Using evolving interface techniques to solve network problems

Yves van Gennip

(joint work with many people, acknowledged through references in this abstract)

In recent years there has been increasing interest from applied analysts in applying the models and techniques from variational methods and partial differential equations (PDEs) to tackle problems on networks. This talk gave an overview of some of the recent developments in this young and growing area.

For the purposes of the talk, [1] kicked off the research in this area. In this paper the authors use graph versions of the Ginzburg-Landau functional for data clustering, data classification, and image segmentation. Minimisation of the classical continuum Ginzburg-Landau functional,

$$F(u) := \varepsilon \int_\Omega |\nabla u|^2 \, dx + \frac{1}{\varepsilon} \int_\Omega W(u) \, dx,$$

provides a model for phase separation. Here $W(u) = u^2(1-u)^2$ is a double well potential with minima at $u = 0$ and $u = 1$, and u describes the relative presence of the two phases $\{u \approx 0\}$ and $\{u \approx 1\}$ in the domain Ω. When F is minimised under some suitable constraints on u (e.g. a mass constraint of the form $\int_\Omega u \, dx = M$) and for small values of the parameter ε, u will take values close to 0 and 1, with transitions between those values occurring in small regions of width $\mathcal{O}(\varepsilon)$.

In [1] the graph functional

$$f(u) := \sum_{i,j \in V} \omega_{ij} (u_i - u_j)^2 + \frac{1}{\varepsilon} \sum_{i \in V} W(u_i)$$

was introduced. This is a functional whose input argument u is a function on the nodes of a given graph, instead of on a continuum set $\Omega \subset \mathbb{R}^n$ and which serves as a graph counterpart of F. Here V is the node set of the (finite, simple, undirected) graph, ω_{ij} is a nonnegative weight on the edge between nodes i and j in the graph, and u_i is the value of the function u on node i. In [1] this functional was used in combination with either a mass constraint or an additional data fidelity term to cluster or classify the nodes of a graph into two groups ('phases' where $u \approx 0$ and $u \approx 1$) based on the pairwise node similarity encoded in the edge weights ω_{ij}. By treating the pixels of an image as nodes in a graph, data classifcation can be used for image segmentation as well.

We can now ask a number of questions:

(1) Can we find graph analogues of properties of the continuum functional?
(2) Is the continuum functional a limit of the graph functionals in some sense?
(3) What can we say about the resulting algorithm and its usage for data analysis/image processing?
(4) Are there other network problems that can be tackled by a PDE inspired approach?
(5) Are there other PDE/variational systems that have interesting network analogues?
    - If the inspiring PDEs are related, are their graph analogues related?

This talk gave a short overview addressing (some aspects of some of) these questions.

(1) Does f have similar properties as F? In [2] we proved that f Γ-converges, when $\varepsilon \to 0$, to the graph total variation functional

$$TV(u) := \frac{1}{2} \sum_{i,j \in V} \omega_{ij} |u_i - u_j|,$$

with as domain the set of node functions u which take values in $\{0,1\}$. This mirrors the well-known continuum result [3, 4]. Moreover, for such $\{0,1\}$-valued functions u, $TV(u)$ reduces to the graph cut [5] of the node partition $V_0 = \{i : u_i = 0\}$, $V_1 = \{i : u_i = 1\}$, i.e. the sum of the edge weights ω_{ij} corresponding to edges that have one node in V_0 and the other in V_1.
(2) Furthermore, when f or TV are defined on certain graphs of which a sensible continuum limit can be defined, they Γ-converge to the continuum total variation in the continuum limit, e.g. on 4-regular graphs obtained by ever finer discretisations of the flat torus [2] and on point clouds obtained by sampling ever more points from an underlying subset of $\mathbb{R}^n$ [6, 7, 8].
(3) Minimisation of f is in practice (approximately) achieved either by solving a gradient flow equation of Allen-Cahn type,

$$\frac{du_i}{dt} = -\sum_{j \in V} \omega_{ij}(u_i - u_j) - \frac{1}{\varepsilon} W'(u_i)$$

(plus additional terms coming from a mass constraint or fidelity term) or by a graph version of the threshold dynamics (or MBO) scheme [9]:

$$u^{k+1} = \begin{cases} 0, & \text{if } \tilde{u}(\tau) < \frac{1}{2}, \\ 1, & \text{if } \tilde{u}(\tau) \geq \frac{1}{2}, \end{cases} \quad \text{where } \tilde{u}(t) \text{ solves} \begin{cases} \tilde{u}(0) = 0, \\ \frac{d\tilde{u}_i}{dt} = -\sum_{j\in V} \omega_{ij}(\tilde{u}_i - \tilde{u}_j). \end{cases}$$

In the (spectral) graph theory literature [5, 10] $(\Delta u)_i := \sum_{j\in V} \omega_{ij}(u_i - u_j)$ is known as the unnormalised or combinatorial graph Laplacian of u. The equations above can also be formulated and solved with normalised versions of the graph Laplacian.

On a given graph, these equations can be solved quickly and accurately using a truncated spectral decomposition based on the eigenfunctions of the graph Laplacian (in combination with a convex splitting scheme in the case of the graph Allen-Cahn equation) [1, 11].

The construction of the underlying graph in the first place can pose a significant computational problem, especially when the number of data points (and thus nodes in the graph) is very large. Matrix completion techniques such as the Nyström extension [12, 13] and fast eigenvalue computation algorithms such as the Rayleigh-Chebychev algorithm [14] make such computations feasible.

This graph Ginzburg-Landau method has found many applications, for example in data clustering and classification and image segmentation [1, 11, 15] and has also been extended to deal with clustering and classification into more than two classes [16, 17, 18, 19, 20]. Recent papers prove convergence of the graph Allen-Cahn algorithm (both the spectrally untruncated and truncated versions) and extend the method to non-smooth potentials and hypergraphs.

This shows that such PDE driven techniques can provide fast approximative alternatives to combinatorial problems whose exact solution is too computationally complex.

(4) Another example of such a problem is the computation of a maximum cut in graphs, i.e. to find a partition of the node set into two sets such that the sum of the edge weights corresponding to edges with one node in each set is maximal. If the graph is bipartite, this corresponds to partitioning the node set according to the bipartite structure. The exact solution of this classical problem is known to be computationally unfeasible for large graphs. Work currently in preparation introduces a fast approximate solution method for this problem using an adaptation of the graph Ginzburg-Landau functional f [21].

(5) The continuum counterparts of both the graph Allen-Cahn equation and graph MBO scheme from point (3) can be viewed as approximating mean curvature flow [22, 23, 24, 25, 26]. This suggests that graph curvature and graph mean curvature flow are interesting concepts to consider as well. In

[27] we introduced both. The graph curvature of a node set S is given by

$$\kappa_i := \begin{cases} \sum_{j\in S^c} \omega_{ij}, & \text{if } i \in S, \\ -\sum_{j\in S} \omega_{ij}, & \text{if } i \in S^c, \end{cases}$$

and the related graph mean curvature flow has a variational formulation along the lines of [28, 29, 30] which leads to a time discrete evolution of node subsets S (given an initial set S_0),

$$S_{n+1} \in \operatorname{argmin}_{\hat{S}} \mathcal{F}(\hat{S}, S_n),$$

where

$$\mathcal{F}(\hat{S}, S_n) := \sum_{i\in S, j\in S^c} \omega_{ij} + \frac{1}{\eth t} \sum_{i\in \hat{S}} d_i sd_i^n.$$

Here d_i is the degree of node i and sd_i^n is the signed graph distance from node i to the boundary of node set S_n. In [27] we started studying the very interesting question if the graph Allen-Cahn equation, graph MBO scheme, and graph mean curvature flow are as intimately connected as their continuum counterparts, but establishing such connections is still mostly an open problem.

Other current work studies a graph version of the Ohta-Kawasaki functional [31], which was originally introduced as a variational model for pattern formation in diblock copolymers [32].

The research on these novel methods has shown that new PDE inspired graph procedures can efficiently (approximately) solve complex graph problems, while at the same time offering fertile ground for proving theoretical connections between the various graph problems (inspired by similar connections their continuum counterparts have) and between the graph problems and their continuum analogues.

References

[1] A.L. Bertozzi, and A. Flenner, *Diffuse interface models on graphs for analysis of high dimensional data*, Multiscale Modeling and Simulation **10(3)** (2012), 1090–1118.

[2] Y. van Gennip and A.L. Bertozzi, *Γ-convergence of graph Ginzburg-Landau functionals*, Advances in Differential Equations **17(11/12)** (2012), 1115–1180.

[3] L. Modica and S. Mortola, *Un esempio di Γ-convergenza*, Bollettino dell'Unione Matematica Italiana **5(14-B)** (1977), 285–299.

[4] L. Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Archive for Rational Mechanics and Analysis **98(2)** (1987), 123–142.

[5] U. von Luxburg, *A tutorial on spectral clustering*, Statistics and Computing **17(4)** (2007), 395–416.

[6] N. Garcia Trillos and D. Slepčev, *Continuum limit of total variation on point clouds*, Archive for Rational Mechanics and Analysis *220(1)* (2016), 193–241.

[7] N. Garcia Trillos, D. Slepčev, J. von Brecht, T. Laurent, and X. Bresson, *Consistency of Cheeger and Ratio Graph Cuts*, Journal of Machine Learning Research **17** (2016), 1–46.

[8] M. Thorpe and F. Theil, *Asymptotic Analysis of the Ginzburg-Landau Functional on Point Clouds*, to appear, arXiv preprint arXiv:1604.04930

[9] B. Merriman, J.K. Bence, and S.J. Osher, *Motion of multiple functions: a level set approach*, Journal of Computational Physics **112(2)** (1994), 334–363.

[10] F. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997.

[11] E. Merkurjev, T. Kostic, and A.L. Bertozzi, *An MBO scheme on graphs for segmentation and image processing*, SIAM Journal on Imaging Sciences **6(4)** (2013), 1903–1930.
[12] E.J. Nyström, *Über die Praktische Auflösung von Linearen Integralgleichungen mit Anwendungen auf Randwertaufgaben der Potentialtheorie*, Commentationes Physico-Mathematicae **4(15)** (1928), 1–52.
[13] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, *Spectral grouping using the Nystrom method*, IEEE transactions on pattern analysis and machine intelligence **26(2)** (2004), 214–225.
[14] C.R. Anderson, *A Rayleigh–Chebyshev procedure for finding the smallest eigenvalues and associated eigenvectors of large sparse Hermitian matrices*, Journal of Computational Physics **229(19)** (2010), 7477–7487.
[15] L. Calatroni, Y. van Gennip, C.-B. Schönlieb, H. Rowland, and A. Flenner, *Graph clustering, variational image segmentation methods and Hough transform scale detection for object measurement in images*, Journal of Mathematical Imaging and Vision **57(2)** (2017), 269–291.
[16] C. Garcia-Cardona, A. Flenner, and A.G. Percus, *Multiclass diffuse interface models for semi-supervised learning on graphs*, Proceedings of the 2nd International Conference on Pattern Recognition Applications and Methods (ICPRAM 2013) (2013), 78–86.
[17] E. Merkurjev, C. Garcia-Cardona, A.L. Bertozzi, A. Flenner, and A.G. Percus, *Diffuse interface methods for multiclass segmentation of high-dimensional data*, Applied Mathematics Letters **33** (2014), 29–34.
[18] , C. Garcia-Cardona, E. Merkurjev, A.L. Bertozzi, A. Flenner, and A.G. Percus, *Multiclass data segmentation using diffuse interface methods on graphs*, IEEE Transactions on Pattern Analysis and Machine Intelligence **36** (2014), 1600–1613.
[19] C. Garcia-Cardona, A. Flenner, and A.G. Percus, *Multiclass semi-supervised learning on graphs using Ginzburg-Landau functional minimization*, Advances in Intelligent Systems and Computing **318** (2015), 119–135.
[20] Z. Meng, E. Merkurjev, A. Koniges, and A.L. Bertozzi, *Hyperspectral Image Classification Using Graph Clustering Methods*, in prep.
[21] B. Keetch and Y. van Gennip, *A Max-Cut approximation using a graph based MBO scheme*, in prep.
[22] K.A. Brakke, *The motion of a surface by its mean curvature*, Princeton University Press (1978)
[23] L. Bronsard and R.V. Kohn, *Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics*, Journal of Differential Equations **90(2)** (1991), 211–237.
[24] L.C. Evans, *Convergence of an algorithm for mean curvature motion*, Indiana University Mathematics Journal **42(2)** (1993), 533-557.
[25] G. Barles, H.M. Soner, and P.E. Souganidis, *Front propagation and phase field theory*, SIAM Journal on Control and Optimization **31(2)** (1993), 439–469.
[26] G. Barles and C. Georgelin, *A simple proof of convergence for an approximation scheme for computing motions by mean curvature*, SIAM Journal of Numerical Analysis **32(2)** (1995), 484–500.
[27] Y. van Gennip, N. Guillen, B. Osting, and A.L. Bertozzi, *Mean Curvature, Threshold Dynamics, and Phase Field Theory on Finite Graphs*, Milan Journal of Mathematics **82(1)** (2014), 3–65.
[28] F. Almgren, J.E. Taylor, and L. Wang, *Curvature-driven flows: a variational approach*, SIAM Journal on Control and Optimization **31(2)** (1993), 387–438.
[29] S. Luckhaus and T. Sturzenhecker, *Implicit time discretization for the mean curvature flow equation*, Calculus of Variations and Partial Differential Equations **3(2)** (1995), 253–271.
[30] J.E. Taylor, *Anisotropic interface motion*, Mathematics of Microstructure Evolution **4** (1996), 135–148.

[31] T. Ohta and K. Kawasaki, *Equilibrium Morpholoy of Block Copolymer Melts*, Macromolecules **19** (1986), 2621–2632.
[32] Y. van Gennip, *A graph Ohta-Kawasaki model*, in prep.