∂S as

$$sd^{\partial S} := (\chi_{S_n^c} - \chi_{S_n}) d^{\partial S}. \tag{20}$$

Lemma 3.6. *Let* $S \subset V$*. Define the exterior boundary of* S *induced by the graph distance* $d^{\partial S}$ *as*

$$\partial_{ext} S := \{i \in S^c : \exists j \in \partial S \text{ such that } d_i^{\partial S} = \omega_{ij}^{q-1}\}.$$

Let $i \in \partial_{ext} S$*, then there is a* $j \in \partial S$ *such that* $(\nabla sd^{\partial S})_{ij} = -\nu_{ij}$*.*

Similarly, let the interior boundary of S *induced by the graph distance be*

$$\partial_{int} S := \{i \in S : \exists j \in \partial S \text{ such that } d_i^{\partial S} = \omega_{ij}^{q-1}\}.$$

If $i \in \partial_{int} S$*, then there is a* $j \in \partial S$ *such that* $(\nabla sd^{\partial S})_{ij} = \nu_{ij}$*.*

[11] Note that in the statement and proof of Lemma 3.3 in [vGB12], it says g_ε twice where w_ε is meant.

Proof. First we note that, for $i \in \partial_{ext} S \subset S^c$, we have $s d_i^{\partial S} = d_i^{\partial S}$. Because $d^{\partial S}$ satisfies equation (13) (with S replaced by ∂S), we have

$$\min_{k \in \mathcal{N}_i} (\nabla d^{\partial S})_{ik} = \min_{k \in \mathcal{N}_i} \omega_{ik}^{1-q} (d_k^{\partial S} - d_i^{\partial S}) = -1.$$

Note that $\partial_{ext} S \subset \partial(S^c)$. Hence, $i \in \partial(S^c)$ and thus there is a $k \in \mathcal{N}_j$ such that $k \in \partial S$ and therefore $d_k^{\partial S} = 0$. Because $d^{\partial S}$ is nonnegative and $i \in \partial_{ext} S$, we deduce

$$1 = \max_{k \in \mathcal{N}_i} \omega_{ik}^{1-q} d_i^{\partial S} = \max_{k \in \mathcal{N}_i} \omega_{ik}^{1-q} \omega_{ij}^{q-1}.$$

Thus the maximum is achieved for $k = j$, and hence so is the minimum in (13), which shows that

$$(\nabla d^{\partial S})_{ij} = -1 = -\nu_{ij}.$$

The proof for $i \in \partial_{int} S$ follows from similar arguments, noting that $s d_i^{\partial S} = -d^{\partial S}$. □

Note that $\partial_{ext} S \subset \partial(S^c)$, but equality does not necessarily hold. If a shortest path from $i \in \partial(S^c)$ to ∂S does not equal $\{i, j\}$, for some $j \in \partial S$, then $i \notin \partial_{ext} S$. This situation does not occur if the graph distances are consistent, in the following sense: if, for all $i, j, k \in V$, $\omega_{ij}^{q-1} \leq \omega_{ik}^{q-1} + \omega_{kj}^{q-1}$. In that case, $\partial_{ext} S = \partial(S^c)$.

3.2 Relation with the continuum nonlocal mean curvature

There is a clear analogy between the expressions in (18) and the (continuum) *nonlocal mean curvature* [CS10, CRS10], as well as between $\mathrm{TV}_{\mathrm{a}}^q$ and continuum nonlocal energy functionals.

Consider an interaction kernel $K : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ with $K(x, y) = K(y, x)$ and

$$\sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} \min\{1, |x - y|^2\} K(x, y)\, dy < +\infty.$$

This kernel K can be thought of as the energy given by a long-range interaction between a particle placed at x with a particle at y. It defines a functional on subsets $S \subset \mathbb{R}^n$, sometimes called "nonlocal perimeter" or "nonlocal energy" and it is given by

$$J_K(S) = L_K(S, S^c),$$

where for any pair $A, B \subset \mathbb{R}^n$ we write

$$L_K(A, B) = \int_A \int_B K(x, y)\, dxdy. \tag{21}$$

Compare this with the graph case, if, for $A, B \subset V$, we write

$$L_G(A, B) = \sum_{i \in A} \sum_{j \in B} \omega_{ij}^q. \tag{22}$$

In terms of this bilinear functional, the anisotropic total variation, defined in (2), can be rewritten

$$\mathrm{TV}_{\mathrm{a}}^q(\chi_S) = L_G(S, S^c).$$

We see that (21) is nothing but the continuum version of (22), and one may rightfully interpret the weight matrix ω_{ij}^q as an interaction kernel between pairs of nodes in G and $L_G(S, S^c)$ as measuring the total "interction energy" between S and S^c.

Now suppose that $S \subset \mathbb{R}^n$ minimizes $J_K(S)$ in some domain Ω, meaning that if S' is such that $S \Delta S' \subset\subset \Omega$ then $J_K(S) \leq J_K(S')$. In this case one can see that the following two conditions must hold

$$\begin{cases} L(A,S) - L(A, S^c \setminus A) & \leq 0, \quad \forall\, A \subset S^c \cap \Omega, \\ L(A, S \setminus A) - L(A, S^c) & \geq 0, \quad \forall\, A \subset S \cap \Omega. \end{cases} \tag{23}$$

If, arguing heuristically, we let A shrink down to any $x \in (\partial S) \cap \Omega$, we find that

$$\int_{\mathbb{R}^n} (\chi_S(y) - \chi_{S^c}(y))\, K(x,y)\, dx = 0.$$

The integral on the left, which is well defined in the principal value sense[12] when $x \in \partial S$ and ∂S is smooth enough (C^2 suffices), is known as the nonlocal mean curvature of S at x with respect to K, or just nonlocal mean curvature of S at x, when K is clear from the context. As with J_K and $\mathrm{TV}_{\mathrm{a}}^q$, we see that

$$\kappa_{\mathrm{nonlocal}}(x) := \int_{\mathbb{R}^n} (\chi_S(y) - \chi_{S^c}(y))\, K(x,y)\, dy$$

is exactly a continuum analogue of the quantity $\left(\sum_{j\in S} - \sum_{j \in S^c}\right) \omega_{nj}^q$ in (18), moreover, the inequalities in (23) are a continuum analogue of those in (18). Note however, that $\left(\sum_{j\in S} - \sum_{j \in S^c}\right) \omega_{nj}^q$ is defined for all of $n \in V$, whereas $\kappa_{\mathrm{nonlocal}}(x)$ above is only defined when $x \in \partial S$ and ∂S is smooth enough.

Known regularity results deal mostly with the case $K_s(x,y) := c_{n,s}|x-y|^{-n-s}$ for $s \in (0,1)$ [CRS10, CG10]. It is worth noting that J_{K_s} is a fractional Sobolev norm of the characteristic function of S

$$J_{K_s}(S) = \tfrac{1}{2} \|\chi_S\|^2_{\dot{H}^{s/2}},$$

where $\|.\|_{\dot{H}^s/2}$ is defined in terms of the Fourier transform of f by

$$\|f\|_{\dot{H}^{s/2}} = \| \, |\xi|^s \hat{f}(\xi) \, \|_{L^2(\mathbb{R}^n)}.$$

Moreover, as $s \to 1^-$ the quantity above gives the perimeter of S, and the corresponding nonlocal mean curvature converges pointwise to the standard mean curvature. In [CS10] it is shown that if we consider the MBO scheme, where instead of the heat equation we use the fractional heat equation, then in the limit we get a set S_t evolving over time with a normal velocity at $x \in \partial S_t$ given by

$$V(x) = c_{n,s} \int_{\mathbb{R}^n} \frac{\chi_S(y) - \chi_{S^c}(y)}{|x-y|^{n+s}}\, dy. \tag{24}$$

3.3 Mean curvature flow

In this section, we define a mean curvature flow on graphs and connect it with the curvature $\kappa_S^{q,r}$ in (14). It is not clear what is the most natural notion for the evolution of a phase in a graph. Do we want to consider a sequence of subsets $\{S_n\}_{n\in\mathbb{N}}$, or a continuous family $\{S_t\}_{t>0}$ which, although piecewise constant in t, may change in arbitrarily small time intervals? How we connect solutions of the graph mean curvature flow to solutions of the graph (MBO_τ) scheme

[12] Note that $K(x,y)$ could have a very strong singularity at $x = y$ making the integral diverge, if taken as a Lebesgue integral. The boundedness of the principal value of this singular integral tells us that ∂S must have some smoothness near x.

from Section 4, or to solutions of the graph Allen-Cahn equation (ACE_ε) from Section 5, will depend on the answer to this question. For now, we shall be content with considering a phase evolution comprised of a discrete sequence of sets $S_n = S(n\eth t)$, $n \in \mathbb{N}$, that correspond to the state of the system at discrete time steps.

Our construction follows the well-known variational formulations for classical mean curvature flow [ATW93, LS95]. Appendix A.2 has a brief overview of mean curvature and the associated flow in the continuum case. An obstacle one encounters when trying to emulate the continuum level set method to express mean curvature flow on graphs is, that, due to the lack of a discrete chain rule, the resulting equation is not independent on the choice of level set function.

Recall the notions of graph distance and boundary of a node set from Section 2.

Definition 3.7. *The mean curvature flow, $S_n = S(n\eth t)$, with discrete time step $\eth t$ for an initial set $S_0 \subset V$, is defined*

$$S_{n+1} \in \arg\min_{\hat{S}\subset V} \mathcal{F}(\hat{S}, S_n), \tag{$\text{MCF}_{\eth t}$}$$

where

$$\mathcal{F}(\hat{S}, S_n) := TV_a^q(\chi_{\hat{S}}) - TV_a^q(\chi_{S_n}) + \frac{1}{\eth t}\langle \chi_{\hat{S}} - \chi_{S_n}, (\chi_{\hat{S}} - \chi_{S_n}) d^{\Sigma_n}\rangle_{\mathcal{V}} \tag{25}$$

and

$$\Sigma_n := \partial S_n \cup \partial(S_n^c) = \{i \in V : \exists\, (i,j) \in E \text{ such that } (i \in S_n \wedge j \in S_n^c) \vee (i \in S_n^c \wedge j \in S_n)\}.$$

Note that, for a given graph G, minimizers of $\mathcal{F}$ may not be unique. In this case different mean curvature flows can be defined, depending on the choice of S_{n+1}. An example of this non-uniqueness on the 4-regular graph is given in Section 6.5.

We choose to use the distance to Σ_n, instead of the distance to either ∂S_n or $\partial(S_n^c)$, so that the mean curvature flow is not *a priori* (independent of curvature) biased to either adding nodes to or removing nodes from S.

Since nodes in Σ_n can be added or removed from S without increasing the last term of (25), every stationary state χ_S of ($\text{MCF}_{\eth t}$) is a minimal surface in the sense that $\text{TV}_{\text{a}}^q(\chi_S) \leq \text{TV}_{\text{a}}^q(\chi_{\{S\cup\{n\}}})$ for $n \in \partial(S^c)$ and $\text{TV}_{\text{a}}^q(\chi_S) \leq \text{TV}_{\text{a}}^q(\chi_{\{S\setminus\{n\}}})$ for $n \in \partial S$ (in the case where the minimizer of $\mathcal{F}$ is unique, the inequalities are strict). In particular, the sequence defined by ($\text{MCF}_{\eth t}$) is not "frozen" for $\eth t$ arbitrarily small.

Remark 3.8. In the last term of $\mathcal{F}(\hat{S}, S_n)$, we use a symmetrized distance to the boundary, $\langle \chi_{\hat{S}} - \chi_{S_n}, (\chi_{\hat{S}} - \chi_{S_n}) d^{\Sigma_n}\rangle_{\mathcal{V}}$. Other choices are possible here, including $\|\chi_{S^c} d^S - \chi_S d^{S^c}\|^2$. For this alternative choice, ($\text{MCF}_{\eth t}$) exhibits "freezing".

We can rewrite the last term in $\mathcal{F}(\hat{S}, S_n)$ in terms of the signed graph distance

$$sd^{\Sigma_n} := (\chi_{S_n^c} - \chi_{S_n}) d^{\Sigma_n},$$

(compare with (20)), which takes nonnegative values in S_n^c and nonpositive values in S_n. We state the precise result in the following lemma.

Lemma 3.9. $\operatorname{argmin}_{\hat{S}\subset V} \mathcal{F}(\hat{S}, S_n) = \operatorname{argmin}_{\hat{S}\subset V} \mathcal{F}'(\hat{S}, S_n)$, *where*

$$\begin{aligned} \mathcal{F}'(\hat{S}, S_n) &:= TV_a^q(\chi_{\hat{S}}) - TV_a^q(\chi_{S_n}) + \frac{1}{\eth t}\langle \chi_{\hat{S}}, sd^{\Sigma_n}\rangle_{\mathcal{V}} \\ &= \langle \kappa_{\hat{S}}^{q,r} + \kappa_{S_n}^{q,r}, \chi_{\hat{S}} - \chi_{S_n}\rangle_{\mathcal{V}} + \frac{1}{\eth t}\langle \chi_{\hat{S}}, sd^{\Sigma_n}\rangle_{\mathcal{V}}. \end{aligned} \tag{26}$$

Proof. The rewriting of $\mathrm{TV}^q_{\mathrm{a}}(\chi_{\hat{S}}) - \mathrm{TV}^q_{\mathrm{a}}(\chi_{S_n})$ in terms of the curvatures, follows directly from (17). For the distance term, we compute

$$\begin{aligned}\langle(\chi_{\hat{S}} - \chi_{S_n})^2, d^{\Sigma_n}\rangle_{\mathcal{V}} &= \langle\chi_{\hat{S}}(1 - 2\chi_{S_n}) + \chi_{S_n}, d^{\Sigma_n}\rangle_{\mathcal{V}} \\ &= \langle\chi_{\hat{S}}, (\chi_{S_n^c} - \chi_{\hat{S}})d^{\Sigma_n}\rangle_{\mathcal{V}} + \langle\chi_{S_n}, d^{\Sigma_n}\rangle_{\mathcal{V}},\end{aligned}$$

where in the last line we used that $1 - \chi_{S_n} = \chi_{S_n^c}$. The proof is completed by noting that the last term above does not depend on $\hat{S}$. □

Theorem 3.10. *Let $u \in \mathcal{V}$ be a minimizer of the convex functional*

$$F(u) := TV^q_a(u) + \frac{1}{\eth t}\langle u, sd^{\Sigma_n}\rangle_{\mathcal{V}}, \tag{27}$$

then, for a.e. $s \in \mathbb{R}$, the superlevel set

$$E(s) := \{i \in V : u_i > s\}$$

is a minimizer $\hat{S}$ of $\mathcal{F}(\cdot, S_n)$ from (25).

Proof. In Lemma B.1 in Appendix B we show that

$$\mathrm{TV}^q_{\mathrm{a}}(u) = \frac{1}{2}\sum_{i,j\in V}\omega^q_{ij}|u_i - u_j| = \frac{1}{2}\int_{\mathbb{R}}\sum_{i,j\in V}\omega^q_{ij}\left|\left(\chi_{E(s)}\right)_i - \left(\chi_{E(s)}\right)_j\right| ds$$

Writing $u_i = \int_{\mathbb{R}}\left(\chi_{E(s)}\right)_i ds$, as in (48), we also see that

$$\langle u, sd^{\Sigma_n}\rangle_{\mathcal{V}} = \int_{\mathbb{R}}\langle\chi_{E(s)}, sd^{\Sigma_n}\rangle_{\mathcal{V}}\, ds.$$

This gives

$$F(u) = \int_{\mathbb{R}}\left[\mathrm{TV}^q_{\mathrm{a}}(\chi_{E(s)}) + \frac{1}{\eth t}\langle\chi_{E(s)}, sd^{\Sigma_n}\rangle_{\mathcal{V}}\right] ds = \int_{\mathbb{R}}\left[\mathcal{F}'(E(s), S_n) - \mathrm{TV}^q_{\mathrm{a}}(S_n)\right] ds,$$

where $\mathcal{F}'$ is as in (26). Hence, if u minimizes F, then a.e. superlevel set $E(s)$ minimizes $\mathcal{F}'(\cdot, S_n)$. Lemma 3.9 now completes the proof. □

Remark 3.11. The function $\mathrm{TV}^q_{\mathrm{a}}(u)$ is a convex function, thus, taking $\partial\mathrm{TV}^q_{\mathrm{a}}(u)$ as the (possibly multivalued) subdifferential of $\mathrm{TV}^q_{\mathrm{a}}(u)$ [ET76], any minimizer of

$$u \mapsto \mathrm{TV}^q_{\mathrm{a}}(u) + \langle u, g\rangle_{\mathcal{V}},$$

for $g \in \mathcal{V}$, will solve the differential inclusion

$$\partial\mathrm{TV}^q_{\mathrm{a}}(u) \ni -g.$$

From the form definition of $\mathrm{TV}^q_{\mathrm{a}}$ we see that its subdifferential is only multivalued for u's such that ∇u vanishes at some node, at all other u, $\mathrm{TV}^q_{\mathrm{a}}$ is pointwise differentiable. In particular, if $u, v \in \mathcal{V}$ and ∇u is never zero, we may differentiate[13]

$$\begin{aligned}\frac{d}{dt}_{|t=0}\mathrm{TV}^q_{\mathrm{a}}(u + tv) &= \frac{d}{dt}_{|t=0}\left(\tfrac{1}{2}\sum_{i,j}\omega^q_{ij}|u_i - u_j + t(v_i - v_j)|\right), \\ &= \tfrac{1}{2}\sum_{ij}\omega^q_{ij}\mathrm{sgn}(u_i - u_j)(v_i - v_j), \\ &= \langle\mathrm{sgn}(\nabla u), \nabla v\rangle.\end{aligned} \tag{28}$$

[13]For further discussion and a generalization of this computation, see Appendix C.

Since div is the adjoint of ∇, it follows that, for all $v \in \mathcal{V}$,

$$\frac{d}{dt}\left(\mathrm{TV}_{\mathrm{a}}^{q}(u+tv) + \langle u+tv, g\rangle_{\mathcal{V}}\right) = \langle \operatorname{div}(\operatorname{sgn}(\nabla u)) + g, v\rangle_{\mathcal{V}}.$$

Therefore, the Euler-Lagrange equation for minimizers of F from (27), is

$$\operatorname{div}\operatorname{sgn}(\nabla u) + \frac{1}{\mathfrak{d}t} sd^{\Sigma_n} = 0,$$

provided ∇u is never zero. Whenever $(\nabla u)_{ij} = 0$ for some i, j, the above equation is replaced by a differential inclusion in terms of the subdifferential of the absolute value function. Concretely, since the subdifferential of the absolute value function at 0 is the interval $[-1, 1]$, there exists $\phi \in \mathcal{V}$ such that $|\phi_i| \leq 1$ for every node $i \in V$, and

$$\operatorname{div}\phi + \frac{1}{\mathfrak{d}t} sd^{\Sigma_n} = 0.$$

Fast computational methods for the solution of ($\mathrm{MCF}_{\mathfrak{d}t}$) based on max flow/min cut algorithms are developed in [CD09]. These methods exploit the homogeneity and submodularity of the total variational functional, $\mathrm{TV}_{\mathrm{a}}^{q}$.

4 Threshold dynamics on graphs

In this section we study the threshold dynamics or Merriman-Bence-Osher algorithm on a graph G. For a short overview of the continuum case we refer to Section A.3 in Appendix A.

4.1 The graph MBO algorithm

The MBO scheme on a graph, describing the evolution of a node subset $S \subset V$, is given as follows.

Algorithm ($\mathbf{MBO}_\tau$): The Merriman-Bence-Osher algorithm on a graph.

Data: An initial node subset $S_0 \subset V$, a time step $\tau > 0$, and the number of time steps $N > 0$.

Output: A sequence of node sets $\{S_k\}_{k=1}^{N}$, which is the (MBO_τ) evolution of S_0.

for $k = 1$ **to** N, **do**

Diffusion step. Let $v = e^{-\Delta\tau}\chi_{S_{k-1}}$ denote the solution at time τ of the initial value problem

$$\dot{v} = -\Delta v, \quad v(0) = \chi_{S_{k-1}}. \tag{29}$$

Here χ_S denotes the characteristic function of the set S.

Threshold step. Define the set $S_k \subset V$ to be

$$S_k = \{i \in V : v_i \geq \frac{1}{2}\}.$$

By the comparison principle, Lemma 2.5(d), we note that the solution to (29) satisfies $v(t) \in [0, 1]^n$ for all $t \in [0, \tau]$.

Remark 4.1. In the thresholding step of the (MBO$_\tau$) scheme, we have arbitrarily chosen to include the level set $\{i \in V : v_i = \frac{1}{2}\}$ in the new set S_k, *i.e.*, the value at nodes i for which $v_i = \frac{1}{2}$ is set to 1.

An alternative description of the algorithm is as follows. Let u_k be the indicator of the set S_k as defined by the (MBO$_\tau$) algorithm. If we define the thresholding function $P : \mathbb{R} \to \{0, 1\}$, which acts by thresholding

$$P(x) := \begin{cases} 1 & \text{if } x \geq \frac{1}{2} \\ 0 & \text{if } x < \frac{1}{2} \end{cases},$$

then the iterates can be succinctly written $u_k = (Pe^{-\Delta\tau})^k u_0$.

Several papers use the MBO algorithm on a graph to approximate motion by mean curvature. For example, in [MKB12, GCMB+13, HLPB13], the MBO algorithm on graphs was implemented and used to study data clustering, community detection, segmentation, object recognition, and inpainting. This is accomplished by simply reinterpreting the Laplacian in (47), or in appropriate extensions of (47), as the graph Laplacian.

4.2 The "step-size" τ in the (MBO$_\tau$) algorithm

As discussed at the end of Section A.3, in a finite difference discretization of the continuum MBO algorithm, the time step τ must be chosen carefully to avoid trivial dynamics. For τ too small, there is not enough diffusion to change the value of u at neighboring grid points beyond the threshold value. In this case, the solution is stationary under an (MBO$_\tau$) iteration and we say that the solution is *frozen* or *pinned*. For τ too large there is so much diffusion that a stationary state is reached after one iteration in the MBO scheme. It is not surprising that these finite difference effects also appear for the MBO algorithm on graphs. From the form of the heat solution operator, $e^{\Delta\tau}$, we expect that τ should be roughly chosen in the interval $(\lambda_n^{-1}, \lambda_2^{-1})$. Theorems 4.2 and 4.3 strengthen this intuition.

The following theorem gives a lower bound on the choice of τ to avoid freezing in the (MBO$_\tau$) algorithm on general graphs.

Theorem 4.2. *Let ρ be the spectral radius of the graph Laplacian, Δ. Then the (MBO$_\tau$) iterations on the graph with initial set S are stationary if either of the two conditions are satisfied:*

$$\tau < \tau_\rho(S) := \rho^{-1} \log\left(1 + \frac{1}{2}\, d_-^{\frac{r}{2}}\, (\text{vol } S)^{-\frac{1}{2}}\right) \tag{30}$$

or

$$\tau \leq \tau_\kappa(S) := \frac{1}{2\|\Delta\chi_S\|_{\mathcal{V},\infty}}. \tag{31}$$

In particular, since vol $S > d_-^{\frac{r}{2}}$, *if* $\tau < \log\frac{3}{2} \cdot \rho^{-1} \approx 0.4 \cdot \rho^{-1}$, *then (30) implies the MBO iterates are pinned for any initial $S \subset V$.*

Proof. To prove (30), let χ_S be the characteristic function on a set $S \subset V$. For a node to be added or removed from S by one iteration of (MBO$_\tau$), it is necessary that $\|e^{-\tau\Delta}\chi_S - \chi_S\|_{\mathcal{V},\infty} \geq \frac{1}{2}$. For the linear operator $A\colon \mathcal{V} \to \mathcal{V}$, let $\|A\|_{\mathcal{V}}$ be the operator norm induced by $\|\cdot\|_{\mathcal{V}}$, *i.e.*,

$$\|A\|_{\mathcal{V}} = \max_{u \in \mathcal{V}\setminus\{0\}} \frac{\|Au\|_{\mathcal{V}}}{\|u\|_{\mathcal{V}}}$$

(see also Lemma 2.4(b)). Using Lemma 2.1, we compute

$$\|e^{-\tau\Delta}\chi_S - \chi_S\|_{\mathcal{V},\infty} \le d_-^{-\frac{r}{2}} \|e^{-\tau\Delta}\chi_S - \chi_S\|_{\mathcal{V}} \le d_-^{-\frac{r}{2}} \|e^{-\tau\Delta} - \mathrm{Id}\|_{\mathcal{V}} \sqrt{\mathrm{vol}\ S}.$$

Using the triangle inequality and the submultiplicative property of $\|\cdot\|_{\mathcal{V}}$ (see, *e.g.*, [HJ90]), we compute

$$\|e^{-\tau\Delta} - \mathrm{Id}\|_{\mathcal{V}} \le \sum_{k=1}^{\infty} \frac{1}{k!}(\tau\|\Delta\|_{\mathcal{V}})^k = e^{\rho\tau} - 1.$$

Thus, if $\tau < \rho^{-1}\log\left(1 + \frac{1}{2}\, d_-^{\frac{r}{2}}\ (\mathrm{vol}\ S)^{-\frac{1}{2}}\right)$, all nodes are stationary under an (MBO$_\tau$) iteration.

To prove (31), we write the solution to the heat equation at time τ,

$$u(\tau) = e^{-\tau\Delta}\chi_S = \chi_S - \int_0^\tau \Delta u(t)dt.$$

This implies

$$\|u(\tau) - \chi_S\|_{\mathcal{V},\infty} = \left\|\int_0^\tau \Delta u(t)dt\right\|_{\mathcal{V},\infty} \le \int_0^\tau \left\|e^{-t\Delta}\Delta\chi_S\right\|_{\mathcal{V},\infty} dt \le \tau\,\|\Delta\chi_S\|_{\mathcal{V},\infty}\,.$$

Here, we used the comparison principle, Lemma 2.5(d). Thus, if $\tau \le \frac{1}{2\|\Delta\chi_S\|_{\mathcal{V},\infty}}$, then $\|u(\tau) - \chi_S\|_{\mathcal{V},\infty} \le \frac{1}{2}$, implying the (MBO$_\tau$) solution is stationary. □

The following corollary of Lemma 2.5(c) shows that an upper bound on τ is necessary to avoid trivial dynamics.

Theorem 4.3. *Let the graph be connected. Consider the (MBO$_\tau$) algorithm with initial condition χ_S, for a node set $S \subset V$. Assume $R_S := \frac{\mathrm{vol}S}{\mathrm{vol}V} \neq \frac{1}{2}$. If*

$$\tau > \tau_t := \frac{1}{\lambda_2}\log\left(\frac{(\mathrm{vol}S)^{\frac{1}{2}}\ (\mathrm{vol}S^c)^{\frac{1}{2}}}{(\mathrm{vol}V)^{\frac{1}{2}}\ |R_S - \frac{1}{2}|\ d_-^{\frac{r}{2}}}\right), \tag{32}$$

where the mass $M(u_0)$ is defined in (10), then

$$Pe^{-\tau\Delta}u_0 = \begin{cases} \chi_V & R_S > \frac{1}{2}, \\ 0 & R_S < \frac{1}{2}. \end{cases}$$

Proof. In Lemma 2.5(c), set $\varepsilon = |(\mathrm{vol}\ V)^{-1}M(u_0) - \frac{1}{2}| = |R_S - \frac{1}{2}|$. This implies that $\|u(\tau) - R_S\|_{\mathcal{V},\infty} \le |R_S - \frac{1}{2}|$, as desired. With $\varepsilon = |R_S - \frac{1}{2}|$, the condition on τ in Lemma 2.5(c) is $\tau > \frac{1}{\lambda_2}\log\left(|R_S - \frac{1}{2}|^{-1}\ d_-^{-\frac{r}{2}}\ \|u_0 - R_S\|_{\mathcal{V}}\right)$. For $u_0 = \chi_S$, $\|u_0 - R_S\|_{\mathcal{V}}^2 = \frac{\mathrm{vol}S\ \mathrm{vol}S^c}{\mathrm{vol}V}$. □

This corollary shows that, if τ is chosen too large, one iteration of (MBO$_\tau$) leads to a trivial state $u = \chi_V$ or $u = 0$, which is stationary under the algorithm (MBO$_\tau$).

The following theorem gives a condition for which there is a gap between the lower and upper bound for τ.

Theorem 4.4. *Consider the (MBO$_\tau$) iterations on a graph with $n \ge 2$. Let τ_p and τ_t be defined as in (30) and (32). If $\frac{\lambda_2}{\lambda_n} < \frac{\log\sqrt{2}}{\log\frac{3}{2}} \approx 0.85$, then $\tau_p < \tau_t$.*

Proof. Since $\left|R_S - \frac{1}{2}\right| \leq \frac{1}{2}$, $d^r_- \leq \text{vol}S$, and $(\text{vol}S)\ (\text{vol}S^c) > d^r_-(\text{vol}V - d^r_-)$ we have

$$\tau_p = \lambda_2^{-1} \log\left(1 + \frac{1}{2}\, d_-^{\frac{r}{2}}\ (\text{vol}\ S)^{-\frac{1}{2}}\right) \geq \lambda_2^{-1}\ \log\left(2\sqrt{1 - \frac{d^r_-}{\text{vol}V}}\right) \geq \lambda_2^{-1}\ \log\sqrt{2}.$$

Since $d^r_- \leq \text{vol}S$, we have

$$\tau_t = \lambda_n^{-1}\ \log\left(1 + \frac{1}{2}\sqrt{\frac{d^r_-}{\text{vol}S}}\right) \leq \lambda_n^{-1}\ \log\frac{3}{2}.$$

The result follows. □

Theorem 4.4 further reenforces our intuition that τ should be chosen in the interval $(\lambda_n^{-1}, \lambda_2^{-1})$. If, for a particular graph, this interval is very small, then Theorems 4.2 and 4.3 cannot provide an interval for which the (MBO$_\tau$) iterations has a chance of being non-stationary after the first iteration. Note however that the interval $[\tau_p, \tau_t]$ given by these theorems, is not necessarily a sharp interval for interesting dynamics.

4.3 A Lyapunov functional for the graph (MBO$_\tau$) algorithm

In this section we introduce a functional which is decreasing on iterations of the (MBO$_\tau$) algorithm. The analogous functional for the continuum setting was recently found in [EO13]. The functional is then used to show that the (MBO$_\tau$) algorithm with any initial condition converges to a stationary state in a finite number of iterations.

Let $\tau > 0$ and consider the functional $J\colon \mathcal{V} \to \mathbb{R}$ defined by

$$J(u) = \langle 1 - u, e^{-\tau\Delta}u\rangle_{\mathcal{V}}. \tag{33}$$

Note that by, Lemma 2.5(a), $J(u) = M(u) - \langle u, e^{-\tau\Delta}u\rangle_{\mathcal{V}}$, where M is the mass from (10).

Lemma 4.5. *The functional J defined in (33) has the following elementary properties.*

1. *J is a strictly concave functional on $\mathcal{V}$.*
2. *J is Fréchet differentiable with derivative in the direction v given by*

$$L_u(v) := \left\langle \left.\frac{\delta J}{\delta u}\right|_u, v\right\rangle_{\mathcal{V}}, \quad \text{where} \quad \left.\frac{\delta J}{\delta u}\right|_u = 1 - 2e^{-\tau\Delta}u.$$

Proof. We compute, for all $v \neq 0$,

$$\frac{d^2}{d\alpha^2} J(u + \alpha v) = -2\langle v, e^{-\tau\Delta}v\rangle_{\mathcal{V}} < 0.$$

Taking the first variation of $J(u) = \langle 1 - u, e^{-\tau\Delta}u\rangle_{\mathcal{V}}$, we find that

$$\left\langle \frac{\delta J}{\delta u}, \delta u\right\rangle_{\mathcal{V}} := \left\langle 1 - u, e^{-\tau\Delta}\delta u\right\rangle_{\mathcal{V}} - \left\langle \delta u, e^{-\tau\Delta}u\right\rangle_{\mathcal{V}} = \left\langle 1 - 2e^{-\tau\Delta}u, \delta v\right\rangle_{\mathcal{V}},$$

as desired. □

Define the convex set $\mathcal{K} := \{\phi \in \mathcal{V} \colon \forall j \in V \ \phi_j \in [0,1]\}$. Is it instructive to consider the optimization problem,

$$\min_{u \in \mathcal{K}} \ J(u). \tag{34}$$

Since the objective function in (34) is concave and the admissible set is a compact and convex set, it follows that the solution to (34) is attained by a vertex function $u \in \mathcal{B} := \{v \in \mathcal{V} \colon \forall j \in V \ v_j \in \{0,1\}\}$. Here $\mathcal{B}$ is the set of binary vertex functions, taking the value 0 or 1 on each vertex. The sequential linear programming approach to solving the system (34) is to consider a sequence of vertex functions $\{u_k\}_{k=0}^{\infty}$ which satisfies

$$u_{k+1} = \arg\min_{v \in K} \ L_{u_k}(v), \qquad u_0 = \chi_S, \text{ for a node set } S \subset V. \tag{35}$$

The optimization problem in (35) may not have a unique solution, so the iterates are not well-defined. The following proposition shows that the iterations of the (MBO$_\tau$) algorithm define a unique sequence satisfying (35). Note that the optimization problem in (35) is the minimization of a linear objective function over a compact and convex set, implying that for any sequence $\{u_k\}_{k=0}^{\infty}$ satisfying (35), $u_k \in \mathcal{B}$ for all $k \geq 0$.

Proposition 4.6. *The iterations defined by the (MBO$_\tau$) algorithm satisfy (35). The functional J, defined in (33), is non-increasing on the iterates $\{u_k\}_{k=1}^{\infty}$, i.e., $J(u_{k+1}) \leq J(u_k)$, with equality only obtained if $u_{k+1} = u_k$. Consequently, the (MBO$_\tau$) algorithm with any initial condition converges to a stationary state in a finite number of iterations.*

Proof. At each iteration k, the objective functional L_{u_k} is linear and thus the minimum is attained by a function

$$u_{k+1} = \begin{cases} 1 & \text{if } 1 - 2e^{-\tau\Delta} u_k \leq 0, \\ 0 & \text{if } 1 - 2e^{-\tau\Delta} u_k > 0 \end{cases} \quad = \quad \chi_{\{e^{-\tau\Delta} u_k \geq \frac{1}{2}\}}.$$

These are precisely the (MBO$_\tau$) iterations. By the strict concavity of J and linearity of L_{u_k}, for $u_{k+1} \neq u_k$,

$$J(u_{k+1}) - J(u_k) < L_{u_k}(u_{k+1} - u_k) = L_{u_k}(u_{k+1}) - L_{u_k}(u_k).$$

Since $u_k \in \mathcal{K}$, $L_{u_k}(u_{k+1}) \leq L_{u_k}(u_k)$ which implies $J(u_{k+1}) < J(u_k)$. The convergence of the algorithm in a finite number of iterations then follows from the fact that $\mathcal{B}$ contains only a finite number of points, the vertices of the unit n-cube. □

Proposition 4.6 shows that J is a *Lyapunov function* for the (MBO$_\tau$) iterates. From the proof of Proposition 4.6, we also note that the non-uniqueness of the iterates in (35) corresponds to the choice in the (MBO$_\tau$) algorithm of thresholding vertices $\{j \in V \colon e^{-\tau\Delta} u_k = \frac{1}{2}\}$ to either 0 or 1 (see Remark 4.1).

Remark 4.7. The framework of [EO13] easily allows for the extension of the MBO algorithm to more phases, however we do not pursue these ideas here.

4.4 A local guarantee for a 'nonfrozen' (MBO$_\tau$) iteration

We begin by observing that the constant τ_κ in Theorem 4.2 depends on the maximum curvature $\kappa_S^{1,r}$ of the indicator set in the graph, $\|\Delta\chi_S\|_{\mathcal{V},\infty}$. In this section, we prove a theorem which gives a condition on τ in terms of the local curvature $(\kappa_S^{1,r})_i$ at a node i, which guarantees that the value of u on that node will change in one iteration of the graph (MBO$_\tau$) scheme.

We first introduce some notation which is needed to state the theorem. Recall that the set of neighbors of a node $i \in V$ is $\mathcal{N}_i = \{j \in V : \omega_{ij} > 0\}$. Let $1 \in V$ be an arbitrary node in the graph G and let $S \subset V$. We define the sets

$$S_1 := \begin{cases} \mathcal{N}_1 \cap S^c & \text{if } 1 \in S, \\ \mathcal{N}_1 \cap S & \text{if } 1 \notin S, \end{cases} \quad \text{and} \quad \overline{S_1} := S_1 \cup \{1\}.$$

The set S_1 contains neighbors of node 1 which are also in either the boundary $\partial(S^c)$ or ∂S (depending on whether or not $1 \in S$). For $u \in \mathcal{V}$, define Δ' as

$$(\Delta' u)_i := \begin{cases} d_i^{-r} \sum_{j \in \overline{S_1}} \omega_{ij}(u_i - u_j) & \text{if } i \in \overline{S_1}, \\ 0 & \text{if } i \notin \overline{S_1}. \end{cases}$$

We see that Δ' on $\overline{S_1}$ is similar to the Laplacian on the subgraph induced by $\overline{S_1}$, with the important distinction that, for each $i \in V$, the degree d_i is the degree of i in the full graph G, not the degree in the subgraph induced by $\overline{S_1}$. In [Chu97, Section 8.4], Δ' is referred to as the Laplacian with Dirichlet conditions on $\partial(\overline{S_1}^c)$. If $v \in \mathcal{V}_1 := \{v \in \mathcal{V} : v = 0 \text{ on } \overline{S_1}^c\}$, then

$$(\Delta' v)_i = \begin{cases} (\Delta v)_i & \text{if } i \in \overline{S_1}, \\ 0 & \text{if } i \notin \overline{S_1}. \end{cases}$$

Note in particular that, if $v \in \mathcal{V}_1$, then $e^{-t\Delta'} v \in \mathcal{V}_1$ for all $t \geq 0$.

Theorem 4.8. *Let $1 \in V$ be an arbitrary node and $S \subset V$ be such that $|(\kappa_S^{1,r})_1|^2 > \|(\Delta')^2 \chi_{S_1}\|_{\mathcal{V},\infty}$. If $\tau \in (\tau_1, \tau_2)$, where*

$$\tau_{1,2} := \frac{1}{\|(\Delta')^2 \chi_{S_1}\|_{\mathcal{V},\infty}} \left(|(\kappa_S^{1,r})_1| \pm \sqrt{|(\kappa_S^{1,r})_1|^2 - \|(\Delta')^2 \chi_{S_1}\|_{\mathcal{V},\infty}} \right) > 0, \tag{36}$$

then

$$|(P e^{-\tau \Delta} \chi_S)_1 - (\chi_S)_1| = 1.$$

That is, the phase at node 1 changes after one (MBO_τ) iteration.

It is important to note that both $|(\kappa_S^{1,r})_1|^2$ and $\|(\Delta')^2 \chi_{S_1}\|_{\mathcal{V},\infty}$ are local quantities, in the sense that they only depend on the structure of G at node 1 and its neighbors. This is in contrast with Theorem 4.2, which depends on the spectrum of the Laplacian on G. The existence of a lower bound τ_1 on τ is unsurprising in the light of this earlier freezing result. The necessity for an upper bound τ_2 can be understood from our wish to only use local quantities in this theorem.

Proof of Theorem 4.8. First we assume that $1 \notin S$, so $S_1 = \mathcal{N}_1 \cap S$. By the comparison principle in Lemma 2.5(d), $\chi_{S_1} \leq \chi_S$ on V implies $(e^{-\tau\Delta} \chi_{S_1})_1 \leq (e^{-\tau\Delta} \chi_S)_1$. In particular, since $(\chi_{S_1})_1 = (\chi_S)_1 = 0$, we have

$$(e^{-\Delta\tau} \chi_S - \chi_S)_1 \geq (e^{-\tau\Delta} \chi_{S_1} - \chi_{S_1})_1.$$

Let v satisfy the heat equation with Dirichlet boundary data,

$$\begin{cases} \dot{v} = -(\Delta' v)_i, \\ v(0) = \chi_{S_1}. \end{cases}$$

As noted above the theorem, $v(t) \in \mathcal{V}_1$ for all $t \geq 0$.

It is easily checked that v is subcaloric, *i.e.*, $\dot{v}_i \leq -(\Delta v)_i$ for all $i \in V$, and $v(0) \leq \chi_S$. In addition, the Laplacian satisfies $-(\Delta u)_i \leq -(\Delta \tilde{u})_i$ if $u_i = \tilde{u}_i$ and $u_j \leq \tilde{u}_j$, for $j \neq i$. Hence, by the theory of differential inequalities (see for example [Sza65, Theorem 8.1(3)]),

$$v_i(t) \leq \left(e^{-t\Delta} v(0)\right)_i = \left(e^{-t\Delta}\chi_{S_1}\right)_i, \quad \text{for all } i \in V.$$

In particular,

$$\left(e^{-\tau\Delta}\chi_{S_1} - \chi_{S_1}\right)_1 \geq v_1(\tau) - v_1(0) = \left(e^{-\tau\Delta'}\chi_{S_1} - \chi_{S_1}\right)_1 = -\tau\left(\Delta'\chi_{S_1}\right)_1 + \tau^2 r(\tau),$$

where

$$|r(\tau)| \leq \frac{1}{2}\sup_{t\in[0,\tau]}\left(e^{-t\Delta'}(\Delta')^2\chi_{S_1}\right)_1 \leq \frac{1}{2}\|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}.$$

Note that $-\left(\Delta'\chi_{S_1}\right)_1 = -(\kappa_S^{1,r})_1 = |(\kappa_S^{1,r})_1|$, where the last equality follows because $1 \notin S$.

We conclude that

$$(e^{-\tau\Delta}\chi_S - \chi_S)_1 \geq \left(e^{-\tau\Delta}\chi_{S_1} - \chi_{S_1}\right)_1 \geq |(\kappa_S^{1,r})_1|\tau - \frac{1}{2}\|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}\tau^2,$$

hence

$$(e^{-\tau\Delta}\chi_S - \chi_S)_1 \geq \frac{1}{2} \Leftrightarrow \tau \in [\tau_1, \tau_2],$$

which proves the result for the case in which $1 \notin S$.

To prove the desired statement if $1 \in S$, we note that

$$(e^{-\tau\Delta} - 1)(\chi_S + \chi_{S^c}) = 0,$$

So the condition $(e^{-\tau\Delta}\chi_S - \chi_S)_1 < -\frac{1}{2}$ is equivalent to $(e^{-\tau\Delta}\chi_{S^c} - \chi_{S^c})_1 > \frac{1}{2}$. Recall that, in this case, $S_1 = \mathcal{N}_1 \cap S^c$, and the same derivation as above holds[14], since $1 \notin S^c$, with the exception that the admissible range of τ becomes the open interval (τ_1, τ_2). This is because, by our definition of the (MBO$_\tau$) algorithm, the thresholding operator thresholds the $\frac{1}{2}$-level set to 1.

□

In the remainder of this section we determine some conditions under which the requirement $|(\kappa_S^{1,r})_1|^2 > \|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}$ in Theorem 4.8 is satisfied. To this end, define the reduced degrees, for $i \in V$, as

$$d_i' := \sum_{j\in S_1}\omega_{ij}. \tag{37}$$

Lemma 4.9. *$|(\kappa_S^{1,r})_1|^2 > \|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}$ if and only if*

$$d_1^{-2r}(d_1')^2 > \max_{i\in\overline{S_1}} d_i^{-r}\left|-d_i^{1-r}d_i' - \sum_{j\in S_1} d_j^{1-r}\omega_{ij} + \sum_{k\in\overline{S_1}} d_k^{-r}d_k'\omega_{ik}\right|. \tag{38}$$

Proof. Consider the $|\mathcal{V}| \times |\mathcal{V}|$ matrix corresponding to Δ'. After possibly relabeling the nodes, it can be written as $\begin{pmatrix} L' & 0 \\ 0 & 0 \end{pmatrix}$, where L' is the $|\overline{S_1}| \times |\overline{S_1}|$ matrix with entries

$$L'_{ij} = d_i^{-r}\begin{cases} -d_i & \text{if } i = j, \\ \omega_{ij} & \text{if } i \neq j. \end{cases}$$

[14] In particular, carefully note that now $-\left(\Delta'\chi_{S_1}\right)_1 = (\kappa_S^{1,r})_1 = |(\kappa_S^{1,r})_1|$ holds.

Then

$$(L'^2)_{ij} = \sum_{k\in\overline{S_1}} (L')_{ik}(L')_{kj} = (L')_{ii}(L')_{ij} + (L')_{ij}(L')_{jj} + \sum_{k\in\overline{S_1}\setminus\{i,j\}} (L')_{ik}(L')_{kj}$$
$$= d_i^{-r}\left[-(d_i^{1-r} + d_j^{1-r})\omega_{ij} + \sum_{k\in\overline{S_1}\setminus\{i,j\}} d_k^{-r}\omega_{ik}\omega_{jk}\right].$$

Thus $\left((\Delta')^2\chi_{S_1}\right)_i = 0$ if $i \notin \overline{S_1}$, and, for $i \in \overline{S_1}$, we have

$$\left((\Delta')^2\chi_{S_1}\right)_i = d_i^{-r}\sum_{j\in S_1}\left[-(d_i^{1-r} + d_j^{1-r})\omega_{ij} + \sum_{k\in\overline{S_1}\setminus\{i,j\}} d_k^{-r}\omega_{ik}\omega_{jk}\right]$$
$$= d_i^{-r}\sum_{j\in S_1}\left[-(d_i^{1-r} + d_j^{1-r})\omega_{ij} + \sum_{k\in\overline{S_1}} d_k^{-r}\omega_{ik}\omega_{jk}\right]$$
$$= d_i^{-r}\left[-d_i^{1-r}d'_i - \sum_{j\in S_1} d_j^{1-r}\omega_{ij} + \sum_{k\in\overline{S_1}} d_k^{-r}d'_k\omega_{ik}\right],$$

where, for the second equality, we used that $\omega_{ii} = \omega_{jj} = 0$.

From (14) and the definition of S_1, we find

$$|(\kappa_S^{1,r})_1|^2 = d_1^{-2r}\sum_{j,k\in S_1}\omega_{1j}\omega_{1k} = d_1^{-2r}d_1'^2.$$

□

Corollary 4.10. *Let $r = 1$ and $d'_1 > 0$. If there exists an $0 \le \varepsilon < 1$, such that, for all $i \in S_1$,*

$$\frac{d'_i}{d_i} \le \varepsilon\frac{d'_1}{d_1} \quad \text{and} \quad \frac{\omega_{i1}}{d_i} < (1-\varepsilon^2)\frac{d'_1}{d_1}, \tag{39}$$

then condition (38) is satisfied, and there exist $0 < \tau_1 < \tau_2$ as in Theorem 4.8.

If the first condition in (39) is satisfied with $0 \le \varepsilon < \frac{1}{2}\left(\sqrt{5}-1\right) \approx 0.618$, then the second condition in (39) can be replaced by the condition that, for all $i \in S_1$, $\omega_{i1} \le d'_i$.

Note, by (14) and (37), that the conditions in (39) can be rewritten as

$$(\kappa_{S_1}^{1,1})_i - 1 \ge \varepsilon(\kappa_{S_1}^{1,1})_1 \quad \text{and} \quad (\kappa_{\{1\}}^{1,1})_i > (1-\varepsilon^2)(\kappa_{S_1}^{1,1})_1,$$

for all $i \in S_1$.

Proof of Corollary 4.10. For $r = 1$, we compute, for $i \in \overline{S_1}$,

$$d_i^{-r}\left|-d_i^{1-r}d'_i - \sum_{j\in S_1} d_j^{1-r}\omega_{ij} + \sum_{k\in\overline{S_1}} d_k^{-r}d'_k\omega_{ik}\right| = \left|\frac{d'_i}{d_i} - \frac{d'_i}{d_i} + \sum_{k\in\overline{S_1}}\frac{d'_k}{d_k}\frac{\omega_{ik}}{d_i}\right| = \sum_{k\in\overline{S_1}}\frac{d'_k}{d_k}\frac{\omega_{ik}}{d_i},$$

hence condition (38) becomes $\left(\frac{d'_1}{d_1}\right)^2 > \max_{i\in\overline{S_1}}\sum_{k\in\overline{S_1}}\frac{d'_k}{d_k}\frac{\omega_{ik}}{d_i}$. If the first condition in (39) is satisfied, we have

$$\sum_{k\in\overline{S_1}}\frac{d'_k}{d_k}\frac{\omega_{ik}}{d_i} = \frac{d'_1}{d_1}\frac{\omega_{i1}}{d_i} + \sum_{k\in S_1}\frac{d'_k}{d_k}\frac{\omega_{ik}}{d_i} \le \frac{d'_1}{d_1}\frac{\omega_{i1}}{d_i} + \varepsilon\frac{d'_1}{d_1}\frac{d'_i}{d_i} = \frac{d'_1}{d_1}\left[\frac{\omega_{i1}}{d_i} + \varepsilon\frac{d'_i}{d_i}\right].$$

Since $d'_1 > 0$, condition (38) reduces to $\max_{i\in\overline{S_1}} \left[\frac{\omega_{i1}}{d_i} + \varepsilon\frac{d'_i}{d_i}\right] < \frac{d'_1}{d_1}$. If the maximum is achieved at $i = 1$, then $\frac{\omega_{i1}}{d_i} + \varepsilon\frac{d'_i}{d_i} = \varepsilon\frac{d'_1}{d_1} < \frac{d'_1}{d_1}$. If, on the other hand, the maximum is achieved at some $i \in S_1$, then $\frac{\omega_{i1}}{d_i} + \varepsilon\frac{d'_i}{d_i} \le \frac{\omega_{i1}}{d_i} + \varepsilon^2\frac{d'_1}{d_1} < \frac{d'_1}{d_1}$. Here, we first used the first condition from (39), and then the second. Combined with Theorem 4.8 and Lemma 4.9, this proves the first claim.

Now assume, instead of the second condition in (39), that $\omega_{i1} \le d'_i$ for all $i \in S_1$. Then $\frac{\omega_{i1}}{d_i} \le \frac{d'_i}{d_i} \le \varepsilon\frac{d'_1}{d_1}$. Hence, if $0 < \varepsilon < 1 - \varepsilon^2$, the second condition in (39) is satisfied. This requirement is met, if $0 < \varepsilon < \frac{1}{2}\left(\sqrt{5} - 1\right)$. □

It is worthwhile to understand the conditions in the corollary. Adding the two conditions in (39) gives $\frac{d'_i+\omega_{i1}}{d_i} \le (1 + \varepsilon - \varepsilon^2)\frac{d'_1}{d_1}$. The ratio $\frac{d'_i+\omega_{i1}}{d_i}$ is a measure of the relative strength of connection of node i within the set $\overline{S}_1$ compared to all its connections in G. Similarly $\frac{d'_1}{d_1}$ is the relative strength of connection of node 1 within $\overline{S}_1$ (or, equivalently, within S_1), compared to all its connections in G. The conditions in (39) thus require node 1 to be a node with comparatively large relative connection strength within $\overline{S}_1$, compared to the other nodes in $\overline{S_1}$. This will allow enough mass to diffuse to or away from node 1 (depending on whether or not $1 \in S$), for it to pass the threshold value $\frac{1}{2}$, without too much of the locally available mass diffusing to other nodes.

Some examples in Section 6 further examines the conditions in (38) and (39).

Remark 4.11. One can interpret the condition $|(\kappa_S^{1,r})_1|^2 > \|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}$ from Theorem 4.8, in terms of the spectral radius of Δ' (similar to [Chu97, Equation (8.7)]). We compute

$$\begin{aligned}
\rho(\Delta') &= \sup_{u\in\mathcal{V}\setminus\{0\}} \frac{\langle u, \Delta' u\rangle_{\mathcal{V}}}{\|u\|_{\mathcal{V}}} = \frac{1}{2}\sup_{u\in\mathcal{V}\setminus\{0\}} \frac{\sum_{i,j\in\overline{S_1}} \omega_{ij}(v_i - v_j)^2}{\sum_{i\in\overline{S_1}} d_i^r v_i^2} \\
&\le \sup_{u\in\mathcal{V}\setminus\{0\}} \frac{\sum_{i,j\in\overline{S_1}} \omega_{ij}(v_i^2 + v_j^2)}{\sum_{i\in\overline{S_1}} d_i^r v_i^2} = 2\sup_{u\in\mathcal{V}\setminus\{0\}} \frac{\sum_{i,j\in\overline{S_1}} \omega_{ij} v_i^2}{\sum_{i\in\overline{S_1}} d_i^r v_i^2} \\
&= 2\sup_{u\in\mathcal{V}\setminus\{0\}} \frac{\sum_{i\in\overline{S_1}} \bar{d}_i v_i^2}{\sum_{i\in\overline{S_1}} d_i^r v_i^2} \le 2(\bar{d}_-)^{-r}\bar{d}_+,
\end{aligned}$$

where

$$\bar{d}_i := \sum_{j\in\overline{S_1}} \omega_{ij}, \quad \bar{d}_+ := \max_{i\in\overline{S_1}} \bar{d}_i, \quad \bar{d}_- := \min_{i\in\overline{S_1}} \bar{d}_i.$$

Because $\left((\Delta')^2\chi_{S_1}\right)_i = 0$ if $i \notin \overline{S_1}$, it is straightforward to adapt the proof of Lemma 2.1, to find $(\bar{d}_-)^{\frac{r}{2}}\|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty} \le \|(\Delta')^2\chi_{S_1}\|_{\mathcal{V}}$.

Combining these results, the condition $|(\kappa_S^{1,r})_1|^2 > \|(\Delta')^2\chi_{S_1}\|_{\mathcal{V},\infty}$ is satisfied if $d_1^{-2r}(d'_1)^2 > 4(\bar{d}_-)^{-\frac{r}{2}}(\bar{d}_-)^{-2r}(\bar{d}_+)^2\sqrt{\text{vol } S_1}$, or, equivalently, if

$$\frac{1}{4}\left(\frac{d'_1}{\bar{d}_+}\right)^2\left(\frac{\bar{d}_-}{d_1}\right)^{2r} > (\bar{d}_-)^{-\frac{r}{2}}\sqrt{\text{vol } S_1}.$$

Using vol $S_1 = \sum_{i\in S_1} d_i^r \ge |S_1|d_-^r$, we can deduce the stronger sufficient condition

$$\frac{1}{4}\left(\frac{d'_1}{\bar{d}_+}\right)^2\left(\frac{\bar{d}_-}{d_1}\right)^{2r} > \left(\frac{d_-^r}{\bar{d}_-}\right)^{-\frac{r}{2}}|S_1|.$$

5 Allen-Cahn equation on graphs

In this section, we investigate the Allen-Cahn equation on graphs. A short overview of the continuum Allen-Cahn equation can be found in Section A.1 in Appendix A.

We propose the following Allen-Cahn equation (ACE) on graphs, for all $i \in V$:

$$\begin{cases} \dot{u}_i &= -(\Delta u)_i - \frac{1}{\varepsilon} d_i^{-r} W'(u_i) \quad \text{for } t > 0, \\ u_i &= (u_0)_i \qquad\qquad\qquad\quad \text{at } t = 0, \end{cases} \tag{ACE$_\varepsilon$}$$

for a given initial condition $u_0 \in \mathcal{V}$ and $\varepsilon > 0$. Here $W \in C^2(\mathbb{R})$ is a double well potential. For definiteness we set W to be the standard double well potential $W(u) = (u+1)^2(u-1)^2$, hence $W'(u) = 4u(u^2-1)$ and W has two stable minima at the wells at $u = \pm 1$ and an unstable local maximum at $u = 0$. Recall that the sign convention for the Laplacian is opposite to the one used in the continuum literature. Note that for ε sufficiently small, this system has 3^n equilibria, of which 2^n are stable.

The (MBO$_\tau$) algorithm is closely related to time-splitting methods applied to the Allen-Cahn evolution (ACE$_\varepsilon$). The diffusion step is precisely the time evolution with respect to the first term of (ACE$_\varepsilon$) and the thresholding step is the asymptotic behavior of evolution with respect to the second term of (ACE$_\varepsilon$).

The case $V = \mathbb{Z}^d$ with weights $\omega_{ij} = \omega(\|i-j\|)$ was considered in [BC99], where it is seen as an approximation to the Ising model, and stationary solutions and traveling waves are constructed for ε small enough. The authors note that, when ω_{ij} corresponds to nearest-neighbors, this equation is known as the discrete Nagumo equation, which is a simplified model of neural networks. In this context, [HPS11] considered the Nagumo equation in $\mathbb{Z}^1$ and derived the existence of traveling waves. In general, we are not aware of any previous works where (ACE$_\varepsilon$) is considered for an arbitrary weighted graph (V, E, ω_{ij}). It would be interesting to see whether the analysis in [BC99] can can be extended to use (ACE$_\varepsilon$) to study phase transitions in general graphs, a topic of interest in other areas of mathematics [Lyo00].

Just as in the continuum case, we arrive at (ACE$_\varepsilon$) as the gradient flow given by the graph Ginzburg-Landau functional, $GL_\varepsilon(u)\colon \mathcal{V} \to \mathbb{R}$,

$$GL_\varepsilon(u) := \frac{1}{2}\|\nabla u\|_{\mathcal{E}}^2 + \frac{1}{\varepsilon}\langle D^{-r} W \circ u, 1\rangle_{\mathcal{V}}, \tag{GL$_\varepsilon$}$$

where $(D^{-r}W \circ u)_i = d_i^{-r} W(u_i)$, and whose first variation is given by

$$\frac{d}{dt} GL_\varepsilon(u + tv)\Big|_{t=0} = \langle \Delta u, v\rangle_{\mathcal{V}} + \frac{1}{\varepsilon}\langle D^{-r} W'(u), v\rangle_{\mathcal{V}}.$$

The factor d_i^{-r} in the potential term is needed to cancel the factor d_i^r in the $\mathcal{V}$-inner product. Equation (ACE$_\varepsilon$) is then the $\mathcal{V}$-gradient flow associated with (GL$_\varepsilon$).

Recall that the Laplacian Δ also depends on r. In fact, the equation in (ACE$_\varepsilon$) can be rewritten as

$$d_i^r \dot{u}_i = -\sum_{j \in V} \omega_{ij}(u_i - u_j) - \frac{1}{\varepsilon} W'(u_i),$$

showing that the factor d_i^r can be interpreted as a node-dependent time rescaling.

By standard ODE arguments and the smoothness of the right hand side of (ACE$_\varepsilon$), for each $\varepsilon > 0$ a unique C^1 solution to (ACE$_\varepsilon$) exists for all $t > 0$.

This continuum case (see Appendix A.1) suggests an approach for finding a valid notion of mean curvature (and its flow) for graphs: Take initial data $u(0) = \chi_S - \chi_{S^c}$, for some node

set $S \subset V$, and consider the corresponding solution $u^\varepsilon(t)$ to (ACE$_\varepsilon$), for all times $t > 0$. The question is whether the limit

$$\bar{u}_i(t) := \lim_{\varepsilon \to 0_+} u_i^\varepsilon(t)$$

exists. Even if it does, it is unlikely that $\bar{u}$ is of the form $\chi_{S(t)} - \chi_{S(t)^c}$, which can be interpreted as a binary indicator function for some evolving set $S(t)$, for all times $t > 0$, but there may be an approximate phase separation: $\bar{u}_i(t) \in [-1-\delta, -1+\delta] \cup [1-\delta, 1+\delta]$, for some small $\delta > 0$. Is there a way to characterize the evolution of the "interface" between the two level sets of $\bar{u}_i(t)$?

However, a little analysis shows the above approach is rather naïve. Indeed, unlike in the continuum case, the graph Laplacian of the indicator function of a set $S \subset V$ is always a well-defined bounded function (in any norm). Thus, for small ε the potential term in the equation will dominate the dynamics, and *pinning* or *freezing* will occur, as proven in Theorem 5.3. This is the dynamics in which the sign of the value of u on each node is fixed by the sign of the initial value, and u at each node just settles into the corresponding well of W.

As discussed at the start of Section 3.3, the question how to connect the sequence of sets evolving by graph mean curvature to the super (or sub) level sets $\{i \in V \colon u_i^\varepsilon(t) > 0\}$ for solutions of (ACE$_\varepsilon$), is still open. See also Question 7.4.

Remark 5.1. Note that in the (MBO$_\tau$) algorithm, the values of u are reinitialized in every iteration to 0 or 1. Our choice of the double well potential W in (ACE$_\varepsilon$) has two equilibria corresponding to the level sets for ± 1. Correspondingly, the unstable equilibrium for (MBO$_\tau$) corresponds to the 1/2 level set, while for (ACE$_\varepsilon$) it corresponds to the 0 level set. This agrees with the now standard notations for Allen-Cahn and MBO.

Below, we show that for all ε below a finite $\varepsilon_0 > 0$ the functions $u_i^\varepsilon(t)$ do not change sign as t varies, so that pinning occurs. Recall that a set which contains the forward orbit of each of its elements is called *positively invariant*, and that the number of nodes in the graph G is $|V| = n$.

Lemma 5.2. *Consider the set $S := \{u \in \mathcal{V} \colon \|u\|_{\mathcal{V}}^2 \leq \frac{17}{4} n\, d_+^{-r}\}$ and let $u(t)$ be the solution to (ACE$_\varepsilon$) for a given $\varepsilon > 0$. Then $t \mapsto \|u(t)\|_{\mathcal{V}}^2$ is decreasing at each t such that $u(t) \in S^c$. As a consequence, the set S is positively invariant and every trajectory of (ACE$_\varepsilon$) enters S in finite time.*

Proof. Define the set $A(t) := \{i \in V \colon u_i^2(t) \leq 2\}$. We compute

$$\begin{aligned}
\tfrac{d}{dt}\|u(t)\|_{\mathcal{V}}^2 &= 2\,\langle u(t), \dot{u}(t)\rangle_{\mathcal{V}} \\
&= -2\|\nabla u(t)\|_{\mathcal{E}}^2 - \tfrac{8}{\varepsilon}\sum_{i\in V} u_i(t)^2(u_i(t)^2 - 1) \\
&= -2\|\nabla u(t)\|_{\mathcal{E}}^2 - \tfrac{8}{\varepsilon}\sum_{i\in A^c(t)} u_i(t)^2(u_i(t)^2 - 1) + \tfrac{8}{\varepsilon}\sum_{i\in A(t)} u_i(t)^2(1 - u_i(t)^2) \\
&< -\tfrac{8}{\varepsilon}\left(\sum_{i\in A^c(t)} u_i(t)^2 - \tfrac{|A(t)|}{4}\right).
\end{aligned}$$

The last inequality follows, since $u_i(t)^2 - 1 > 1$ for $i \in A^c(t)$, and $\max\{x^2(1-x^2) \colon x^2 \leq 2\} = \frac{1}{4}$. Note that $\|u(t)\|_{\mathcal{V}}^2 \leq d_+^r \sum_{i\in V} u_i(t)^2$. Thus, if $u(t) \in S^c$, then $\sum_{i\in V} u_i(t)^2 > \frac{17}{4} n$, and hence

$$\sum_{i\in A^c(t)} u_i(t)^2 = \sum_{i\in V} u_i(t)^2 - \sum_{i\in A(t)} u_i(t)^2 > \tfrac{17}{4} n - 4|A(t)|.$$

Therefore,

$$\tfrac{d}{dt}\|u(t)\|_{\mathcal{V}}^2 < -\tfrac{8}{\varepsilon}\left(\tfrac{17}{4} n - 4|A(t)| - \tfrac{1}{4}|A(t)|\right) < 0,$$

where we have used that $|A(t)| \leq n$. This shows $\|u(t)\|_{\mathcal{V}}^2$ is decreasing in the region S^c, as desired. The other statements in the lemma now follow. □

Theorem 5.3. *Assume $|u_i(0)| > 0$ for all $i \in V$. There exist an ε_ρ and an ε_κ (depending on the spectral radius of Δ via (41), and on $\sup_{t\geq 0} \|\Delta u(t)\|_{\mathcal{V},\infty} < \infty$ via (42), respectively), such that, if either $\varepsilon \leq \varepsilon_\rho$ or $\varepsilon \leq \varepsilon_\kappa$, then the solution $u(t)$ to (ACE$_\varepsilon$) is such that sign$(u_i(t))$ is constant in time, for all $i \in V$.*

Proof. By Lemma 5.2, $\|u(t)\|_{\mathcal{V}}^2 \leq \frac{17}{4} n\, d_+^{-r}$ for t large enough. Hence, by continuity of $u(t)$, there is a C (depending on the initial condition) such that, for all $t \geq 0$,

$$\|u(t)\|_{\mathcal{V}} \leq C.$$

Thus, if $\rho > 0$ denotes the spectral radius of Δ, we get, for all $i \in V$,

$$d_i^r |\Delta u_i(t)|^2 \leq \|\Delta u(t)\|_{\mathcal{V}}^2 \leq \rho^2 C^2.$$

In particular

$$|\Delta u_i(t)| \leq \rho C d_i^{-\frac{r}{2}}, \tag{40}$$

for all $i \in V$, thus, we have the inequalities

$$-\rho C d_i^{-\frac{r}{2}} - \tfrac{1}{\varepsilon} d_i^{-r} W'(u_i) \leq \dot{u}_i \leq \rho C d_i^{-\frac{r}{2}} - \tfrac{1}{\varepsilon} d_i^{-r} W'(u_i).$$

Without loss of generality, we can assume that there is a number $\alpha \in (0,1)$ such that $|u_i(0)| \geq \alpha$ for all $i \in V$. If there is an $i \in V$ such that $|u_i(t)| = \alpha$ for a given t, then we have that $|W'(u_i(t))| = 4\alpha(1-\alpha^2)$, with a sign opposite to that of u_i. Thus, if

$$\varepsilon \leq \varepsilon_0 := C^{-1}\rho^{-1} 4\alpha(1-\alpha^2) d_+^{-\frac{r}{2}} \leq C^{-1}\rho^{-1} 4\alpha(1-\alpha^2) d_i^{-\frac{r}{2}}, \tag{41}$$

then $\dot{u}_i \leq 0$ if $u_i(0) < 0$, and $\dot{u}_i \geq 0$ if $u_i(0) > 0$. Hence $u_i(t)$ can never reach zero, and by continuity in t it does not change sign.

Alternatively, instead of (40), we can estimate $|\Delta u_i(t)| \leq \sup_{t\geq 0} \|\Delta u(t)\|_{\mathcal{V},\infty} < \infty$. The finitude of the right hand side follows from (40). Following the same reasoning as above, we then conclude

$$\varepsilon_\kappa := \left(\sup_{t\geq 0} \|\Delta u(t)\|_{\mathcal{V},\infty} \right)^{-1} d_+^{-r} 4\alpha(1-\alpha)^2. \tag{42}$$

□

The constant ε_κ in Theorem 5.3 involves $\|\Delta u\|_{\mathcal{V},\infty}$, which is "curvature-like". Compare this to the constant τ_κ in Theorem 4.2, which depends on the maximum curvature of the indicator set in the graph, $\|\Delta \chi_S\|_{\mathcal{V},\infty}$, as also discussed in Section 4.4. This tentative similarity makes us suspect, that a condition on the local curvature, similar to those for the (MBO$_\tau$) algorithm given in Theorem 4.8, guarantees a phase change in the Allen-Cahn flow. We discuss this further in Question 7.3.

We see that the discrete nature of the graph, manifest in the finite spectral radius of the Laplacian, makes the limit behavior of (ACE$_\varepsilon$) as $\varepsilon \to 0$ much different than that for the continuum case. In particular, this means that we ought to look for a notion of mean curvature flow on graphs more carefully.

Remark 5.4. For ε small enough, but not smaller than the ε_0 from Theorem 5.3 above, we expect interesting asymptotic behavior for the motion of the phases in (ACE_ε) on intermediate time scales. Such asymptotics might be connected to the graph curvature of the phases, which would match the situation in the continuum setting, where the solution has phases that for large times behave as if they were evolving by mean curvature flow, while the solution itself becomes stationary in the limit $t \to +\infty$. This phenomenon is known as dynamic metastability (see for instance [BK91] and the references therein). See also Question 7.4.

6 Explicit and computational examples

In this section we give several examples of graphs where the mean curvature, MBO, and Allen Cahn evolutions can be compared either explicitly or computationally.

6.1 Complete graph

Consider the complete graph, K_n, on n nodes with $\omega_{ij} = \omega$ for all $i, j \in V$. See Figure 1a. In this case, the matrix representation of the graph Laplacian is given by the circulant matrix,

$$L = \omega[(n-1)\omega]^{-r}\left(n\ \mathrm{Id}_n - 1_n 1_n^t\right),$$

where 1_n denotes the vector in $\mathbb{R}^n$ with all entries equal to 1, and Id_n the identity matrix in $\mathbb{R}^{n\times n}$. The eigenvalues of L are given by 0 and $\omega n[(n-1)\omega]^{-r}$ (with multiplicity $n-1$). In particular, $\lambda_2 = \lambda_n = \rho$. Note that the normalized eigenvector corresponding to eigenvalue 0 is given by $(\mathrm{vol}\ V)^{-\frac{1}{2}}\chi_V$.

Let $S \subset V$ be a set with volume ratio $R_S = \frac{\mathrm{vol}\ S}{\mathrm{vol}\ V}$ (see also Theorem 4.3). Using the spectral decomposition from (11), the evolution of χ_S by the heat equation can be explicitly written as

$$e^{-t\Delta}\chi_S = R_S\chi_V + e^{-\rho t}\left(\chi_S - R_S\chi_V\right).$$

Assume $R_S \neq \frac{1}{2}$. Then there exists a critical time step τ_c depending only on vol S, vol V, and ρ such that $\tau < \tau_c$ implies the solution to the (MBO_τ) evolution is pinned and $\tau \geq \tau_c$ implies exactly one iteration of the (MBO_τ) evolution gives a stationary solution, either 0 or χ_V depending on the initial mass, $M(\chi_S) = \mathrm{vol}\ S$ (see (10)). From the solution, the critical time step τ_c can be directly computed,

$$\tau_c = \frac{1}{\rho}\log\frac{\max\{Rs, 1-R_S\}}{|\frac{1}{2} - R_S|}$$

If $R_S = \frac{1}{2}$, symmetry pins the (MBO_τ) evolution for all $\tau > 0$.

The bound from Theorem 4.2 states that pinning occurs if $\tau < \tau_p = \rho^{-1}\log\left(1 + \frac{1}{2}(nR_S)^{-\frac{1}{2}}\right)$, where we have used the fact that, for all $i \in V$, $d_i^r = \frac{\mathrm{vol}V}{n}$. The bound in Theorem 4.3 states that trivial dynamics occur if $\tau > \tau_t = \rho^{-1}\log\left(\frac{n^{\frac{1}{2}}R_S^{\frac{1}{2}}R_{S^c}^{\frac{1}{2}}}{|\frac{1}{2}-R_S|}\right)$. Note that for $n > 2$, $R_S > \frac{1}{n}$ and $\tau_t > \tau_c > \tau_p$.

By symmetry, both the Allen-Cahn equations and mean curvature flows reduce to two-dimensional systems, with one variable governing the value of the nodes in S and the other the nodes in S^c. Critical parameters ϵ and $\eth t$ exist for which below the phase remains the same for all nodes and above the phase simultaneously changes.

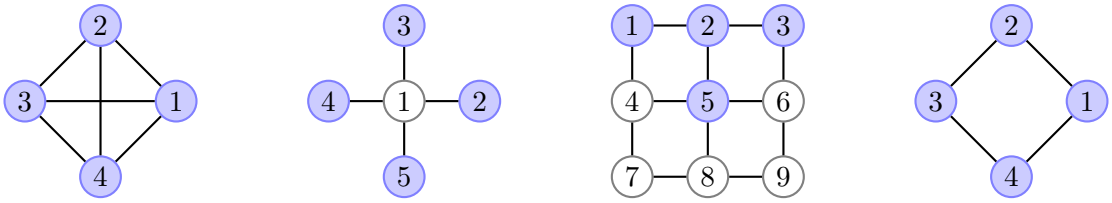


Figure 1: Some small graphs, discussed in the examples of Section 6. **(a)** The complete graph K_4, **(b)** the star graph SG_5, **(c)** a small grid, and **(d)** a cycle graph C_4; see Sections 6.1, 6.2, 6.4, and 6.5.

6.2 Star graph

Consider a star graph SG_n as in Figure 1b with $n \geq 3$ nodes. Here the central node (say node 1) is connected to all other nodes and the other $n-1$ nodes are only connected to the central node. Hence, for all $i \in \{2, \dots, n\}$, $\omega_{1i} = \omega_{i1} > 0$, and all the other ω_{jk} are zero.

We consider the unnormalized graph Laplacian $L = D - A$ ($r = 0$ in (8)). Since $d_1 = \sum_{j=2}^{n} \omega_{1j}$ and $d_i = \omega_{1i}$, for $i \in \{2, \dots, n\}$, we can explicitly compute the characteristic polynomial of L:

$$p(\lambda) = \Big(-\lambda + \sum_{k=2}^{n} \omega_{1k} \Big) \prod_{j=2}^{n} (\omega_{1j} - \lambda) - \sum_{k=2}^{n} \omega_{1k}^2 \prod_{j \geq 2,\, j \neq k} (\omega_{1j} - \lambda).$$

If all non-zero edge weights have the same value ω, this simplifies considerably to

$$p(\lambda) = \big((n-1)\omega - \lambda\big)(\omega - \lambda)^{n-1} - \omega^2 (n-1)(\omega - \lambda)^{n-2}.$$

Hence, in this case, the eigenvalues are $\lambda_1 = 0$, $\lambda_i = \omega$ for $i \in \{2, \dots, n-1\}$, and $\lambda_n = n\omega$. A choice of corresponding (normalized) eigenvectors $\{v^i\}_{i=1}^n$ is given by[15]

$$v^1 = n^{-\frac{1}{2}} \chi_V, \qquad v_j^i = 2^{-\frac{1}{2}} \begin{cases} 1 & \text{if } j = i, \\ -1 & \text{if } j = i+1, \\ 0 & \text{else,} \end{cases} \quad \text{for } i \in \{2, \dots, n-1\}$$

$$v_j^n = \frac{1}{\sqrt{n(n-1)}} \begin{cases} n-1 & \text{if } j = 1, \\ -1 & \text{if } j \neq 1. \end{cases}$$

We now let $S = \{1\}$ and note that χ_S has the explicit expansion in terms of these eigenvectors,

$$\chi_S = n^{-\frac{1}{2}} v^1 + (n-1)^{\frac{1}{2}} n^{-\frac{1}{2}} v^n.$$

We now consider the (MBO$_\tau$) iterates of χ_S. We compute

$$e^{-\Delta\tau} \chi_S = n^{-1} \chi_V + (n-1)^{\frac{1}{2}} n^{-\frac{1}{2}} (e^{-n\omega\tau}) v^n.$$

Thus pinning occurs if $\tau < \tau_c := \frac{1}{n\omega} \log\left(2\frac{n-1}{n-2}\right)$. If $\tau > \tau_c$, the solution to the (MBO$_\tau$) evolution gives the stationary solution, 0, after exactly one iteration. The bound from Theorem 4.2 states that pinning occurs if $\tau < \frac{1}{n\omega} \log \frac{3}{2}$. The bound in Theorem 4.3 states that trivial dynamics

[15] Here, subscripts j denote the components of the vectors.

occur if $\tau > \frac{1}{\omega}\log\left(2\frac{n-1}{n-2}\right)$. Qualitatively, this example shows that it is easier for a solution to be pinned on nodes with smaller degree.

We now consider an implication of Theorem 4.8 in the case where the graph induced by $\overline{S_1}$ is a star graph with node 1 as center, *i.e.*, $d'_i = 0$ for all $i \in S_1$. This is certainly true for the case when the graph is a star graph and $S = \{2, \ldots, n\}$. The following lemma states, in the case where $r = 1$, a simple criterion on the degrees for which there exists a τ interval in which node 1 switches phase in a single iteration of (MBO$_\tau$). We will see an application of the lemma in Section 6.3.

Lemma 6.1. *Let $r = 1$ and consider the case where the subgraph induced by $\overline{S_1}$ is a star graph with node 1 as center, i.e., $d'_i = 0$ for all $i \in S_1$. If either*

- $|S_1| = 1$ *and* $d_i < d_1$, *for* $i \in S_1$, *or*
- $|S_1| \geq 2$ *and* $d_i \leq d_1$, *for all* $i \in S_1$,

then there exists a nontrivial τ interval (as in Theorem 4.8) such that node 1 will change phase in the next (MBO$_\tau$) iterate.

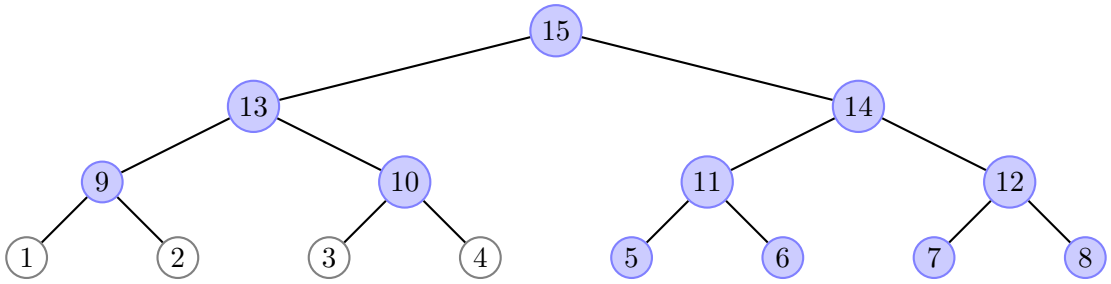
Proof. We see from condition (39) (or via direct computation from (38)) that a sufficient condition for $\tau_1 < \tau_2$, is to have, for all $i \in S_1$, $\dfrac{\omega_{i1}}{d_i} < \dfrac{d'_1}{d_1}$ or equivalently $\dfrac{\omega_{i1}}{d'_1} < \dfrac{d_1}{d_i}$. If $i \in S_1$ is the only node in S_1, then $\omega_{i1} = d'_1$ and we find the condition $d_i < d_1$. If however $|S_1| \geq 2$, we have $\dfrac{\omega_{i1}}{d'_1} \leq 1$, because $d'_1 = \omega_{i1} + \displaystyle\sum_{j \in S_1 \setminus \{i\}} \omega_{ij}$, and thus the condition on $\frac{d_1}{d_i}$ can be replaced by the simpler (but stronger) condition $d_i < d_1$, for all $i \in S_1$. □

6.3 A regular tree

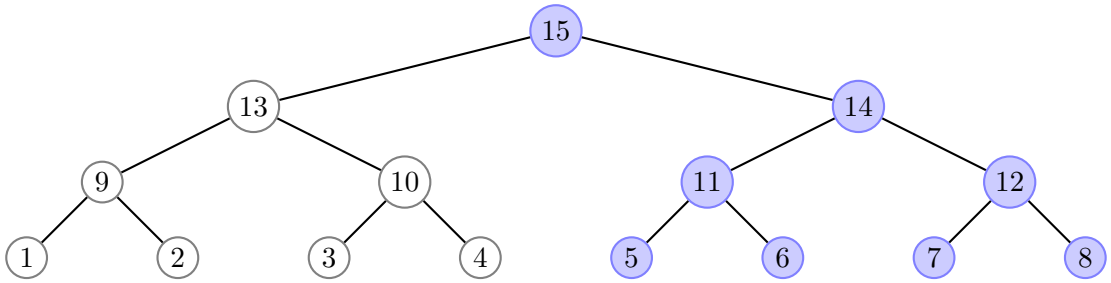
We consider the (MBO$_\tau$) iterations on a regular tree as in Figure 2. Let $\omega_{ij} = \omega$, for all $(i, j) \in E$, and $r = 1$. As in Figure 2a, we consider the case where the initial set S consists of the leaves of a branch. We first observe that the subgraph induced by $\overline{S_j}$, for any $j \in V$, is a star graph with node j as center, *i.e.*, $d'_i = 0$ for all $i \in S_j$ (for an example of a star graph with five nodes, see Figure 1b), so that the hypothesis of Lemma 6.1 are satisfied with nodes 9 and 10 each playing the role of "node 1" in the lemma.

Applying Lemma 6.1 to node 9 in Figure 2a where $S = \{1, 2, 3, 4\}$, we see that there exists a τ such that node 9 will change in the next iteration. By symmetry, node 10 will change in the same iteration. If node 13 doesn't change in the first MBO iteration, Lemma 6.1 can be applied again (because node 13 has two children, the hypotheses of the lemma are again satisfied with $9, 10 \in S_{13}$) to show that there exists a τ such that node 13 will be added to the set. After node 13 has been added to the set, S, as in Figure 2b the MBO iterates are stationary. To see that node 15 cannot be added to S, assume that it were. Then the value of the Lyapunov functional, (33), must have decreased. But by symmetry, in the next MBO iteration, node 15 will be removed from S, again decreasing the value of the Lyapunov functional, a contradiction. The final configuration in Figure 2b minimizes the normalized cut, as defined in Section 2.2.

This argument is easily generalized to trees where each node, excluding leaves, has the same number of children $c \geq 2$.



(a) Initial configuration



(b) Final configuration

Figure 2: The initial and final configurations for an evolution by the (MBO$_\tau$) scheme on a tree graph; see Section 6.3.

6.4 A small square grid

Here we construct an explicit example where Theorem 4.8 can be applied to show that there exists a time interval (τ_1, τ_2) such that a node is guaranteed to change in one iteration of the (MBO$_\tau$) algorithm.

Consider a three by three (nonperiodic) square grid as in Figure 1c with unit edge weights, with nodes numbered 1 through 9 from left to right, top to bottom. Let $S = \{4, 6, 7, 8, 9\}$. We focus on node 5. We have $\mathcal{N}_5 = \{2, 4, 6, 8\}$ and $S_5 = \mathcal{N}_5 \cap S = \{4, 6, 8\}$. We then compute $d'_5 = 3$, $d_5 = 4$, $d'_4 = d'_6 = d'_8 = 0$ and $d_4 = d_6 = d_8 = 3$. It is easily checked that, for $i \in S_5$,

$$\frac{d'_i}{d_i} = 0 \le \frac{3}{4} = \frac{d'_5}{d_5} \quad \text{and} \quad \frac{\omega_{i5}}{d_i} = \frac{1}{3} < \frac{3}{4} = \frac{d'_5}{d_5},$$

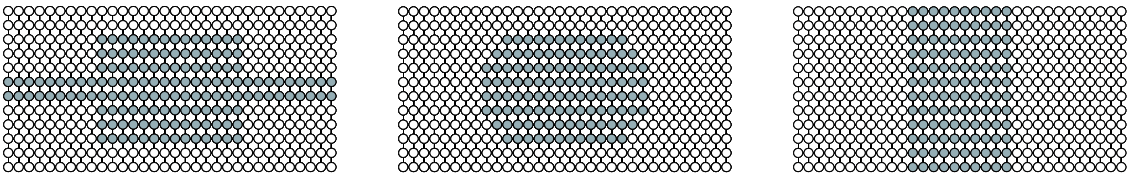
such that conditions (39) are satisfied. Furthermore, with $\overline{S_5} := S_5 \cup \{5\}$,

$$\sum_{k \in \overline{S_5}} \frac{d'_k}{d_k} \frac{\omega_{ik}}{d_i} = \begin{cases} 0 & \text{if } i = 5, \\ \frac{1}{4} & \text{if } i \in S_5, \end{cases}$$

thus $\left(\frac{d'_5}{d_5}\right)^2 - \max_{i \in \overline{S_5}} \sum_{k \in \overline{S_5}} \frac{d'_k}{d_k} \frac{\omega_{ik}}{d_i} = \left(\frac{3}{4}\right)^2 - \frac{1}{4} = \frac{5}{16} > 0$, and even the full condition (38), for $r = 1$, is satisfied. From (36) we can then compute

$$\tau_{1,2} = \frac{3/4}{1/4} \pm 4\sqrt{\frac{5}{16}} = 3 \pm \sqrt{5},$$

for the time interval (τ_1, τ_2) of Theorem 4.8.



(a) (b) (c)

Figure 3: Two (MBO$_\tau$) evolutions on the 2-torus graph, $T^2_{32,12}$. The top and bottom 'border' nodes are connected (not shown) as are the left and right 'border' nodes. **(a)** Initial condition. **(b)** For $\tau = 1.12$, the stationary state shown is reached in 4 iterations. **(c)** For $\tau = 4$, the stationary state shown is reached in 5 iterations. See Section 6.5.

6.5 Torus graph

Consider the n-cycle, C_n with n nodes. The nodes are arranged in a circle and each node is connected to its 2 neighbors. We take $\omega_{ij} = \omega$ for $i \sim j$ and zero otherwise. See Figure 1d. We consider the unnormalized graph Laplacian $L = D - A$ ($r = 0$ in (8)). In this case, L is a circulant matrix $\mathrm{diag}(\{-1, 2, -1\}, \{-1, 0, 1\})$. The eigenpairs $\{(\lambda_j, v^j)\}_{j=1}^n$ are given by

$$\begin{aligned} \lambda_j &= 2\omega - 2\omega \cos \frac{2\pi(j-1)}{n} \\ v^j_i &= \exp\left(2\pi i(j-1)/n\right). \end{aligned}$$

We then consider the 2-torus graph, $T^2_{n_1,n_2}$ which is the Kronecker (tensor) product of the n_1- and n_2-cycles. See Figure 3. In particular, if u and v are eigenfunctions of the graph Laplacian on C_{n_1} and C_{n_2} with corresponding eigenvalues α and β respectively, then $w = u \otimes v$ (with $w_{i,j} = u_i v_j$) is an eigenvector of $T^2_{n_1,n_2}$ with corresponding eigenvalue $\alpha + \beta$. In particular, the spectral radius of the Laplacian is $\rho = 8\omega$.

Consider for a moment $T^2_{n_1,n_2}$ as a discretization of the torus, $\mathbb{T}^2$. The nontrivial minimal-perimeter subsets of $\mathbb{T}^2$ are given by "strips". Thus we might expect that for some initial condition, χ_S, $S \subset V$ the evolution by MBO, Allen-Cahn, or MC would converge to a strip.

We consider the (MBO$_\tau$) evolution on a 32×12 torus with $\omega = 1$ and initial condition, as in Figure 3a. For $\tau = 1.12$, the solution is stationary after 4 iterations once the "high curvature corners" have been removed, as in Figure 3b. For $\tau = 4$, the solution evolves into a minimal-perimeter "strip" in 5 iterations, as in Figure 3c.

For the parameters in Figure 3b, we compute the guaranteed stationarity bounds in (30) and (31) to be $\tau_\rho \approx 0.0057$ and $\tau_\kappa = \frac{1}{4}$, respectively, showing these bounds are not sharp. figs Consider (MCF$_{\eth t}$), with S_n equal to the minimal-perimeter strip in Figure 3c. Then $S_{n+1} = S_n$ is a minimizer of $\mathcal{F}(\cdot, S_n)$, but so are $S_{n+1} = S_n \cup \partial(S_n^c)$ and $S_{n+1} = S_n \setminus \partial S_n$ (or variations in which only one 'vertical line' of the boundary is added or removed). This illustrates a possible type of non-uniqueness for (MCF$_{\eth t}$), which occurs when S_n is totally geodesic (*i.e.*, its boundary is a geodesic). To reiterate, the stationary solution in Figure 3b is frozen (due to the smallness of τ), while the solution in Figure 3c is totally geodesic.

6.6 Buckyball graph

Consider the buckyball graph with 60 nodes and 90 edges with $\omega_{ij} = \omega$ for all edges (i, j) as in Figure 4. The graph is regular; each node has degree 3ω.

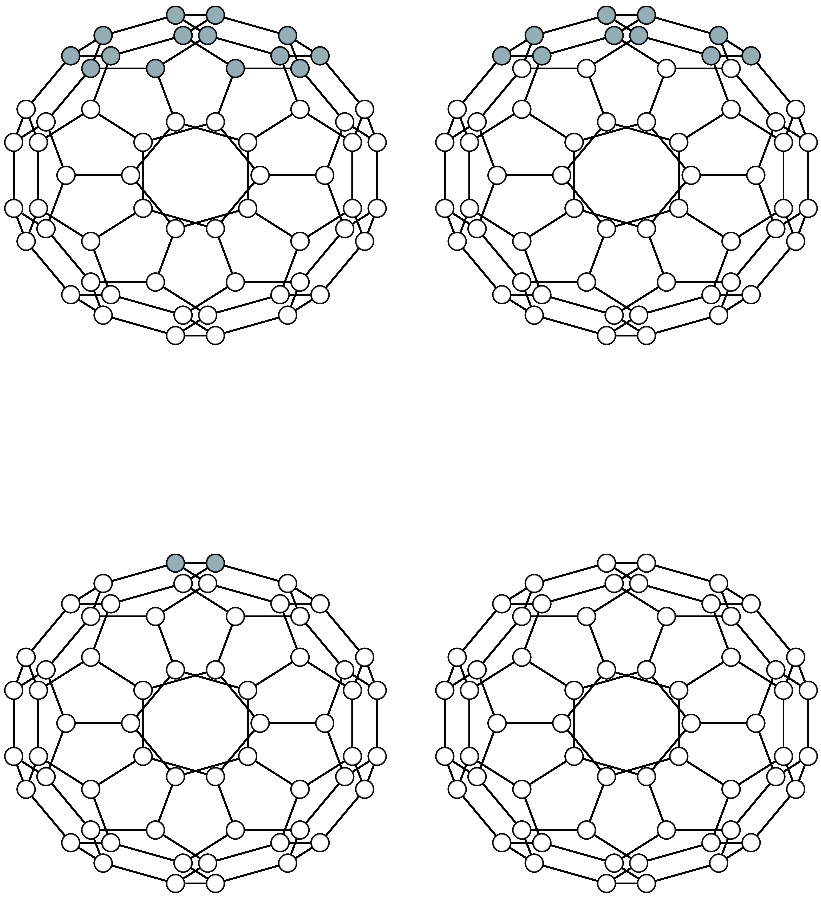


Figure 4: An (MBO$_\tau$) evolution with $\tau = 2$ on the buckyball graph. The solution at each iteration is the characteristic function of the gray nodeset. See Section 6.6.

Consider for a moment the buckyball graph as a (coarse) discretization of the sphere, $\mathbb{S}^2$. There are no nontrivial minimal-perimeter subsets of $\mathbb{S}^2$. Great circles are the only nontrivial stationary submanifolds of $\mathbb{S}^2$ (and have constant curvature). In fact, great circles are totally geodesic. Thus we might expect that for any initial condition, χ_S, $S \subset V$ such that $|S| \neq |\mathbb{S}^2|/2$, the evolution by MBO, Allen-Cahn, or MC would converge to a stationary solution, either 0 or χ_V depending on the initial mass, $M(\chi_S) = \text{vol } S$. If S is chosen to be a symmetric partitioning of the nodes for the buckyball graph, we expect that the (MBO$_\tau$) evolution will be stationary for all values of τ.

The bound from Theorem 4.2 states that pinning in (MBO$_\tau$) occurs if

$$\tau < \rho^{-1} \log\left(1 + \frac{1}{2}|S|^{-\frac{1}{2}}\right).$$

The bound in Theorem 4.3 states that trivial dynamics occur if

$$\tau > \lambda_2^{-1} \log \left(\frac{(3\omega)^{\frac{r}{2}} |S|(n-|S|)}{\left||S| - \frac{n}{2}\right|} \right).$$

We find numerically that $\lambda_2 \approx \omega^{1-r} 3^{-r} \cdot 0.2434$ and $\lambda_n \approx \omega^{1-r} 3^{-r} \cdot 5.6180$.

For initial condition χ_S, with $|S| = 14$, as given in Figure 4 (top left), and $r = 0$ and $\omega = 1$, Theorems 4.2 and 4.3 predict that pinning occurs if $\tau < 0.0223$ and trivial dynamics occur if $\tau > 15.1811$. We find numerically that this initial condition is pinned if $\tau < 1.89$ and trivial dynamics occur if $\tau > 3.54$. For intermediate values of τ, the iterates shrink to the empty node set. For $\tau = 2$, the iterates take 3 iterations to reach steady state, as illustrated in Figure 4.

For the initial condition χ_S where S is taken to be a symmetric partitioning of the nodes, (MBO$_\tau$) evolution is pinned for all values of τ.

6.7 Adjoining regular lattices

We consider the graph which is composed by adjoining a square and triangular lattice. See Figure 5. We take $r = 0$ and $\omega_{ij} = 1$ for $i \sim j$ and zero otherwise. Note that the degree of a node in the triangular lattice is 6 and the degree of a node in the square lattice is 4.

To test the intuition from the star graph (see Section 6.2) that it is easier for the solution to pin on nodes with smaller degree, we consider the initial condition given in the top left panel of Figure 5. The mass is initially distributed over both the square and triangular lattice sites. We consider the (MBO$_\tau$) evolution with $\tau = 0.8$. The solution moves freely on the lattice sites with degree > 4, *i.e.*, on the triangular lattice. However, on the square lattice, the solution only 'rounds corners'.

The nodes on the 'border' of the graph (where the regular lattice was cut) have smaller degree. In Figure 6, we demonstrate that the solution can also be pinned on the border. Again, the initial condition is given in the top left panel. In this simulation, we take $\tau = 0.9$. Away from the boundary, the solution set can again shrink freely. However, the solution becomes pinned on the border.

6.8 Two moons graph

In this last example, we consider a graph which is widely used as a benchmark problem for partitioning algorithms. Our construction of the graph follows [BH09]. The graph is generated by first randomly distributing 600 points in a region described by two half arcs in $\mathbb{R}^2$—referred to as "two moons". See Figure 7 (top left). The points are then embedded in $\mathbb{R}^{100}$ and randomly perturbed by i.i.d. Gaussian noise with mean zero and standard deviation $\sigma = 0.1$. Let $k = 10$. The edge weights are chosen to be

$$w_{ij} = \max\{s_i(j), s_j(i)\}, \quad \text{where } s_i(j) = e^{-\frac{4}{d_i^2}\|x_i - x_j\|^2},$$

and d_i is the Euclidean distance between x_i and its k-th nearest neighbor. We then take the symmetrized k-nearest neighbors graph. This is given in Figure 7 (top right).

We consider the (MBO$_\tau$) evolution with $\tau = 5$ and initial condition as shown in Figure 7 (bottom left). After 9 iterations, the (MBO$_\tau$) evolution converges to the state in Figure 7 (bottom right).

We want to stress that the two moons example is meant as an illustration of the (MBO$_\tau$) algorithm on a more complex toy graph. In this paper we do not aim to compete in terms of accuracy or efficiency with existing clustering methods, hence we will not focus on those aspects of the two moons example.

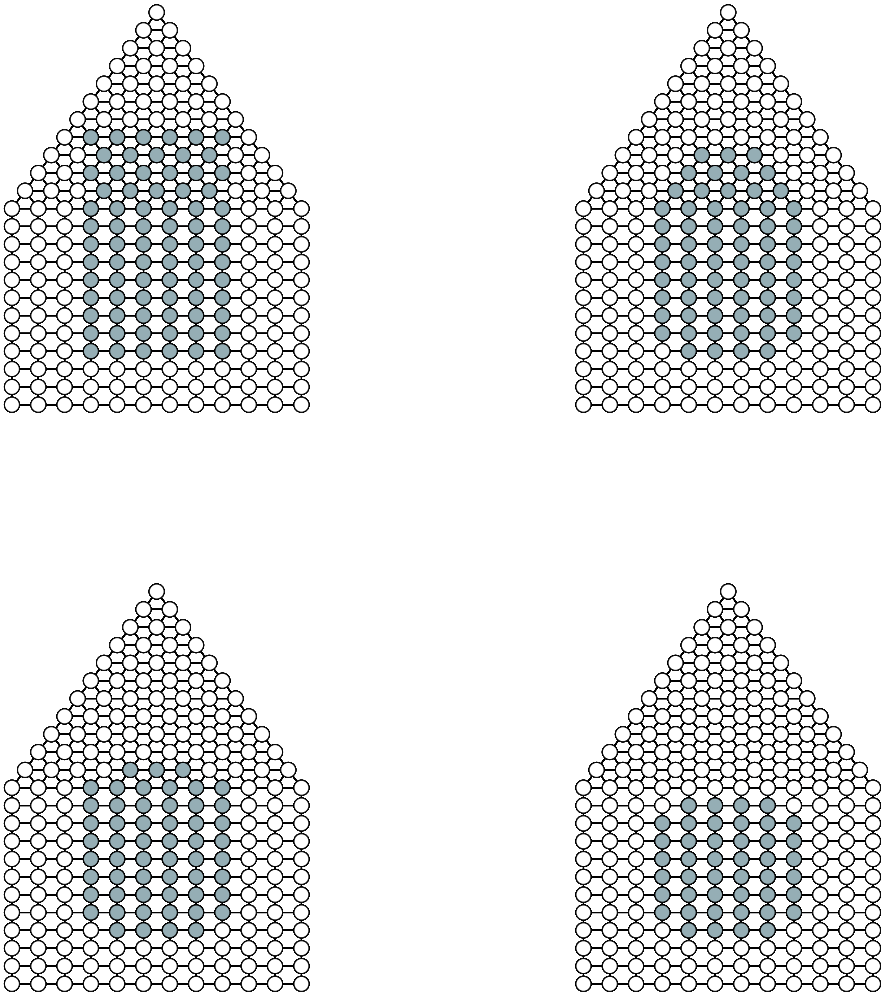


Figure 5: An (MBO$_\tau$) evolution with $\tau = 0.8$ on a graph consisting of adjoining regular lattices. The solution at chosen iterations is the characteristic function of the gray nodeset. For the initial condition, given by the top left panel, the evolution reaches a steady state in 9 iterations. Iterations 3 (top right), 6 (bottom left), and 9 (bottom right) are shown. This example strengthens the 'rule of thumb' that it is easier for a solution to become pinned on nodes with smaller degree. See Section 6.7.

7 Discussion and open questions

Motivated by curvature flows in continuum mechanics, we described several analogous processes on graphs. In particular we used the graph total variation, or graph cut, to define curvature on graphs, which we then related to the graph Allen-Cahn equation, graph MBO scheme, and graph mean curvature flow. The continuum intuition for these problems suggests many results, some of which we proved in this paper, some which we have shown cannot hold on a graph because of the lack of infinitesimal length scales, and some which we state below as, still unproven, open

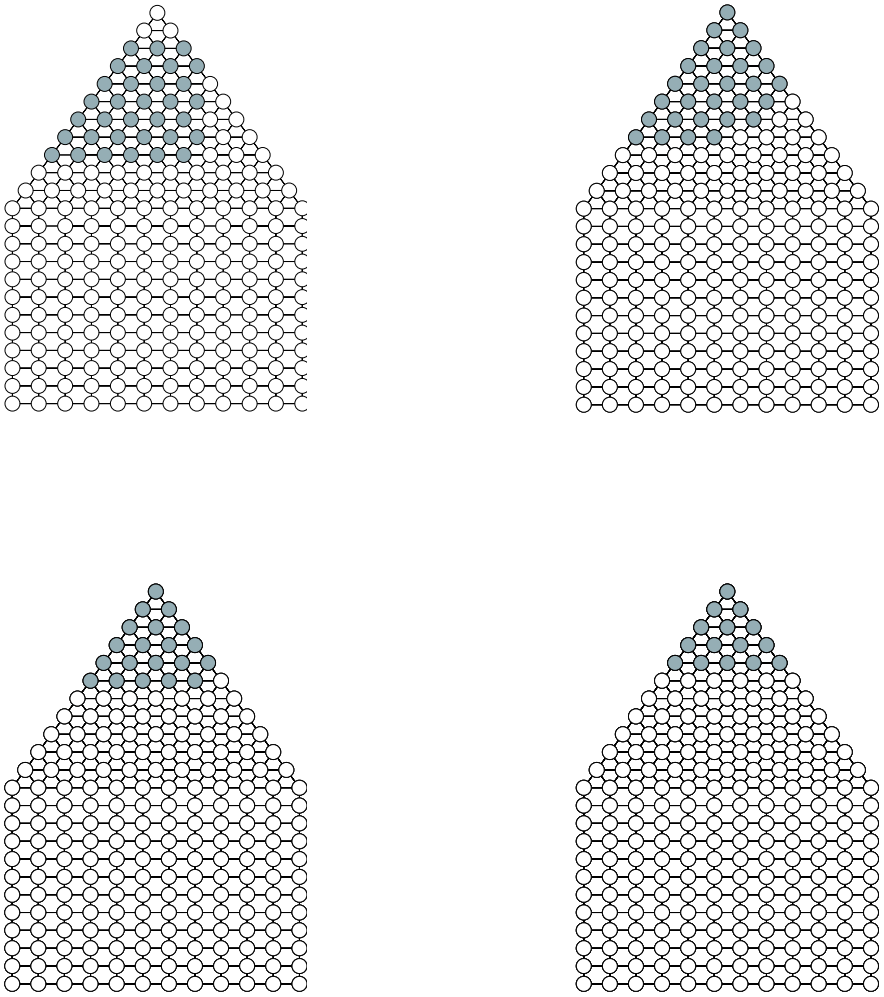


Figure 6: An (MBO$_\tau$) evolution with $\tau = 0.9$ on a graph consisting of adjoining regular lattices. The solution at chosen iterations is the characteristic function of the gray nodeset. For the initial condition, given by the top left panel, the evolution reaches a steady state in 13 iterations. Iterations 4 (top right), 9 (bottom left), and 13 (bottom right) are shown. This example strengthens the 'rule of thumb' that it is easier for a solution to become pinned on nodes with smaller degree. See Section 6.7.

questions.

In a sense to be made precise, for a suitable choice of τ (not too small, not too large, depending on ω, most likely depending on the graph's spectrum), the dynamics of (MBO$_\tau$) are expected to approximate those of graph MCF.

Question 7.1 (MBO and graph Mean Curvature Flow)**.** *Is there an interval of τ (depending on $\eth t$), for which a single (MBO$_\tau$) iteration minimizes the (MCF$_{\eth t}$) functional $\mathcal{F}$ from (25)? For such a τ, the graph mean curvature flow (MCF$_{\eth t}$) would coincide with the (MBO$_\tau$) scheme (up to a time rescaling).*

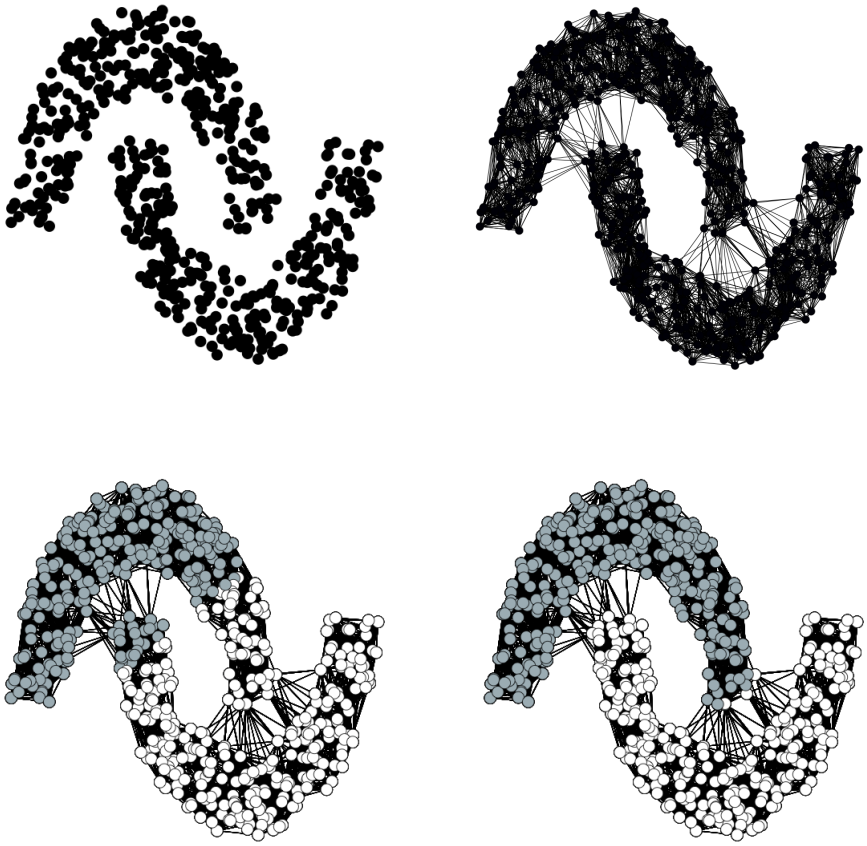


Figure 7: (top) Construction of the two moons graph. (top left) Some random points in the shape of two moons. (top right) The connectivity of the graph resulting from connecting nearest neighbors after adding high dimensional noise. (bottom) An (MBO_τ) evolution for $\tau = 5$, starting with initial condition on the left and terminating at the stationary solution on the right in 9 iterations. See Section 6.8.

An approach to Question 7.1, uses the Taylor series expansion for the solution of the graph heat equation: $e^{-t\Delta}\chi_S = \sum_{k=0}^{\infty} \frac{1}{k!} (-t\Delta)^k \chi_S$, for $S \subset V$. Hence, we can rewrite the Lyapunov functional J from (33) as

$$J(\chi_S) = \langle 1 - \chi_S, \chi_S - \tau\Delta\chi_S\rangle_{\mathcal{V}} + R_S(\tau), \quad \text{where } R_S(\tau) := \sum_{k=2}^{\infty} \frac{(-\tau)^k}{k!} \langle \chi_{S^c}, \Delta^k \chi_S\rangle_{\mathcal{V}}.$$

Using $\langle 1 - \chi_S, \chi_S\rangle_{\mathcal{V}} = 0$, $\langle 1, \Delta\chi_S\rangle_{\mathcal{V}} = 0$, and (3), we find

$$J(\chi_S) = \tau \mathrm{TV}_{\mathrm{a}}^1(\chi_S) + R_S(\tau).$$

This connection between the Lyapunov functional J and the total variation, and hence the MCF functional $\mathcal{F}$ from (25), strengthens the plausibility of a positive answer to Question 7.1. A more difficult question, which could be of great use in numerical problems, is how we can estimate the number iterations of (MBO_τ) needed to go from some initial data to a minimizer of the Ginzburg-Landau functional or graph cut functional.

Question 7.2 (Minimizing graph cut)**.** *For any, a priori specified, approximation error, is there a local, quantitative bound on the number of iterations of* (MBO_τ) *needed to approximate a minimizer of the graph cut functional* TV_a^q *up to the specified accuracy? "Local" means here that the bound does not rely on the spectrum of the graph, but instead uses quantities such as graph curvature* $\kappa^{q,r}$ *or the total variation* $TV_a^q(\chi_N)$ *for some local graph neighborhood* $N \subset V$*. The analogous question can be asked for* ($\text{MCF}_{\eth t}$)*.*

In Theorem 4.8 it was shown that, if the curvature at a given node is sufficiently large and the time step τ in (MBO_τ) is chosen in the right interval, then the value at the node will change in one (MBO_τ) iteration. The next question is the analogous statement for the Allen-Cahn equation, (ACE_ε).

Question 7.3 (Non-freezing for Allen-Cahn)**.** *Let* ε *be in some positive interval and let* u^ε *be a solution of the Allen-Cahn equation* (ACE_ε) *for this choice of* ε*. Suppose that the curvature* $(\kappa_S^{1,r})_i$ *of* $S = \{j \mid u_j^\varepsilon(t_0) \leq 0\}$ *at a node* $i \in S$*, is sufficiently large (possibly depending on* ε*). Is there is some interval of positive times such that* $u_i^\varepsilon(t_0 + h) > 0$ *for* h *in this interval?*

Because (ACE_ε) is derived from the graph functional (GL_ε), we suspect that the correct curvature in Question 7.3 is $\kappa_S^{1,r}$, the curvature related to the anisotropic functional $\frac{1}{2}\text{TV}_\text{a}^1$, which was identified as the Γ-limit of (GL_ε) for $\varepsilon \to 0$ in earlier work [vGB12], and not the curvature which can be derived from the isotropic total variation functional TV as the continuum case might suggest at first glance. Since we have seen that pinning occurs for small enough ε, full convergence is not expected here, but the numerical examples of Section 6 suggest an approximate result for small ε is feasible.

Question 7.4 (Allen-Cahn and graph Mean Curvature Flow)**.** *Is there an* $\varepsilon > 0$ *such that, given the solution* u^ε *to* (ACE_ε) *for some* $\epsilon > 0$*, there is an increasing sequence of times* t_n *for which either the sets* $S_n := \{j \mid u_j^\varepsilon(t_n) \leq 0\}$ *or the sets* $S_n := \{j \mid u_j^\varepsilon(t_n) \geq 0\}$ *form a solution to the graph MCF?*

Furthermore, among sequences with this property is there exactly one sequence $\{t_n\}$ *that is maximal in the following sense: there exists no sequence* $\{t'_n\}$*, of which* $\{t_n\}$ *is a strict subsequence, such that* $\{S_{t'_n}\}_n \not\subset \{S_{t_n}\}_n$ *and* $\{S_{t'_n}\}$ *is still a solution to the graph MCF?*

A different question is how the graph MCF behaves in the continuum limit, when it is formulated on a sequence of graphs which are ever finer discretizations of some continuum space. We expect that it should give back the usual MCF in the continuum limit, or some anisotropic MCF, as the convergence results in [vGB12] show the final limit could crucially depend on the scaling in $\eth t$ and the discretization parameter (which will show up in the graph weights). This question is similar (and perhaps equivalent) to the convergence of discretization schemes for the usual MCF. Similar questions can be asked about the graph Allen-Cahn equation and graph MBO scheme.

Question 7.5 (Stability of graph MCF, MBO, and ACE, in the continuum limit)**.** *Suppose we are given any sequence of graphs* (V^k, ω_{ij}^k)*,* $k \in \mathbb{N}$*, converging in the Gromov-Hausdorff sense to a Riemannian manifold* (M, g)*. Is there a fixed time interval such that, as* $k \to \infty$*, any sequence generated by* (MBO_τ) *with* τ *in this interval converges to a sequence generated by the (possibly anisotropic) continuum MBO algorithm in* M *(with the Laplacian induced by* g*)? Accordingly, do solutions of* (ACE_ε) *converge to solutions of the (possibly anisotropic) continuum Allen-Cahn equation in* M*, and do solutions to* ($\text{MCF}_{\eth t}$) *converge to viscosity solutions (via the level set formulation) of (possibly anisotropic) continuum MCF in* M*, with initial data given by the limit of the initial data in each* V^n*?*

As explained in Appendix A, MCF is closely related to certain models of continuum phase transitions, particularly Allen-Cahn and Ginzburg-Landau dynamics. However, another important connection with statistical mechanics involves Ising models and other interacting particle systems, which are known to converge in the mesoscopic limit to flow by mean curvature. In work of Katsoulakis and Souganidis [KS94, KS97] convergence to a viscosity solution of MCF is first proved. See also the related work of Funaki and Spohn [FS97] where MCF is derived as a deterministic limit of stochastic Ginzburg-Landau dynamics. On the other hand, there is vast literature concerned with (for instance) the Ising model (and its generalizations) on graphs [Lyo89, Lyo00], see also Durrett's book [Dur07]. This suggests the following question.

Question 7.6 (Possible probabilistic interpretations of graph MCF, MBO, and AC)**.** *Is (MCF$_{\partial t}$) related to an interacting particle system on the underlying graph? Also, are there interacting particle systems or stochastic processes in V that are related to (MBO$_\tau$) or (ACE$_\varepsilon$)?*

Finding such a system would partly resolve the issue that a front moving on a graph in continuum time necessarily does so in a way that, from a continuum point of view, looks discontinuous (as discussed previously in this paper, in particular in Sections 3.3 and 5.), as the particle dynamics would be continuous in time and stochastic The convergence results in [KS94, KS97, FS97] show that the above question has an a priori higher chance of having a positive answer for a large graph, as it already holds in the continuum limit.

Acknowledgements

This work was supported by the W. M. Keck Foundation, UC Lab Fees Research Grant 12-LR-236660, ONR grants N000141210838 and N000141210040, AFOSR MURI grant FA9550-10-1-0569, and NSF grants DMS-1118971, DMS-0968309, and DMS-0914856. Braxton Osting is supported in part by a National Science Foundation (NSF) Postdoctoral Fellowship DMS-1103959. Nestor Guillen is supported in part by National Science Foundation (NSF) grant NSF-DMS-1201413. We would like to acknowledge helpful conversations with Jérôme Darbon, Selim Esedoḡlu, Inwon Kim, Tijana Kostić, Stanley J. Osher, and Richard Tsai.

A The continuum case

In this appendix, we briefly review and provide references for the Allen-Cahn equation, the MBO algorithm, and mean curvature flow in the continuum setting.

A.1 The continuum Allen-Cahn equation

The Allen-Cahn equation is a reaction-diffusion equation, given by

$$u_t = \Delta u + f(u), \tag{43}$$

where $u\colon \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ and Δ is the standard Laplacian (although other linear elliptic operators can be considered as well), and f is a non-linear function of the form $f = -W'$ where $W\colon \mathbb{R} \to \mathbb{R}$ is a double well potential with two global minima. For simplicity, take $W(u) = (u+1)^2(u-1)^2$, where the minima are at ± 1.

A question which is always of interest is understanding the way that solutions to (43) converge to equilibrium. For each fixed x, one expects that $u(x,t)$ approaches either 1 or -1, as $t \to +\infty$, as these values correspond to the minima of W. This indicates that for very large t the function u defines two regions of $\mathbb{R}^n$, where it is very close to either 1 or to -1, separated with a smooth transition layer in between.

This asymptotic behavior is well understood nowadays. Rescaling (x,t) as $(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2})$, we obtain the equation

$$u_t^\varepsilon = \Delta u^\varepsilon + \varepsilon^{-2} f(u^\varepsilon). \tag{44}$$

Note that for very small ε the function u^ε describes the long time behavior of the original u. Then, it is well known (see [BSS93, BK91] for background and discussion) that, as $\varepsilon \to 0^+$, the solutions $u^\varepsilon(x,t)$ converge to a function which takes the value -1 in some set S_t (depending on time) and takes the value 1 in S_t^c. Here S_t is a set whose boundary is evolving by mean curvature flow (see Section A.2).

Although the original motivation for studying (43) was phase transitions, it is also the gradient flow of the Ginzburg-Landau functional. Precisely, equation (44) is the L^2 gradient flow of the functional

$$GL_\varepsilon(u) := \int \frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon} W(u)\, dx. \tag{45}$$

It is expected that solutions to (44) converge to a local minimum of this functional, as $t \to +\infty$, thus schemes for (44) could be used for approximating minima of (45). This is the application that serves as the biggest motivation in the graph setting.

For more information about reaction-diffusion equations with a polynomial nonlinearity we refer to [Tem97, Section 1.1].

A.2 Continuum mean curvature flow

Mean curvature flow (MCF) consists of the evolution of a closed, oriented hypersurface $\Sigma_t \subset \mathbb{R}^d$ over time, such that the inner normal velocity at a given point of Σ_t is equal to the mean curvature of Σ_t at that point. The study of such a flow has been greatly motivated by phase transition models in crystal growth and materials science, in particular since the important work of Allen and Cahn [AC79]. Starting with the seminal work of Brakke [Bra78], the mathematical study of this flow has been vast, and has involved areas of mathematics ranging from differential geometry to stochastic control. The use of MCF is now widespread in the modeling of moving fronts [CS94]. The reason why MCF is so ubiquitous in the phase transitions literature, is that many singular limits of reaction diffusion equations (*i.e.,* singular limits of Ginzburg-Landau dynamics) converge to motion by mean curvature. See [BK91, Peg89, BG95] for precise convergence theorems and further discussion.

A well known feature of MCF is both the formation of singularities and the occurrence of topological changes, regardless of the smoothness of the initial data. A significant portion of the literature on MCF deals with notions of weak solutions, the first of which goes back to Brakke [Bra78]. Partial regularity for weak solutions as well as regularity up to the first singular time have been widely studied [Eck04].

An equivalent formulation of the flow looks not only at the hypersurface Σ_t, but at the entire domain Ω_t bounded by it, so that $\partial\Omega_t = \Sigma_t$. Accordingly, it is said that Ω_t itself is evolving by mean curvature flow. This perspective is natural for phase transitions.

Let $\phi(t,\cdot)\colon \mathbb{R}^d \to \mathbb{R}$ be the signed distance function to the set Ω at time t. From the level set method perspective [OF03], the motion by mean curvature[16] of Ω_t corresponds to an initial value problem for a fully non-linear degenerate parabolic equation,

$$\phi_t = F(D^2\phi, \nabla\phi), \quad \phi(.,0) = \phi_0, \quad \text{where } F(D^2\phi, \nabla\phi) = -|\nabla\phi|\mathrm{div}\frac{\nabla\phi}{|\nabla\phi|}. \tag{46}$$

[16] In the literature two related, but different, concepts of mean curvature appear. One corresponds with the factor $\mathrm{div}\frac{\nabla\phi}{|\nabla\phi|}$ in (46), the other has a normalization factor $\frac{1}{d-1}$, where d is the dimension of the space. This normalization by the dimension of the hypersurface justifies the "mean" part of "mean curvature".

Then, when there is a smooth solution $\phi(x,t)$, the domains given by $\Omega_t := \{\phi(\cdot,t) < 0\}$ will be evolving by mean curvature flow and will start from the original domain Ω. In general, even for an initial domain with a smooth boundary, a smooth solution might not exist for all times, and one must work with viscosity solutions. In this context, the convergence of the MBO scheme (47) (explained in Section A.3) to such viscosity solutions was proved by Evans [Eva93].

It is worth remarking that Soner and Touzi in [ST03] interpret MCF as a stochastic control problem. In this interpretation, one controls a Brownian motion for which one is allowed to turn off diffusion in one given direction. The surface Σ_t in this case arises as the set of points that can be reached with probability 1. This probabilistic interpretation is quite different from those mentioned in the discussion at the end of Section 7.

Finally, given the affinity with the graph setting, it is worthwhile to comment briefly on the more recent nonlocal mean curvature flow. Caffarelli and Souganidis [CS10] arrive at this flow by following a nonlocal and continuum analogue of (MBO$_\tau$), where instead of using the Laplacian one uses a fractional power of the Laplacian $(-\Delta)^s$ with $s \in (0, 1/2)$. A level set formulation based on viscosity solutions was developed later by Imbert [Imb09].

A.3 The continuum MBO algorithm

The Merriman, Bence, and Osher (MBO) algorithm [MBO92, MBO93, MBO94], also known as the threshold dynamics algorithm, approximates the dynamics of mean curvature flow (46) by alternatively applying diffusion and thresholding operators. Let $\chi(t,\cdot)$ be the characteristic function of the set Ω_t at time t. Define the diffusion operator $\chi_0 \mapsto u(t,\cdot) := e^{-t\Delta}\chi_0$ to be the solution of the initial value problem

$$\dot{u} = -\Delta u, \quad u(0) = \chi_0(\cdot).$$

Define the threshold operator

$$Pu(x) = \begin{cases} 1 & u(x) \geq \frac{1}{2} \\ 0 & u(x) < \frac{1}{2} \end{cases}.$$

The MBO evolution of a set described by u at time T can then be succinctly written

$$\chi(T,\cdot) = \left(Pe^{-\tau\Delta}\right)^k \chi_0, \qquad \text{where } \tau = T/k \tag{47}$$

is the 'time step' and k is a parameter. In [Eva93, BG95] convergence of the MBO algorithm to motion by mean curvature, defined in (46), as $k \uparrow \infty$, is proven.

The MBO scheme and its implementation has evolved considerably since its original proposal. We provide a small, non-exhaustive, overview here. In [Mas92, Ruu98a], the MBO scheme was extended to multiple-phase problems. In [Ruu96, Ruu98b], a spectral discretization of the MBO scheme for motion by mean curvature was proposed, which is much more efficient then finite difference approaches. This approach can be applied to both two-phase and multiphase problems. In [ERT08], diffusion generated motion was applied to higher order geometric motions. In [ET06], the MBO scheme was extended to a thresholding method for approximating the evolution by gradient descent of the Mumford-Shah functional and applied to image segmentation problems. In [ERT10] the authors study MBO-like schemes which use the signed distance function. Recent work [EO13] presents new algorithms for multiphase mean curvature flow, based on a variational description of the MBO scheme.

It is well-known that, in a finite difference scheme for the MBO algorithm, the time step τ (equivalently k) in (47) must be chosen carefully and in the limit as $k \uparrow \infty$, the discretized MBO evolution is stationary. In fact, when discussing the numerical implementation of the algorithm on a discrete grid, Merriman, Bence, and Osher [MBO92] observe:

> "The basic requirement is that [the time step, τ,] be short enough so that the local analysis ... is valid, but also long enough so that the boundary curve moves by at least one grid cell on the spatial grid (otherwise the curve would be stuck)."

They derive heuristic upper and lower bounds on the time step, τ, for the algorithm to approximate motion by mean curvature.

B Coarea formula for anisotropic graph total variation

In this appendix we prove a layer cake or coarea formula for the discrete total variation in $\mathrm{TV}_{\mathrm{a}}^q$. Such a formula is useful when trying to minimize the total variation over binary valued functions. The formula shows that such minimization can be achieved, by first minimizing over real-valued functions and then selecting (almost any) sublevel set of the resulting minimizer. This will give a binary minimizer of the total variation. Minimizing the total variation by itself gives a trivial solution, so the usefulness of this layer cake formula hinges on the ability to find a similar formula for the fidelity term or constraint that is applied to the minimization. An example where this technique is used for the continuum isotropic total variation is [CEN06]. In [CvGO11] it is used for a continuum anisotropic total variation.

Recall from Section 2 that

$$\begin{aligned}\mathrm{TV}_{\mathrm{a}}^q(u) &:= \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} \colon \varphi \in \mathcal{E},\ \|\varphi\|_{\mathcal{E},\infty} \le 1\} \\ &= \langle \nabla u, \varphi^a)\rangle_{\mathcal{E}} = \frac{1}{2}\sum_{i,j\in V} \omega_{ij}^q |u_i - u_j|.\end{aligned}$$

where a particular choice of φ^a is given in (1): $\varphi^a := \operatorname{sgn}(\nabla u)$. Since, for given $u \in V$, and for all $i, j \in V$, $(u_i - u_j)\varphi_{ij}^a \le 0$, an equivalent definition of $\mathrm{TV}_{\mathrm{a}}^q$ is

$$\mathrm{TV}_{\mathrm{a}}^q(u) = \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} \colon \varphi \in \mathcal{E}_-(u)\},$$

where

$$\mathcal{E}_-(u) := \{\phi \in \mathcal{E} \colon\ \|\phi\|_{\mathcal{E},\infty} \le 1,\ \forall i, j \in V\ (u_i - u_j)\phi_{ij} \le 0\}.$$

Lemma B.1. *Let $u \in \mathcal{V}$ and define*

$$E(t) := \{i \in V \colon u_i > t\}.$$

Then

$$TV_a^q(u) = \frac{1}{2}\sum_{i,j\in V} \omega_{ij}^q |u_i - u_j| = \frac{1}{2}\int_{\mathbb{R}} \sum_{i,j\in V} \omega_{ij}^q \left|\left(\chi_{E(t)}\right)_i - \left(\chi_{E(t)}\right)_j\right| dt.$$

Proof. Fix $u \in \mathcal{V}$. Note that

$$u_i = \int_{\mathbb{R}} \chi_{[0,u_i]}(t)\, dt = \int_{\mathbb{R}} \left(\chi_{E(t)}\right)_i dt, \tag{48}$$

from which it immediately follows that

$$\sum_{i,j\in V} \omega_{ij}^q |u_i - u_j| = \sum_{i,j\in V} \omega_{ij}^q \left|\int_{\mathbb{R}} \left[\left(\chi_{E(t)}\right)_i - \left(\chi_{E(t)}\right)_j\right] dt\right| \le \int_{\mathbb{R}} \sum_{i,j\in V} \omega_{ij}^q \left|\left(\chi_{E(t)}\right)_i - \left(\chi_{E(t)}\right)_j\right| dt.$$

To prove the opposite inequality, we follow closely the argument used to prove the analogous result in the continuum case, in [EG92, 5.5 Theorem 1]. Let

$$k(t) := \sum_{i \in V \setminus E(t)} \sum_{j \in V} \omega_{ij}^q |u_i - u_j| = \sum_{i \in V\,:\,u_i \le t} \sum_{j \in V} \omega_{ij}^q |u_i - u_j|.$$

Since k is a monotone nondecreasing function, the derivative

$$k'(t) = \lim_{r \to 0} \frac{1}{r} \sum_{i \in V\,:\,t < u_i \le t+r} \sum_{j \in V} \omega_{ij}^q |u_i - u_j| \tag{49}$$

exists almost everywhere. Hence

$$\int_{\mathbb{R}} k'(t)\,dt \le \sum_{i,j \in V} \omega_{ij}^q |u_i - u_j|. \tag{50}$$

We claim that, for $\varphi \in \mathcal{E}_-(u)$,

$$\int_{\mathbb{R}} \sum_{i,j \in V} \omega_{ij}^q \left(\chi_{E(t)}\right)_i (\varphi_{ji} - \varphi_{ij})\,dt \le \int_{\mathbb{R}} k'(t)\,dt. \tag{51}$$

If this claim is true, then by taking the maximum over all $\varphi \in \mathcal{E}_-$, just as in the discussion immediately preceding this lemma, it follows that

$$\int_{\mathbb{R}} \sum_{i,j \in V} \omega_{ij}^q \left| \left(\chi_{E(t)}\right)_i - \left(\chi_{E(t)}\right)_j \right| dt \le \int_{\mathbb{R}} k'(t)\,dt.$$

By (50) the result then follows.

We now prove the claim (51). First note that

$$\sum_{i,j \in V} \omega_{ij}^q \left(\chi_{E(t)}\right)_i (\varphi_{ji} - \varphi_{ij}) = \sum_{i \in E(t)} \sum_{j \in V} \omega_{ij}^q (\varphi_{ji} - \varphi_{ij}).$$

Fix $t \in \mathbb{R}$ and $r > 0$, and define the function $\eta \in C(\mathbb{R})$ as

$$\eta(s) = \begin{cases} 0 & \text{if } s \le t, \\ \frac{s-t}{r} & \text{if } t \le s \le t+r, \\ 1 & \text{if } s \ge t+r. \end{cases}$$

Clearly

$$\eta'(s) = \begin{cases} 0 & \text{if } s < t \text{ or } s > t+r, , \\ \frac{1}{r} & \text{if } t < s < t+r. \end{cases}$$

Let $\phi \in C^\infty(\mathbb{R})$ be a mollifier with $\operatorname{supp}\phi \in [-1,1]$ and $\int_{\mathbb{R}} \phi = 1$, and define for $\delta > 0$, $\phi_\delta(x) := \delta^{-1}\phi\left(\frac{x}{\delta}\right)$. Let $\eta_\delta := \phi_\delta * \eta$, then $\eta_\delta \in C^\infty(\mathbb{R})$, $\eta_\delta \to \eta$ a.e. as $\delta \to 0$, and $\eta_\delta \to \eta$ uniformly on compact subsets of $\mathbb{R}$ as $\delta \to 0$ (*e.g.*, [Eva10, Appendix C, Theorem 7]). The Taylor series with remainder gives us that for all $u_i, u_j \in \mathbb{R}$ there exists a $\xi_{ij} \in [u_i, u_j] \cup [u_j, u_i]$ such that

$$\eta_\delta(u_j) = \eta_\delta(u_i) + \eta_\delta'(u_j)(u_j - u_i) + \frac{1}{2}\eta_\delta''(\xi_{ij})(u_j - u_i)^2.$$

Writing δ_t for the Dirac delta distribution centered at $t \in \mathbb{R}$, we compute

$$\eta_\delta''(s) = \eta'' * \phi_\delta(s) = \frac{1}{r}(\delta_t - \delta_{t+r}) * \phi_\delta(s) = \frac{1}{r}(\phi_\delta(s-t) - \phi_\delta(s-t-r)).$$

We note that since $\operatorname{supp}\phi'_\delta \subset [-\delta,\delta]$ we have $\operatorname{supp}\eta'_\delta \subset [t-\delta, t+r+\delta]$. Using the symmetry of ω_{ij} and antisymmetry of φ_{ij} we now find

$$\begin{aligned}
\sum_{i,j\in V} \eta_\delta(u_i)\omega^q_{ij}(\varphi_{ji}-\varphi_{ij}) &= \frac{1}{2}\sum_{i,j\in V} \omega^q_{ij}\left(\eta_\delta(u_i)-\eta_\delta(u_j)\right)(\varphi_{ji}-\varphi_{ij}) \\
&= \sum_{i,j\in V} \omega^q_{ij}\left(\eta_\delta(u_i)-\eta_\delta(u_j)\right)\varphi_{ji} \\
&= -\sum_{i,j\in V} \omega^q_{ij}\left(\eta'_\delta(u_i)(u_j-u_i)+\frac{1}{2}\eta''_\delta(\xi_{ij})(u_j-u_i)^2\right)\varphi_{ji} \\
&= \sum_{i\in V\,:\,t-\delta<u_i<t+r+\delta}\ \sum_{j\in V} \eta'_\delta(u_i)\omega^q_{ij}(u_i-u_j)\varphi_{ji} \qquad (52)\\
&\quad + \frac{1}{2r}\sum_{i,j\in V} \omega^q_{ij}(\phi_\delta(\xi_{ij}-t-r)-\phi_\delta(\xi_{ij}-t))(u_i-u_j)^2\varphi_{ji}
\end{aligned}$$

Integrating over $\mathbb{R}$ and taking the limit $\delta\to 0$, for the first term in (52) we compute

$$\begin{aligned}
&\lim_{\delta\to 0}\int_{\mathbb{R}} \sum_{i\in V\,:\,t-\delta<u_i<t+r+\delta}\ \sum_{j\in V} \eta'_\delta(u_i)\omega^q_{ij}(u_i-u_j)\varphi_{ji}\,dt \\
\le &\lim_{\delta\to 0}\int_{\mathbb{R}} \sum_{i\in V\,:\,t-\delta<u_i<t+r+\delta}\ \sum_{j\in V} \frac{1}{r}\omega^q_{ij}(u_i-u_j)\varphi_{ji}\,dt \\
= &\int_{\mathbb{R}} \lim_{\delta\to 0} \sum_{i\in V\,:\,t-\delta<u_i<t+r+\delta}\ \sum_{j\in V} \frac{1}{r}\omega^q_{ij}(u_i-u_j)\varphi_{ji}\,dt \\
= &\int_{\mathbb{R}} \sum_{i\in V\,:\,t<u_i<t+r}\ \sum_{j\in V} \frac{1}{r}\omega^q_{ij}(u_i-u_j)\varphi_{ji}\,dt,
\end{aligned}$$

where we have used, that by definition of $\mathcal{E}_-(u)$, $(u_i-u_j)\varphi_{ji}\ge 0$, and the monotone convergence theorem.

For the second term in (52) we find

$$\begin{aligned}
&\frac{1}{2r}\lim_{\delta\to 0}\int_{\mathbb{R}} \sum_{i,j\in V} \omega^q_{ij}(\phi_\delta(\xi_{ij}-t-r)-\phi_\delta(\xi_{ij}-t))(u_i-u_j)^2\varphi_{ji}\,dt \\
=&\frac{1}{2r}\lim_{\delta\to 0}\sum_{i,j\in V} \omega^q_{ij}\varphi_{ji}(u_i-u_j)^2\int_{\mathbb{R}}(\phi_\delta(\xi_{ij}-t-r)-\phi_\delta(\xi_{ij}-t))\,dt = 0,
\end{aligned}$$

since $\int_{\mathbb{R}}\phi_\delta = 1$.

Combining the above we have

$$\begin{aligned}
\int_{\mathbb{R}}\sum_{i,j\in V}\eta(u_i)(\varphi_{ji}-\varphi_{ij})\omega^q_{ij}\,dt &= \int_{\mathbb{R}}\lim_{\delta\to 0}\sum_{i,j\in V}\eta_\delta(u_i)(\varphi_{ji}-\varphi_{ij})\omega^q_{ij}\,dt \\
&= \lim_{\delta\to 0}\int_{\mathbb{R}}\sum_{i,j\in V}\eta_\delta(u_i)(\varphi_{ji}-\varphi_{ij})\omega^q_{ij}\,dt \\
&\le \int_{\mathbb{R}}\sum_{i\in V\,:\,t<u_i<t+r}\ \sum_{j\in V}\frac{1}{r}\omega^q_{ij}(u_i-u_j)\varphi_{ji}\,dt, \qquad (53)
\end{aligned}$$

where we used the dominated convergence theorem in the second line.

Finally, we conclude, by the definition of η, (49), and (53),

$$\begin{aligned}
\int_{\mathbb{R}} \sum_{i\in E(t)} \sum_{j\in V} (\varphi_{ji}-\varphi_{ij})\omega_{ij}^q \, dt &= \int_{\mathbb{R}} \lim_{r\to 0} \sum_{i,j\in V} \eta(u_i)(\varphi_{ji}-\varphi_{ij})\omega_{ij}^q \, dt \\
&= \lim_{r\to 0} \int_{\mathbb{R}} \sum_{i,j\in V} \eta(u_i)(\varphi_{ji}-\varphi_{ij})\omega_{ij}^q \, dt \\
&\le \lim_{r\to 0} \int_{\mathbb{R}} \sum_{i\in V\,:\, t<u_i<t+r} \sum_{j\in V} \frac{1}{r}\omega_{ij}^q (u_i-u_j)\varphi_{ji} \, dt \\
&= \int_{\mathbb{R}} \lim_{r\to 0} \sum_{i\in V\,:\, t<u_i<t+r} \sum_{j\in V} \frac{1}{r}\omega_{ij}^q (u_i-u_j)\varphi_{ji} \, dt \\
&\le \int_{\mathbb{R}} k'(t)\, dt,
\end{aligned}$$

which proves (51). The second and third identities follow from Lebesgue's dominated convergence theorem, using repeatedly the fact that u is fixed and hence, by the finitude of the graph G, u has compact support and a finite range. The final inequality follows since φ satisfies $\|\varphi\|_{\mathcal{E},\infty} \le 1$. □

C Calculation of the first variation for graph total variation

In (28) we computed $\frac{d}{dt}|_{t=0} \mathrm{TV}_{\mathrm{a}}^q(u+tv) = \langle \mathrm{sgn}(\nabla u), \nabla v\rangle$ using the convexity of $\mathrm{TV}_{\mathrm{a}}^q$. In this section we review this fact to other kinds of graph total variation, which are expressible as

$$\mathcal{TV}(u) := \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} : \varphi \in \mathcal{A}\}, \tag{54}$$

where $\mathcal{A} \subset \mathcal{E}$ is some admissible set of edge functions, such that a maximizer $\varphi^u \in \mathcal{A}$ exists (even if it might not be unique). The key fact is that such a $\mathcal{TV}(u)$ is convex and might be studied via convex analysis. The convexity of $\mathcal{TV}$ is evident from its definition: $u \to \mathcal{TV}(u)$ is a scalar valued function given as the maximum of a family of linear functions $u \mapsto \langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}}$. Let us recall some concepts from convex analysis [ET76, Chapter 1, Section 5], in particular, the subdifferential of $\mathcal{TV}$ at u. This set valued function is denoted by $\partial\mathcal{TV}(u)$ and given by

$$\partial\mathcal{TV}(u) := \{v \in \mathcal{V} \mid \mathcal{TV}(u) < \infty \text{ and } \forall\, w \in \mathcal{V}\ \ \mathcal{TV}(w) \ge \mathcal{TV}(u) + \langle v, w-u\rangle_{\mathcal{V}}\}.$$

That is, $v \in \mathcal{TV}(u)$ if and only if it is the slope of an affine function which is tangent to the graph of $\mathcal{TV}$ at u. In particular, at the points where $\mathcal{TV}(u)$ is differentiable $\partial\mathcal{TV}(u)$ consists of a single element: the gradient of $\mathcal{TV}(u)$ at u.

In the particular case of $\mathcal{TV}$, it follows that

$$\partial\mathcal{TV}(u) = \{v \in \mathcal{V} \mid \langle v, u\rangle_{\mathcal{V}} = \mathcal{TV}(u)\}. \tag{55}$$

Indeed, note that by (54), for any $w \in \mathcal{V}$, $\mathcal{TV}(w) = \langle \operatorname{div}\varphi^w, w\rangle_{\mathcal{V}}$, where $\varphi^w \in \mathcal{A}$ is a maximizer in (54). It follows that, if $u \in \mathcal{V}$ is given, then

$$v \in \partial\mathcal{TV}(u) \Leftrightarrow \langle \operatorname{div}\varphi^w, w\rangle_{\mathcal{V}} \ge \langle \operatorname{div}\varphi^u, u\rangle_{\mathcal{V}} + \langle \operatorname{div} v, w-u\rangle_{\mathcal{V}}.$$

By choosing $w = 0$ and $w = 2u$, respectively, we find $\langle v, u\rangle_{\mathcal{V}} = \langle \operatorname{div}\varphi^u, u\rangle_{\mathcal{V}} = \mathcal{TV}(u)$. This proves the set identity (55) for any $u \in \mathcal{V}$.

On the other hand, the convexity of $\mathcal{TV}(u)$ implies it is a locally Lipschitz function, making it differentiable for a.e.[17] $u \in \mathcal{V}$ by Rademacher's theorem [EG92, Chapter 6, Section 6.2, Theorem 2]. Therefore, for a.e. $u \in \mathcal{V}$, $\partial\mathcal{TV}(u)$ contains a single element $v = \operatorname{div}\varphi^u$. Then, for a.e. u, it follows that

$$\frac{d}{dt}_{|t=0} \mathcal{TV}(u + tv) = \langle \operatorname{div}\varphi^u, v\rangle_{\mathcal{V}}$$

Now we can consider particular choices for $\mathcal{TV}$ and hence for $\mathcal{A}$. For $\mathcal{TV} = \mathrm{TV}^q_{\mathrm{a}}$, we have $\varphi^u = \varphi^a$ from (1), as already explained in Remark 3.11. Similar computations can be done if we take $\mathcal{A}$'s corresponding to φ^u given respectively by

$$\varphi^u(u) = \varphi^{\mathcal{E}}(u) = \begin{cases} \frac{\nabla u}{\|\nabla u\|_{\mathcal{E}}} & \text{if } \|\nabla u\|_{\mathcal{E}} \neq 0, \\ 0 & \text{if } \|\nabla u\|_{\mathcal{E}} = 0, \end{cases} \quad \text{and} \quad \varphi^u_{ij}(u) = \varphi^{\mathrm{TV}}_{ij}(u) = \begin{cases} \frac{(\nabla u)_{ij}}{|\nabla u|_i} & \text{if } |\nabla u|_i \neq 0, \\ 0 & \text{if } |\nabla u|_i = 0, \end{cases}$$

i.e., optimal φ's for $\mathcal{TV}(u) = \|\nabla u\|_{\mathcal{E}}$[18] and $\mathcal{TV}(u) = \mathrm{TV}(u)$, respectively (see Section 2).The previous analysis shows the first variations in these cases are given by the $\mathcal{V}$-inner product with

$$\operatorname{div}\varphi^{\mathcal{E}} = \begin{cases} \frac{\Delta u}{\|\nabla u\|_{\mathcal{E}}} & \text{if } \|\nabla u\|_{\mathcal{E}} \neq 0, \\ 0 & \text{if } \|\nabla u\|_{\mathcal{E}} = 0, \end{cases}$$

$$(\operatorname{div}\varphi^{\mathrm{TV}})_i = \begin{cases} \frac{1}{2} d_i^{-r} \Big[\sum_{j\in V\colon |\nabla u|_j \neq 0} \omega^q_{ij} \frac{(\nabla u)_{ji}}{|\nabla u|_j} \\ \qquad\qquad - \sum_{j\in V} \omega^q_{ij} \frac{(\nabla u)_{ij}}{|\nabla u|_i} \Big] & \text{if } |\nabla u|_i \neq 0, \\ \frac{1}{2} d_i^{-r} \sum_{j\in V\colon |\nabla u|_j \neq 0} \omega^q_{ij} \frac{(\nabla u)_{ji}}{|\nabla u|_j} & \text{if } |\nabla u|_i = 0. \end{cases}$$

For the latter we can also write

$$(\operatorname{div}\varphi^{\mathrm{TV}})_i = \frac{1}{2} d_i^{-r} \sum_{j\in V} \omega^q_{ij} \left(\frac{u_i - u_j}{|\nabla u|_j} + \frac{u_i - u_j}{|\nabla u|_i} \right) = \frac{1}{2} d_i^{-r} \sum_{j\in V} \omega^q_{ij} \left(\frac{1}{|\nabla u|_j} + \frac{1}{|\nabla u|_i} \right) (u_i - u_j),$$

where we have to remember that $\omega^q_{ij} \frac{u_i - u_j}{|\nabla u|_j}$ and $\frac{1}{|\nabla u|_j}$ are to be interpreted as 0 whenever $|\nabla u|_j = 0$ for any $j \in V$ (including $j = i$). Because the node function $\operatorname{div}\varphi^{\mathrm{TV}}$ is the first variation of the isotropic graph total variation, in the literature it is sometimes referred to as curvature or 1-Laplacian. In this paper we have argued why the use of the anisotropic total variation TV_{a} to define curvature, as in (14), is a more natural choice on graphs.

References

[AC79] Samuel M. Allen and John W. Cahn, *A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metallurgica **27** (1979), no. 6, 1085–1095.

[Ada75] R. A. Adams, *Sobolev spaces*, first ed., Pure and applied mathematics; a series of monographs and textbooks, vol. 65, Academic Press, Inc, New York, 1975.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, first ed., Oxford Mathematical Monographs, Oxford University Press, Oxford, 2000.

[ATW93] Fred Almgren, Jean E. Taylor, and Lihe Wang, *Curvature-driven flows: a variational approach*, SIAM J. Control Optim. **31** (1993), no. 2, 387–438. MR 1205983 (94h:58067)

[17] Here a.e. is with respect to Lebesgue measure in $\mathcal{V}$ (which is uniquely determined by its inner product).

[18] Note that we can write $\|\nabla u\|_{\mathcal{E}} = \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} \colon \varphi \in \mathcal{E},\ \|\varphi\|_{\mathcal{E}} \leq 1\}$.

[BBPSV04] Alain Barrat, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani, *The architecture of complex weighted networks*, Proceedings of the National Academy of Sciences of the United States of America **101** (2004), no. 11, 3747–3752.

[BC99] Peter W. Bates and Adam Chmaj, *A discrete convolution model for phase transitions*, Arch. Ration. Mech. Anal. **150** (1999), no. 4, 281–305. MR 1741258 (2001c:82026)

[BF12] A. L. Bertozzi and A. Flenner, *Diffuse interface models on graphs for analysis of high dimensional data*, Multiscale Modeling and Simulation **10** (2012), no. 3, 1090–1118.

[BG95] Guy Barles and Christine Georgelin, *A simple proof of convergence for an approximation scheme for computing motions by mean curvature*, SIAM J. Numer. Anal. **32** (1995), no. 2, 484–500.

[BH09] T. Bühler and M. Hein, *Spectral clustering based on the graph p-laplacian*, Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 81–88.

[BK91] Lia Bronsard and Robert V. Kohn, *Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics*, J. Differential Equations **90** (1991), no. 2, 211–237. MR 1101239 (92d:35037)

[BLS91] A. Björner, L. Lovász, and P. W. Shor, *Chip-firing games on graphs*, European J. Combin **12** (1991), no. 4.

[BLS07] T. Biyikoglu, J. Leydold, and P. F. Stadler, *Laplacian eigenvectors of graphs*, Springer, 2007.

[Bra78] Kenneth A. Brakke, *The motion of a surface by its mean curvature*, Mathematical Notes, vol. 20, Princeton University Press, Princeton, N.J., 1978. MR 485012 (82c:49035)

[Bra02] A. Braides, Γ-*convergence for beginners*, first ed., Oxford Lecture Series in Mathematics and its Applications, vol. 22, Oxford University Press, Oxford, 2002.

[BSS93] G. Barles, H. M. Soner, and P. E. Souganidis, *Front propagation and phase field theory*, SIAM J. Control Optim. **31** (1993), no. 2, 439–469. MR 1205984 (94c:35005)

[CD09] Antonin Chambolle and Jérôme Darbon, *On total variation minimization and surface evolution using parametric maximum flows*, International Journal of Computer Vision **84** (2009), no. 3, 288–307.

[CEN06] Tony F. Chan, Selim Esedoḡlu, and Mila Nikolova, *Algorithms for finding global minimizers of image segmentation and denoising models*, SIAM J. Appl. Math. **66** (2006), no. 5, 1632–1648.

[CG10] M. Cristina Caputo and Nestor Guillen, *Regularity for non-local almost minimal boundaries and applications*, arXiv preprint arXiv:1003.2470 (2010).

[Chu97] F. R. K. Chung, *Spectral graph theory*, AMS, 1997.

[CMOP11] Ozan Candogan, Ishai Menache, Asuman Ozdaglar, and Pablo A. Parrilo, *Flows and decompositions of games: harmonic and potential games*, Math. Oper. Res. **36** (2011), no. 3, 474–503.

[CRS10] L. Caffarelli, J.-M. Roquejoffre, and O. Savin, *Nonlocal minimal surfaces*, Comm. Pure Appl. Math. **63** (2010), no. 9, 1111–1144. MR 2675483 (2011h:49057)

[CS94] G. Caginalp and Eduardo Socolovsky, *Phase field computations of single-needle crystals, crystal growth, and motion by mean curvature*, SIAM Journal on Scientific Computing **15** (1994), no. 1, 106–126.

[CS10] Luis A. Caffarelli and Panagiotis E. Souganidis, *Convergence of nonlocal threshold dynamics approximations to front propagation*, Arch. Ration. Mech. Anal. **195** (2010), no. 1, 1–23. MR 2564467 (2011c:80008)

[CvGO11] Rustum Choksi, Yves van Gennip, and Adam Oberman, *Anisotropic total variation regularized L^1 approximation and denoising/deblurring of 2D bar codes*, Inverse Probl. Imaging **5** (2011), no. 3, 591–617.

[DEL11] Xavier Desquesnes, Abderrahim Elmoataz, and Olivier Lezoray, *PDEs level sets on weighted graphs*, Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE, 2011, pp. 3377–3380.

[DEL12a] Xavier Desquesnes, Abderrahim Elmoataz, and Olivier Lézoray, *Eikonal equation adaptation on weighted graphs: Fast geometric diffusion process for local and non-local image and data processing*, Journal of Mathematical Imaging and Vision (2012), 1–20.

[DEL12b] ———, *Generalized fronts propagation on weighted graphs*, Proceedings of ALGORITMY, 2012, pp. 371–381.

[DEL12c] ———, *PdEs-based morphology on graphs for cytological slides segmentation and clustering*, Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE, 2012, pp. 1619–1622.

[DELT10] Xavier Desquesnes, Abderrahim Elmoataz, Olivier Lézoray, and Vinh-Thong Ta, *Efficient algorithms for image and high dimensional data processing using eikonal equation on graphs*, Advances in Visual Computing, Springer, 2010, pp. 647–658.

[DM93] G. Dal Maso, *An introduction to Γ-convergence*, first ed., Progress in Nonlinear Differential Equations and Their Applications, vol. 8, Birkhäuser, Boston, 1993.

[Dur07] Richard Durrett, *Random graph dynamics*, vol. 20, Cambridge university press, 2007.

[Eck04] Klaus Ecker, *Regularity theory for mean curvature flow*, vol. 57, Springer, 2004.

[EDL12] Abderrahim Elmoataz, Xavier Desquesnes, and Olivier Lézoray, *Non-local morphological PDEs and p-Laplacian equation on graphs with applications in image processing and machine learning*, IEEE Journal of Selected Topics in Signal Processing **6** (2012), no. 7, 764–779.

[EG92] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, first ed., Studies in Advanced Mathematics, CRC Press LLC, Boca Raton, Florida, 1992.

[ELB08] Abderrahim Elmoataz, Olivier Lezoray, and Sébastien Bougleux, *Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing*, Image Processing, IEEE Transactions on **17** (2008), no. 7, 1047–1060.

[EO13] S. Esedoglu and F. Otto, *Threshold dynamics for networks with arbitrary surface tensions.*

[ERT08] S. Esedoglu, S. J. Ruuth, and R. Tsai, *Threshold dynamics for high order geometric motions*, Interfaces and Free Boundaries **10** (2008), no. 3, 263–282.

[ERT10] Selim Esedoḡlu, Steven Ruuth, and Richard Tsai, *Diffusion generated motion using signed distance functions*, Journal of Computational Physics **229** (2010), no. 4, 1017–1042.

[ET76] Ivar Ekeland and Roger Temam, *Convex analysis and variational problems*, North-Holland Publishing Co., Amsterdam, 1976, Translated from the French, Studies in Mathematics and its Applications, Vol. 1. MR MR0463994 (57 #3931b)

[ET06] S. Esedoglu and Y.-H. R. Tsai, *Threshold dynamics for the piecewise constant Mumford-Shah functional*, Journal of Computational Physics **211** (2006), no. 1, 367–384.

[Eva93] Lawrence C. Evans, *Convergence of an algorithm for mean curvature motion*, Indiana Univ. Math. J. **42** (1993), no. 2, 533–557.

[Eva10] ———, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943 (2011c:35002)

[Fie73] M. Fiedler, *Algebraic connectivity of graphs*, Czech. Math. J. **23** (1973), 298–305.

[FS97] T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model*, Comm. Math. Phys. **185** (1997), no. 1, 1–36. MR 1463032 (98f:60206)

[FT04] J. Friedman and J. P. Tillich, *Calculus on graphs*, Arxiv preprint cs/0408028 (2004).

[GCMB+13] Cristina Garcia-Cardona, Ekaterina Merkurjev, Andrea L Bertozzi, Arjuna Flenner, and Allon Percus, *Fast multiclass segmentation using diffuse interface methods on graphs*, submitted (2013).

[GM98] Stephen Guattery and Gary L. Miller, *On the quality of spectral separators*, SIAM J. Matrix Anal. Appl. **19** (1998), no. 3, 701–719. MR 1611179 (99e:68125)

[GO09] G. Gilboa and S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Modeling & Simulation **7** (2009), no. 3, 1005–1028.

[GP10] Leo J. Grady and Jonathan R. Polimeni, *Discrete calculus*, Springer-Verlag London Ltd., 2010.

[HJ90] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

[HLPB13] Huiyi Hu, Thomas Laurent, Mason A. Porter, and Andrea L. Bertozzi, *A method based on total variation for network modularity optimization using the MBO scheme*, submitted (2013).

[HPS11] H. J. Hupkes, D. Pelinovsky, and B. Sandstede, *Propagation failure in the discrete Nagumo equation*, Proc. Amer. Math. Soc. **139** (2011), no. 10, 3537–3551. MR 2813385

[HS98] H. S. Hele-Shaw, *Flow of water*, Nature **58** (1898), 520.

[Imb09] Cyril Imbert, *Level set approach for fractional mean curvature flows*, Interfaces and Free Boundaries **11** (2009), no. 1, 153–176.

[KS94] Markos A. Katsoulakis and Panagiotis E. Souganidis, *Interacting particle systems and generalized evolution of fronts*, Archive for rational mechanics and analysis **127** (1994), no. 2, 133–157.

[KS97] Markos A Katsoulakis and Panagiotis E Souganidis, *Stochastic ising models and anisotropic front propagation*, Journal of statistical physics **87** (1997), no. 1-2, 63–89.

[KVV00] Ravi Kannan, Santosh Vempala, and Adrian Vetta, *On clusterings—good, bad and spectral*, 41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000), IEEE Comput. Soc. Press, Los Alamitos, CA, 2000, pp. 367–377. MR 1931834

[KVV04] ———, *On clusterings: good, bad and spectral*, J. ACM **51** (2004), no. 3, 497–515. MR 2145863 (2006m:05153)

[LS95] Stephan Luckhaus and Thomas Sturzenhecker, *Implicit time discretization for the mean curvature flow equation*, Calc. Var. Partial Differential Equations **3** (1995), no. 2, 253–271. MR 1386964 (97e:65085)

[Lyo89] Russell Lyons, *The ising model and percolation on trees and tree-like graphs*, Communications in Mathematical Physics **125** (1989), no. 2, 337–353.

[Lyo00] ———, *Phase transitions on nonamenable graphs.(english summary)*, J. Math. Phys **41** (2000), no. 3, 1099–1126.

[Mas92] P. Mascarenhas, *Diffusion generated motion by mean curvature*, UCLA Department of Mathematics CAM report CAM 92–33 (1992).

[MBO92] B. Merriman, J. K. Bence, and S. Osher, *Diffusion generated motion by mean curvature*, UCLA Department of Mathematics CAM report CAM 06–32 (1992).

[MBO93] B. Merriman, J.K. Bence, and S. Osher, *Diffusion generated motion by mean curvature*, AMS Selected Letters, Crystal Grower's Workshop (1993), 73–83.

[MBO94] Barry Merriman, James K. Bence, and Stanley J. Osher, *Motion of multiple functions: a level set approach*, J. Comput. Phys. **112** (1994), no. 2, 334–363.

[MKB12] Ekaterina Merkurjev, Tijana Kostić, and Andrea L. Bertozzi, *An MBO scheme on graphs for segmentation and image processing*, submitted (2012).

[MM77] L. Modica and S. Mortola, *Un esempio di* Γ*-convergenza*, Bollettino U.M.I. **5** (1977), no. 14-B, 285–299.

[Mod87] Luciano Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Arch. Rational Mech. Anal. **98** (1987), no. 2, 123–142. MR 866718 (88f:76038)

[Moh91] B. Mohar, *The Laplacian spectrum of graphs*, Graph Theory, Combinatorics, and Applications, vol. 2, Wiley, 1991, pp. 871–898.

[MOS12] J. J. Manfredi, A. M. Oberman, and A. P. Sviridov, *Nonlinear elliptic partial differential equations and p-harmonic functions on graphs*, Preprint (2012).

[MS64] William W. Mullins and R. F. Sekerka, *Stability of a planar interface during solidification of a dilute binary alloy*, Journal of Applied Physics **35** (1964), no. 2, 444–451.

[NJW02] A. Ng, M. Jordan, and Y. Weiss, *On spectral clustering: Analysis and an algorithm*, Dietterich, T., Becker, S., Ghahramani, Z. (eds.) Advances in Neural Information Processing Systems 14, MIT Press, Cambridge, 2002, pp. 849–856.

[OBO13] B. Osting, C. Brune, and S. Osher, *Enhanced statistical rankings via targeted data collection*, JMLR, W&CP **28** (2013), no. 1, 489–497.

[OF03] S. J. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*, vol. 153, Springer Verlag, 2003.

[OSFM07] R. Olfati-Saber, A. Fax, and R. M. Murray, *Consensus and cooperation in networked multi-agent systems*, Proceedings of the IEEE **95** (2007), no. 1, 215–233.

[Peg89] Robert L. Pego, *Front migration in the nonlinear cahn-hilliard equation*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **422** (1989), no. 1863, 261–278.

[Pet00] Norbert Peters, *Turbulent combustion*, Cambridge university press, 2000.

[RS79] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, I: Functional Analysis*, Academic Press, 1979.

[Ruu96] Steven J. Ruuth, *Efficient algorithms for diffusion-generated motion by mean curvature*, ProQuest LLC, Ann Arbor, MI, 1996, Thesis (Ph.D.)–The University of British Columbia (Canada).

[Ruu98a] S. J. Ruuth, *A diffusion-generated approach to multiphase motion*, Journal of Computational Physics **145** (1998), no. 1, 166–192.

[Ruu98b] ———, *Efficient algorithms for diffusion-generated motion by mean curvature*, Journal of Computational Physics **144** (1998), no. 2, 603–625.

[SB10] Arthur Szlam and Xavier Bresson, *Total variation and cheeger cuts*, Proceedings of the 27th International Conference on Machine Learning (ICML-10) (Haifa, Israel) (Johannes Fürnkranz and Thorsten Joachims, eds.), Omnipress, June 2010, pp. 1039–1046.

[SBXD04] J. Sun, S. Boyd, L. Xiao, and P. Diaconis, *The fastest mixing Markov process on a graph and a connection to a maximum variance unfolding problem*, SIAM Review **48** (2004), 2006.

[SM00] Jianbo Shi and Jitendra Malik, *Normalized cuts and image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000), no. 8, 888–905.

[ST96] Daniel A. Spielman and Shang-Hua Teng, *Spectral partitioning works: planar graphs and finite element meshes*, 37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996), IEEE Comput. Soc. Press, Los Alamitos, CA, 1996, pp. 96–105. MR 1450607 (98d:65144)

[ST03] H. Mete Soner and Nizar Touzi, *A stochastic representation for mean curvature type geometric flows*, The Annals of probability **31** (2003), no. 3, 1145–1165.

[ST07] Daniel A. Spielman and Shang-Hua Teng, *Spectral partitioning works: planar graphs and finite element meshes*, Linear Algebra Appl. **421** (2007), no. 2-3, 284–305. MR 2294342 (2008h:15011)

[Sza65] Jacek Szarski, *Differential inequalities*, PWN Warsaw, 1965.

[Tay92] Jean E. Taylor, *II—mean curvature and weighted mean curvature*, Acta Metallurgica et Materialia **40** (1992), no. 7, 1475–1485.

[TEL08] Vinh-Thong Ta, Abderrahim Elmoataz, and Olivier Lézoray, *Partial difference equations over graphs: Morphological processing of arbitrary discrete data*, Computer Vision–ECCV 2008, Springer, 2008, pp. 668–680.

[TEL11] ———, *Nonlocal PDEs-based morphology on weighted graphs for image and data processing*, Image Processing, IEEE Transactions on **20** (2011), no. 6, 1504–1516.

[Tem97] Roger Temam, *Infinite-dimensional dynamical systems in mechanics and physics*, second ed., Applied Mathematical Sciences, vol. 68, Springer-Verlag, New York, 1997. MR 1441312 (98b:58056)

[vGB12] Y. van Gennip and A. L. Bertozzi, Γ*-convergence of graph Ginzburg-Landau functionals*, Adv. Differential Equations **17** (2012), no. 11–12, 1115–1180.

[vL07] Ulrike von Luxburg, *A tutorial on spectral clustering*, Statistics and Computing **17** (2007), no. 4, 395–416.

[WS98] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks.*, Nature **393** (1998), no. 6684, 440–442.

[WW93] Dorothea Wagner and Frank Wagner, *Between min cut and graph bisection*, Mathematical foundations of computer science 1993 (Gdańsk, 1993), Lecture Notes in Comput. Sci., vol. 711, Springer, Berlin, 1993, pp. 744–750. MR 1265105 (94k:05195)

[XY11] Jack Xin and Yifeng Yu, *Analysis and comparison of large time front speeds in turbulent combustion models*, arXiv preprint arXiv:1105.5607 (2011).

[ZD09] J. Zhang and Q. Du, *Numerical studies of discrete approximations to the Allen-Cahn equation in the sharp interface limit*, SIAM Journal on Scientific Computing **31** (2009), no. 4, 3042–3063.