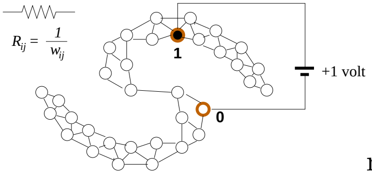
p-voltages: Laplacian Regularization for Semi-Supervised Learning on High-Dimensional Data

Nick Bridle, Xiaojin Zhu. University of Wisconsin-Madison

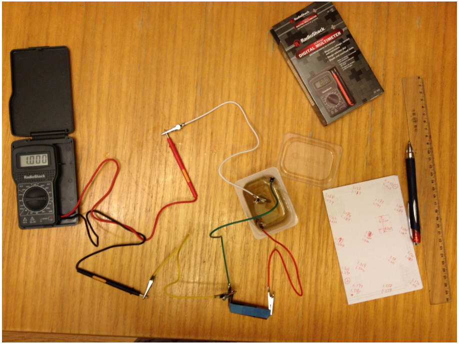
Classification on graphs

label propagation = harmonic functions = electric network

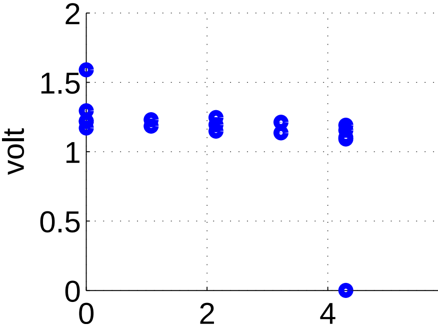


$$\min_v \left\{ \sum_{ij \in \text{edges}} \frac{(v_i - v_j)^2}{R_{ij}} \middle| v_s - v_t = 1 \right\}$$

Spike-flat pathology when $n \to \infty$



$n \approx 10^{24}$ salt water molecules



x distance in water dish

Pathology first noticed in [Nadler,Srebro,Zhou NIPS09]

Alternative world: p-electric networks

- p-Ohm's law

$$v_i - v_j = \mathrm{sign}(I_{ij})|I_{ij}|^{p-1} R_{ij}$$

- p-voltages:

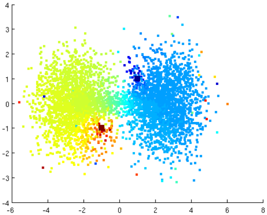
$$\min_v \left\{ \sum_{ij \in \text{edges}} \frac{|v_i - v_j|^{\frac{p}{p-1}}}{R_{ij}^{\frac{1}{p-1}}} \middle| v_s - v_t = 1 \right\}$$

  whose properties are not well-studied before.

- Our world: $p = 2$
- Suggested: $p < p^* := \frac{d}{d-1}$ [Alamgir, von Luxburg NIPS11]
  - As $p \to 1$, current concentrates itself on fewer paths

In this paper

- We prove that p-voltages are
  - not spiky around labeled nodes
  - not flat over unlabeled nodes



$(p = 1.4)$

- computationally faster than alternatives such as p-resistance
- has the potential for graph-based learning
- but, empirically does not outperform state-of-the-art Iterated Laplacian [Zhou,Belkin AISTATS11]