DIFFUSE INTERFACE METHODS ON GRAPHS

Andrea Bertozzi
University of California, Los Angeles
www.math.ucla.edu/~bertozzi

Thanks to J. Dobrosotskaya, S. Esedoglu, A. Flenner, Y. van Gennip, A. Gillette, C. Schoenlieb, Martin Short, Alexey Stomakhin
SPARSITY AND L1 COMPRESSIVE SENSING – DIFFUSE INTERFACES

\[ \int |\nabla u| \, dx \sim \frac{\epsilon}{2} \int |\nabla u|^2 + \frac{1}{\epsilon} \int W(u) \, dx \]

Total variation \quad \text{Ginzburg-Landau functional}

\( W \) is a double well potential with two minima

Total variation measures length of boundary between two constant regions. This is the most heavily used metric for sparse data in e.g. image processing because it identifies edges and allows for efficient grouping of information.

GL energy is a diffuse interface approximation of TV for binary functionals
CAHN-HILLIARD INPAINTING


Patent pending. Transitioned to NGA for road inpainting.
Transitioned to InQtel for document exploitation.
Continue edges in the same direction – higher order method for local inpainting.
Fast method using convexity splitting and FFT

\[ u_t = -\nabla^2 \left( \varepsilon \nabla^2 u - \frac{1}{\varepsilon} W'(u) \right) + \lambda(\mathbf{x})(f - u) \]

where

\[ \lambda(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \in D \\ \lambda_0 & \text{if } \mathbf{x} \in \Omega \setminus D. \end{cases} \]

\( \text{H}^{-1} \) gradient flow for diffuse TV
\( \text{L}^2 \) fidelity with known data
  • Transitioned to NGA for road inpainting.
  • Transitioned to InQtel for document exploitation.
• Nonlocal wavelet basis replaces Fourier basis in classical diffuse interface method.
• Analysis theory in Besov spaces.
• Gamma convergence to anisotropic TV.
Theorem:

\[ GL_\varepsilon(f) = \frac{\varepsilon}{2} \int |\nabla f(x)|^2 dx + \frac{1}{4\varepsilon} \int W(f(x))dx, \quad W(f) = (f^2 - 1)^2 \]

\[ WGL_\varepsilon(f) = \frac{\varepsilon}{2} |f|_B^2 + \frac{1}{4\varepsilon} \int W(f(x))dx, f \in H^1, \]

\[ WGL_\varepsilon(u_\varepsilon) \xrightarrow{\Gamma} G_\infty(u), \quad G_\infty(u) = \frac{\sqrt{2}}{3} C(u)|u|_{TV}, \]

\[ G_\infty(\chi_E) = \int_{\partial E} \rho(\bar{n}(x), \psi)dl(x), \]
GAMMA CONVERGE OF WAVELET GINZBURG-LANDAU ENERGY

Dobrosotskaya and Bertozzi, IFB 2010
DIFFUSE INTERFACES ON GRAPHS

Joint work with Arjuna Flenner, China Lake
Paper to appear in MMS 2012
In a typical application we have data supported on the graph, possibly high dimensional. The above weights represent comparison of the data.

Examples include:

voting records of Congress – each person has a vote vector associated with them.

Nonlocal means image processing – each pixel has a pixel neighborhood that can be compared with nearby and far away pixels.
GRAPH BASED GL FUNCTIONAL

\[ L(\nu, \mu) = \begin{cases} 
  d(\nu) & \text{if } \nu = \mu, \\
  -w(\nu, \mu) & \text{otherwise}. 
\end{cases} \]

\[ \langle u, Lu \rangle = \frac{1}{2} \sum_{\mu, \nu \in V} w(\nu, \mu)(u(\nu) - u(\mu))^2 \]

\[ L_s = D^{-1/2} LD^{-1/2} = I - D^{-1/2} WD^{-1/2}. \]

\[ E(u) = \frac{\epsilon}{2} \langle u, L_s u \rangle + \frac{1}{4\epsilon} \sum_{z \in Z} (u^2(z) - 1)^2 + \sum_{z \in Z} \frac{\lambda(z)}{2} (u(z) - u_0(z))^2. \]
This example uses nonlocal means-type weights between pixels
Application of a fast method for computing eigenfunctions of a linear operator – e.g. graph Laplacian.

Useful for fully connected graphs.

Example on right uses k-means clustering of eigenfunctions to segment the image.
Basic idea:

$$E(u) = E_c(u) - E_e(u)$$

$$U_{k+1} - U_k = -\Delta t(\nabla E_c(U_{k+1}) - \nabla E_e(U_k))$$

Art is to choose $E_c$ to give an implicit problem that is easy to solve
- e.g. $E_c$ is H1 semi norm – can be solved using FFT
- in wavelet case $E_c$ is wavelet Laplace operator

Constraints on $E_c$ and $E_e$ so that splitting is unconditionally stable

Proof of convergence of splitting schemes for various higher order inpainting methods.
Convex Splitting for the Graph Laplacian

1. Input $\leftarrow$ an initial function $u_0$ and the eigenvalue-eigenvector pairs $(\tilde{\lambda}_k, \phi_k(x))$ for the graph Laplacian $L_s$ from Equation (2.7).
2. Set convexity parameter $c$ and interface scale $\epsilon$ from Equation (3.2).
3. Set the time step $dt$.
4. Initialize $a_k^{(0)} = \int u(x) \phi_k(x) \, dx$.
5. Initialize $b_k^{(0)} = \int [u_0(x)]^3 \phi_k(x) \, dx$.
6. Initialize $d_k^{(0)} = 0$.
7. Calculate $D_k = 1 + dt \left( \epsilon \tilde{\lambda}_k + c \right)$.
8. For $n$ less than a set number of iterations $M$
   (a) $a_k^{(n+1)} = D_k^{-1} \left[ (1 + \frac{dt}{\epsilon} + c \, dt) \, a_k^{(n)} - \frac{dt}{\epsilon} b_k^{(n)} - dt d_k^{(n)} \right]$  
   (b) $u^{(n+1)}(x) = \sum_k a_k^{(n+1)} \phi_k(x)$
   (c) $b_k^{(n+1)} = \int [u^{(n+1)}(x)]^3 \phi_k(x) \, dx$
   (d) $d_k^{(n+1)} = \int \lambda(x) \left( u^{(n+1)}(x) - u_0(x) \right) \phi_k(x) \, dx$
9. end for
10. Output $\leftarrow$ the function $u^{(M)}(x)$.
TWO MOONS EXAMPLE

Data embedded in $\mathbb{R}^{100}$

Replaces Laplace operator with a weighted graph Laplacian in the Ginzburg Landau Functional

Allows for segmentation using L1-like metrics due to connection with $G$.

Compared with L1 methods of Hein and Beuhler NIPS 2010.
US HOUSE OF REPRESENTATIVES VOTING RECORD CLASSIFICATION OF PARTY AFFILIATION FROM VOTING RECORD

98th US Congress 1984
Assume knowledge of party affiliation of 5 of the 435 members of the House
Infer party affiliation of the remaining 430 members from voting records
Gaussian similarity weight matrix for vector of votes (1, 0, -1)
MACHINE LEARNING IDENTIFICATION OF SIMILAR REGIONS IN IMAGES

High dimensional fully connected graph – use Nystrom extension methods for fast computation methods.
Alexey Stomakhin, Martin B. Short, and Andrea L. Bertozzi, Reconstruction of Missing Data in Social Networks Based on Temporal Patterns of Interactions, accepted in Inverse Problems, 2011.
HAWKES PROCESS

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]
HAWKES PROCESS

rivalry intensity

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]
rivalry intensity

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]

background rate of violence
HAWKES PROCESS

rivalry intensity

self-excitation

\[ \lambda(t) = \mu + k_0 \sum_{t>t_i} \omega e^{-\omega(t-t_i)} \]

background rate of violence
HAWKES PROCESS

rivalry intensity

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]

time since the most recent incident

background rate of violence
HAWKES PROCESS

rivalry intensity

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]

background rate of violence

retaliation strength
HAWKES PROCESS

rivalry intensity

\[ \lambda(t) = \mu + k_0 \sum_{t > t_i} \omega e^{-\omega(t-t_i)} \]

background rate of violence

retaliation strength

retaliation duration
<table>
<thead>
<tr>
<th>Rivalry</th>
<th>$k_0$</th>
<th>$w^{-1}$ (days)</th>
<th>$\mu$</th>
<th>expect. cluster duration</th>
<th>expect. cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locke St-Lowell St</td>
<td>0.584</td>
<td>14.7</td>
<td>0.011</td>
<td>15.106</td>
<td>1.369</td>
</tr>
<tr>
<td>Clover St-Eastlake</td>
<td>0.209</td>
<td>1.1</td>
<td>0.025</td>
<td>0.278</td>
<td>0.261</td>
</tr>
<tr>
<td>Lincoln St-Eastlake</td>
<td>0.121</td>
<td>10.4</td>
<td>0.013</td>
<td>1.295</td>
<td>0.133</td>
</tr>
<tr>
<td>KAM-State St</td>
<td>0.259</td>
<td>9.9</td>
<td>0.014</td>
<td>3.226</td>
<td>0.356</td>
</tr>
<tr>
<td>Tiny Boys-State St</td>
<td>0.124</td>
<td>0.03</td>
<td>0.013</td>
<td>0.004</td>
<td>0.136</td>
</tr>
<tr>
<td>MCF-ELA</td>
<td>0.513</td>
<td>12.5</td>
<td>0.008</td>
<td>10.731</td>
<td>1.055</td>
</tr>
<tr>
<td>VNE-Opal St</td>
<td>0.390</td>
<td>7.5</td>
<td>0.012</td>
<td>3.958</td>
<td>0.610</td>
</tr>
<tr>
<td>VNE-8th St</td>
<td>0.488</td>
<td>30.3</td>
<td>0.008</td>
<td>23.647</td>
<td>0.969</td>
</tr>
<tr>
<td>TMC-Cuatro Flats</td>
<td>0.074</td>
<td>47.6</td>
<td>0.023</td>
<td>3.697</td>
<td>0.079</td>
</tr>
</tbody>
</table>
OVERVIEW OF HOLLENBECK GANGS

<table>
<thead>
<tr>
<th>Rivalry</th>
<th>Poisson $\mu$</th>
<th>$k_0$</th>
<th>$w$</th>
<th>$\mu$</th>
<th>Poisson AIC</th>
<th>Hawkes AIC</th>
<th>Best Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loc-Low</td>
<td>0.026</td>
<td>0.584</td>
<td>0.068</td>
<td>0.011</td>
<td>253.3</td>
<td>242.2</td>
<td>Hawkes</td>
</tr>
<tr>
<td>Clo-Eas</td>
<td>0.032</td>
<td>0.209</td>
<td>0.921</td>
<td>0.025</td>
<td>296.0</td>
<td>288.8</td>
<td>Hawkes</td>
</tr>
<tr>
<td>Lin-Eas</td>
<td>0.014</td>
<td>0.121</td>
<td>0.096</td>
<td>0.013</td>
<td>159.3</td>
<td>162.7</td>
<td>Poisson</td>
</tr>
<tr>
<td>KAM-Sta</td>
<td>0.018</td>
<td>0.259</td>
<td>0.101</td>
<td>0.014</td>
<td>192.2</td>
<td>193.1</td>
<td>Poisson</td>
</tr>
<tr>
<td>Tin-Sta</td>
<td>0.015</td>
<td>0.124</td>
<td>28.6</td>
<td>0.013</td>
<td>167.7</td>
<td>157.6</td>
<td>Hawkes</td>
</tr>
<tr>
<td>MCF-ELA</td>
<td>0.017</td>
<td>0.513</td>
<td>0.080</td>
<td>0.008</td>
<td>184.2</td>
<td>173.6</td>
<td>Hawkes</td>
</tr>
<tr>
<td>VNE-Opa</td>
<td>0.019</td>
<td>0.390</td>
<td>0.134</td>
<td>0.012</td>
<td>200.2</td>
<td>192.8</td>
<td>Hawkes</td>
</tr>
<tr>
<td>VNE-8th</td>
<td>0.015</td>
<td>0.488</td>
<td>0.033</td>
<td>0.008</td>
<td>167.7</td>
<td>167.0</td>
<td>Hawkes</td>
</tr>
<tr>
<td>TMC-Cua</td>
<td>0.025</td>
<td>0.074</td>
<td>0.021</td>
<td>0.023</td>
<td>246.0</td>
<td>247.6</td>
<td>Poisson</td>
</tr>
</tbody>
</table>

Here $k_0$ is the expected number of retaliations per attack, 
1/$w$ is the expected waiting time for retaliation (in days)
event dependence is a common process driving repeat victimization across all crime types

specific behavioral mechanism—street smarts/street justice—may differ in detail, but outcome is the same

Hawkes Process is a flexible representation of self-excitation
RECONSTRUCTION OF MISSING DATA IN SOCIAL NETWORKS BASED ON TEMPORAL PATTERNS OF INTERACTIONS

Stomakhin, Short, ALB, Inverse Problems 2011

\[ \lambda_{\alpha\beta}(t) = \mu_{\alpha\beta} + \theta_{\alpha\beta} \sum_{t_{i_{\alpha\beta}} < t} \omega_{\alpha\beta} e^{-\omega_{\alpha\beta} (t - t_{i_{\alpha\beta}})} \]
THE VARIATIONAL PROBLEM

- Likelihood function

\[ L = \prod_{\alpha \beta} \prod_{t_i^\alpha} \lambda_{\alpha \beta}(t_i^{\alpha \beta}). \]

- Energy functional

\[ \Lambda = \sum_{\alpha \beta} \sum_{t_i^\alpha} \lambda_{\alpha \beta}(t_i^{\alpha \beta}). \]

\[
\begin{aligned}
\max \left\{ \sum_{\alpha \beta} \sum_{i,j} \left[ \delta_{ij} \mu_{\alpha \beta} m_i^{\alpha \beta} + \\
+ \frac{1}{2}(1 - \delta_{ij}) \theta_{\alpha \beta} \omega_{\alpha \beta} e^{-\omega_{\alpha \beta} |t_i^{\alpha \beta} - t_j^{\alpha \beta}| m_i^{\alpha \beta} m_j^{\alpha \beta} \right] \right\}, \\
\sum_{\alpha \beta} (m_i^{\alpha \beta})^2 = 1, \quad \forall i = 1, \ldots, n \\
\sum_{\alpha \beta} m_i^{\alpha \beta} \geq 0, \quad \forall i = 1, \ldots, n, \quad \forall \alpha \beta
\end{aligned}
\]
For the variational problem with L2 regularization:

- There exists a global maximizer
- Every local maximizer is a global maximizer
- Under certain conditions of the problem the global maximizer is unique.
Table 2. Continuous model (8) performance results. The first three columns describe the dimensions of the network and the data the method was applied to, and the last three indicate how often, on average, a ground-truth unknown pair was in the top one, top two, and top three weights of the predicted distribution. The $\star$ value of $k$ corresponds to the real Los Angeles gang network, see Figure 3, which is not a fully connected graph. The “Guessing” rows show the results that would be obtained by random guessing.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$N$</th>
<th>$n$</th>
<th>Top 1</th>
<th>Top 2</th>
<th>Top 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>400</td>
<td>50</td>
<td>57%</td>
<td>80%</td>
<td>92%</td>
</tr>
<tr>
<td>5</td>
<td>400</td>
<td>100</td>
<td>56%</td>
<td>79%</td>
<td>91%</td>
</tr>
<tr>
<td>5</td>
<td>400</td>
<td>200</td>
<td>54%</td>
<td>76%</td>
<td>90%</td>
</tr>
<tr>
<td>5</td>
<td>Guessing</td>
<td></td>
<td>25%</td>
<td>50%</td>
<td>75%</td>
</tr>
<tr>
<td>7</td>
<td>400</td>
<td>50</td>
<td>47%</td>
<td>69%</td>
<td>82%</td>
</tr>
<tr>
<td>7</td>
<td>400</td>
<td>100</td>
<td>46%</td>
<td>68%</td>
<td>80%</td>
</tr>
<tr>
<td>7</td>
<td>400</td>
<td>200</td>
<td>45%</td>
<td>65%</td>
<td>77%</td>
</tr>
<tr>
<td>7</td>
<td>Guessing</td>
<td></td>
<td>17%</td>
<td>33%</td>
<td>50%</td>
</tr>
<tr>
<td>9</td>
<td>400</td>
<td>50</td>
<td>42%</td>
<td>62%</td>
<td>73%</td>
</tr>
<tr>
<td>9</td>
<td>400</td>
<td>100</td>
<td>41%</td>
<td>60%</td>
<td>72%</td>
</tr>
<tr>
<td>9</td>
<td>400</td>
<td>200</td>
<td>39%</td>
<td>57%</td>
<td>69%</td>
</tr>
<tr>
<td>9</td>
<td>Guessing</td>
<td></td>
<td>13%</td>
<td>25%</td>
<td>38%</td>
</tr>
<tr>
<td>$\star$</td>
<td>400</td>
<td>50</td>
<td>50%</td>
<td>72%</td>
<td>83%</td>
</tr>
<tr>
<td>$\star$</td>
<td>400</td>
<td>100</td>
<td>49%</td>
<td>71%</td>
<td>82%</td>
</tr>
<tr>
<td>$\star$</td>
<td>400</td>
<td>200</td>
<td>48%</td>
<td>68%</td>
<td>80%</td>
</tr>
</tbody>
</table>

Parameters for Los Angeles Hollenbeck including graph network
Andrea L. Bertozzi and Arjuna Flenner,
Diffuse interface models on graphs for classification of high dimensional data
To appear in MMS 2012.

Carola-Bibiane Schoenlieb and Andrea Bertozzi,
Unconditionally stable schemes for higher order inpainting,

Julia A. Dobrosotskaya and Andrea L. Bertozzi,
Wavelet analogue of the Ginzburg-Landau energy and its Gamma-convergence,

Julia A. Dobrosotskaya and Andrea L. Bertozzi,
A Wavelet-Laplace Variational Technique for Image Deconvolution and Inpainting,

Andrea Bertozzi, Selim Esedoglu, and Alan Gillette,
Analysis of a two-scale Cahn-Hilliard model for image inpainting,

Andrea Bertozzi, Selim Esedoglu, and Alan Gillette,
Inpainting of Binary Images Using the Cahn-Hilliard Equation
Rachel A. Hegemann, Erik A. Lewis, and Andrea L. Bertozzi,  
An "Estimate & Score Algorithm" for simultaneous parameter estimation and reconstruction of missing data on social networks submitted 2012.

Alexey Stomakhin, Martin B. Short, and Andrea L. Bertozzi, 
Reconstruction of Missing Data in Social Networks Based on Temporal Patterns of Interactions, Inverse Problems, 27(11), 115013, 2011.