Random Walks and Evolving Sets: Faster Convergences and Limitations

Siu On Chan[1] Tsz Chiu Kwok[2] Lap Chi Lau[2]

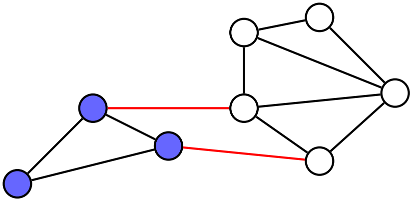
[1]The Chinese University of Hong Kong

[2]University of Waterloo

Jan 18, 2017

Finding small clusters

Given 1-regular graph with edge weights w



(Edge) expansion $\phi(S) \triangleq \dfrac{w(S, V \setminus S)}{|S|}$

Task: compute $\phi_\delta(G) \triangleq \min_{|S| \leqslant \delta n} \phi(S)$

Algorithm 1: Random walk

For finding small clusters [Spielman–Teng'04, Andersen–Chung–Lang'06, ...]

Theorem ([Kwok–Lau'12])

Random walk returns S with

$$\phi(S) = O(\sqrt{\phi(S^*)/\varepsilon}) \quad \textit{and} \quad |S| = O(|S^*|^{1+\varepsilon})$$

in poly time, when start from nice vertices in S^*

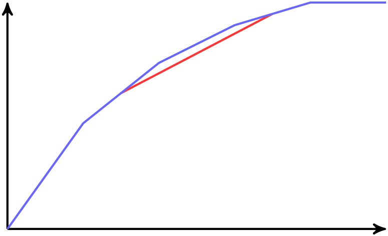
When $\varepsilon = 1/\log|S^*|$

$$\phi(S) = O(\sqrt{\phi(S^*)\log|S^*|}) \quad \text{and} \quad |S| = O(|S^*|)$$

Analysis: Lovász–Simonovits ['90]

Measure progress of lazy random walk mixing with a curve
Large $\phi(G)$ implies fast mixing

(lazy = self-loop of weight $\geqslant 1/2$ at every vertex)



More on this later

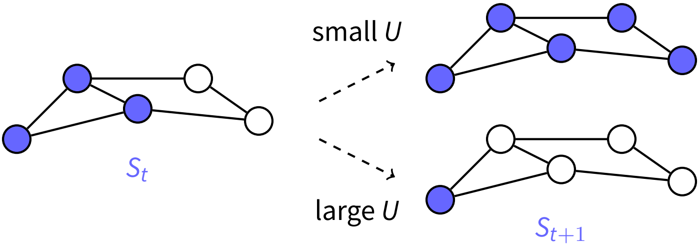
Evolving Set Process [Morris'02]

Yields strong bounds on mixing times [Morris–Peres'05]

Evolving Set Process is Markov Chain $\{S_t\}$ on subsets of V

Given S_t, choose U uniformly from $[0,1]$

$$S_{t+1} \triangleq \{v \in V \mid w(S_t, v) \geqslant U\}$$



Algorithm 2: Evolving Set Process

Can find small clusters [Andersen–Peres'09, Oveis Gharan–Trevisan'12]

Theorem ([Oveis Gharan–Trevisan'12, AOPT'16])

Evolving Set Process returns S with

$$\phi(S) = O(\sqrt{\phi(S^*)/\varepsilon}) \quad \textit{and} \quad |S| = O(|S^*|^{1+\varepsilon})$$

in poly time, when start from nice vertices in S^*

Algorithm 2: Evolving Set Process

Can find small clusters [Andersen–Peres'09, Oveis Gharan–Trevisan'12]

Theorem ([Oveis Gharan–Trevisan'12, AOPT'16])

Evolving Set Process returns S with

$$\phi(S) = O(\sqrt{\phi(S^*)/\varepsilon}) \quad \textit{and} \quad |S| = O(|S^*|^{1+\varepsilon})$$

in poly time, when start from nice vertices in S^*

Conjecture [Oveis Gharan'13, AOPT'16]

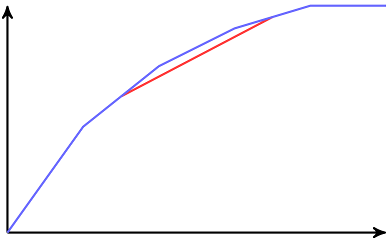
With non-trivial probability, in fact $|S| = O(|S^*|)$

If true, will refute Small-Set-Expansion Hypothesis
that $\phi_\delta(G)$ is hard to approximate [Raghavendra–Steurer'10]
(cousin of Unique-Games Conjecture)

Our results: Generalized Lovász–Simonovits analysis

Also measure random walk mixing with LS curve

1. Large combinatorial gap $\varphi(G)$ implies fast mixing
2. Large robust vertex expansion $\phi^V(G)$ + laziness imply fast mixing



Our results: Combinatorial gap

Large combinatorial gap $\varphi(G)$ implies fast mixing

$$\varphi(G) \triangleq \min_{|S|=|T|\leqslant n/2} 1 - \frac{w(S,T)}{|S|}$$

1. Similar definition as $\phi(G)$ (which only allows $T = S$)
2. $\varphi(G) = \phi(G)$ for lazy graphs
3. $\varphi(G)$ small if G has near-bipartite component

Corollary (Expansion of graph powers)

$$\phi_{\delta/4}(G^t) = \min\{\Omega(\sqrt{t}\phi_\delta(G)), 1/20\}$$

without laziness assumption of [Kwok–Lau'14]

Our results: Vertex expansion

Robust vertex expansion ϕ^V defined by [Kannan–Lovász–Montenegro'06]

Theorem

Evolving Set Process returns S with

$$\phi(S) = O(\phi(S^*)/(\varepsilon\phi^V(S))) \quad \textit{and} \quad |S| = O(|S^*|^{1+\varepsilon})$$

in poly time, when start from nice vertices in S^*

Compare: $\phi(S) = O(\sqrt{\phi(S^*)/\varepsilon})$ in [APOT'16]

Our results: Vertex expansion

Robust vertex expansion ϕ^V defined by [Kannan–Lovász–Montenegro'06]

Theorem

Evolving Set Process returns S with

$$\phi(S) = O(\phi(S^*)/(\varepsilon\phi^V(S))) \quad \textit{and} \quad |S| = O(|S^*|^{1+\varepsilon})$$

in poly time, when start from nice vertices in S^*

Compare: $\phi(S) = O(\sqrt{\phi(S^*)/\varepsilon})$ in [APOT'16]

Implies constant factor approximation when

$$\phi^V(G) \triangleq \min_{|S| \leqslant n/2} \phi^V(S) \quad \text{is } \Omega(1)$$

Evolving Set Process analog of spectral partitioning result in [Kwok–Lau–Lee'16]

Our results: Hard instances

Theorem

For arbitrarily small δ, some graphs have a hidden small cluster S^*

$$\phi(S^*) \leqslant \varepsilon \quad \textit{and} \quad |S^*| = \delta n$$

but Evolving Set Process never returns S with

$$\phi(S) \leqslant 1 - \varepsilon \quad \textit{and} \quad |S| \leqslant \delta_\varepsilon n$$

Refute the conjecture in [Oveis Gharan'13, AOPT'16]
Also limitation of random walk, PageRank, etc

Details:
Generalized Lovász–Simonovits Analysis

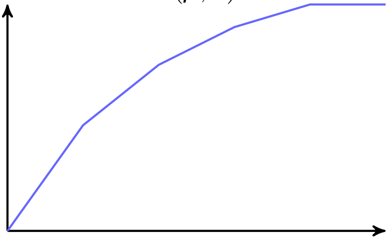
Lovász–Simonovits

Let p be probability vector

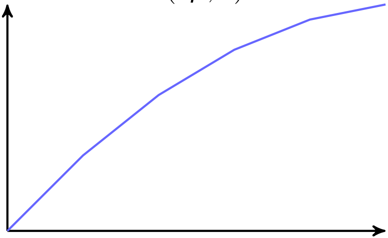
Measure random walk mixing with curve $C(p, x)$

$C(p, x) \triangleq$ sum of first x largest elements in p for integer x

$C(p, x)$



$C(Ap, x)$

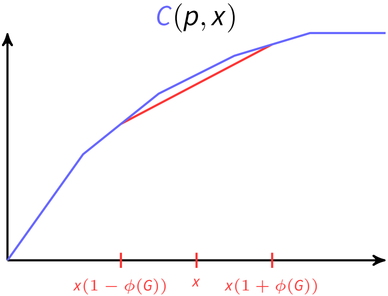


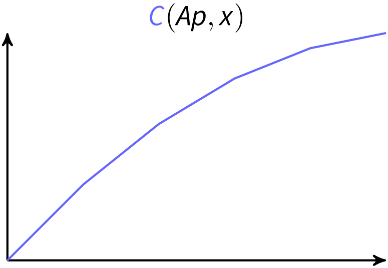
Lovász–Simonovits

Let p be probability vector

Measure random walk mixing with curve $C(p, x)$

$C(p, x) \triangleq$ sum of first x largest elements in p for integer x





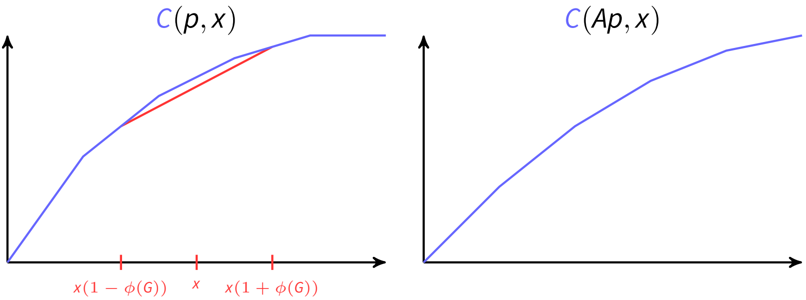
When lazy $\quad C(Ap, x) \leqslant \frac{1}{2}\left(C(p, x(1-\phi(G))) + C(p, x(1+\phi(G)))\right)$

Lovász–Simonovits

Let p be probability vector

Measure random walk mixing with curve $C(p, x)$

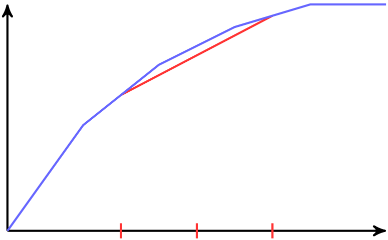
$C(p, x) \triangleq$ sum of first x largest elements in p for integer x



When lazy $$C(Ap, x) \leqslant \frac{1}{2}\left(C(p, x(1-\phi(G))) + C(p, x(1+\phi(G)))\right)$$

By induction: $$C(A^t p, x) \leqslant \frac{x}{n} + \sqrt{x}\left(1 - \frac{\phi(G)^2}{8}\right)^t$$

Generalizing Lovász–Simonovits



$$C(Ap,x) \leqslant \frac{1}{2}\left(C(p,x(1-\varphi(G))) + C(p,x(1+\varphi(G)))\right)$$

$\varphi(G)$ in place of $\phi(G)$ laziness not required

$\varphi(G) = \phi(G)$ for lazy graphs $\Rightarrow$ generalizing LS to non-lazy

More intuitive analysis

Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

Sort vertices by decreasing $d_S(i) \triangleq w(i, S)$

$$\begin{aligned}(Ap)(S) &= \sum_{0\leqslant i\leqslant n} d_S(i)p(i) \\ &= \sum_{0\leqslant i\leqslant n} (d_S(i) - d_S(i+1)) \sum_{1\leqslant j\leqslant i} p(j)\end{aligned}$$

Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

Sort vertices by decreasing $d_S(i) \triangleq w(i, S)$

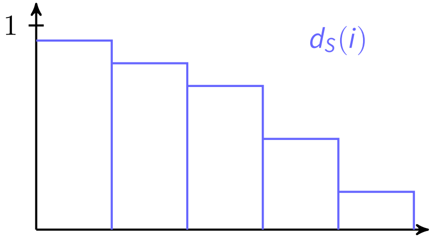
$$\begin{aligned}(Ap)(S) &= \sum_{0 \leqslant i \leqslant n} d_S(i)p(i) \\ &= \sum_{0 \leqslant i \leqslant n} (d_S(i) - d_S(i+1)) \underbrace{\sum_{1 \leqslant j \leqslant i} p(j)}_{\leqslant C(p,i)}\end{aligned}$$

Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

Sort vertices by decreasing $d_S(i) \triangleq w(i, S)$

$$\begin{aligned}(Ap)(S) &= \sum_{0\leqslant i\leqslant n} d_S(i)p(i) \\ &= \sum_{0\leqslant i\leqslant n} (d_S(i) - d_S(i+1)) \underbrace{\sum_{1\leqslant j\leqslant i} p(j)}_{\leqslant C(p,i)}\end{aligned}$$

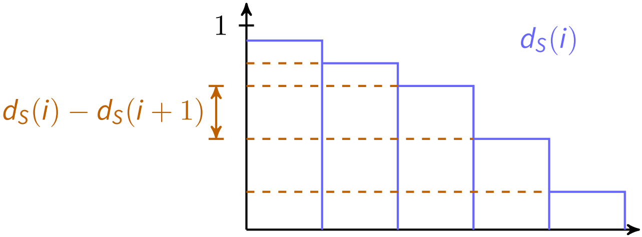


Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

Sort vertices by decreasing $d_S(i) \triangleq w(i, S)$

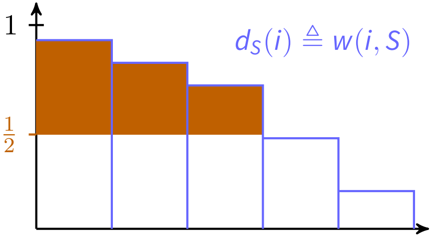
$$\begin{aligned}(Ap)(S) &= \sum_{0\leqslant i\leqslant n} d_S(i)p(i) \\ &= \sum_{0\leqslant i\leqslant n} (d_S(i) - d_S(i+1)) \underbrace{\sum_{1\leqslant j\leqslant i} p(j)}_{\leqslant C(p,i)}\end{aligned}$$



Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

From earlier $\quad (Ap)(S) \leqslant \sum_{0 \leqslant i \leqslant n} (d_S(i) - d_S(i+1)) C(p, i)$

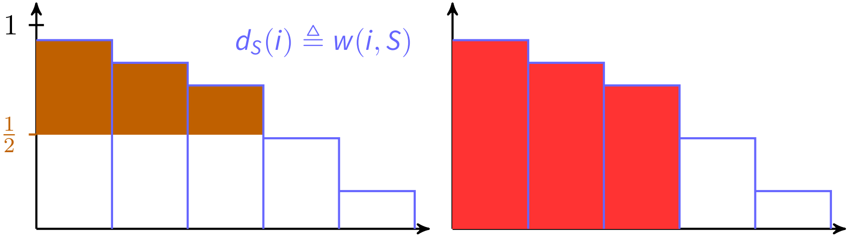


Upper Area $\leqslant \quad \frac{1}{2}|S|(1-\varphi(G))$

Sketch of main ideas

$$C(Ap, |S|) \leqslant \frac{1}{2}\left(C(p, |S|(1-\varphi(G))) + C(p, |S|(1+\varphi(G)))\right)$$

From earlier $$(Ap)(S) \leqslant \sum_{0\leqslant i\leqslant n} (d_S(i) - d_S(i+1))C(p,i)$$



$$\text{Upper Area} \leqslant \frac{1}{2}\sum_{1\leqslant i\leqslant |T|} d_S(i) \leqslant \frac{1}{2}|S|(1-\varphi(G))$$

Details:
Hard instances for Evolving Set Process

Hard instances: Noisy hypercube

Noisy k-ary hypercube $\quad (k = 1/\delta)$

- k^d vertices, each represented by a string of length d over $[k]$
- Transition probability from x to y:

When $x = x_1 x_2 \ldots x_d$, $\quad y_i = \begin{cases} x_i & \text{prob } \varepsilon \\ \text{uniformly from } [k] & \text{prob } 1-\varepsilon \end{cases}$

Coordinate cut $S = \{x \mid x_1 = 0\}$ satisfies

$$\phi(S) \leqslant \varepsilon \quad \text{and} \quad |S| = \delta n$$

Why local algorithms fail

Lets start from $S_0 = \{\vec{0}\}$

Evolving Set Process treats all vertices with the same Hamming weight equally

The process only explores sets S_t that are symmetric under coordinate permutations (in fact, Hamming balls)

Small Hamming balls on noisy k-ary hypercube are expanding
[Chan–Mossel–Neeman'14]

Hamming balls B expand

$$B = \{x \in [k]^d \mid |x| \leqslant r\}$$

$$\phi(B) = \frac{\Pr_{x \sim y}[x, y \in B]}{\Pr_x[x \in B]}$$

Hamming balls B expand

$$B = \{x \in [k]^d \mid |x| \leqslant r\}$$

$$\phi(B) = \frac{\Pr_{x\sim y}[x, y \in B]}{\Pr_x[x \in B]}$$

$$\Pr_x[x \in B] = \mathbb{E}_x[\mathbb{1}_{x_1 \neq 0} + \cdots + \mathbb{1}_{x_d \neq 0} \leqslant r]$$

$$\approx \Pr[g \leqslant r'] \qquad \text{Gaussian } g$$

$$\Pr_{x\sim y}[x, y \in B] = \mathbb{E}_{x,y}\left[\sum_{1 \leqslant i \leqslant d} \mathbb{1}_{x_i \neq 0} \leqslant r \text{ and } \sum_{1 \leqslant i \leqslant d} \mathbb{1}_{y_i \neq 0} \leqslant r\right]$$

$$\approx \Pr[g \leqslant r' \text{ and } h \leqslant r'] \qquad \text{Gaussians } g, h$$

Hamming balls B expand

$$B = \{x \in [k]^d \mid |x| \leqslant r\}$$

$$\phi(B) = \frac{\Pr_{x \sim y}[x, y \in B]}{\Pr_x[x \in B]} \approx \frac{\Pr_{g,h}[g, h \leqslant r']}{\Pr_g[g \leqslant r']}$$

$$\Pr_x[x \in B] = \mathbb{E}_x[\mathbb{1}_{x_1 \neq 0} + \cdots + \mathbb{1}_{x_d \neq 0} \leqslant r]$$

$$\approx \Pr[g \leqslant r'] \qquad \text{Gaussian } g$$

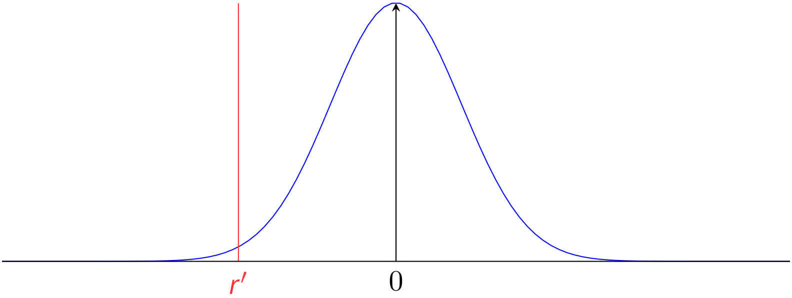
$$\Pr_{x \sim y}[x, y \in B] = \mathbb{E}_{x,y}\left[\sum_{1 \leqslant i \leqslant d} \mathbb{1}_{x_i \neq 0} \leqslant r \text{ and } \sum_{1 \leqslant i \leqslant d} \mathbb{1}_{y_i \neq 0} \leqslant r\right]$$

$$\approx \Pr[g \leqslant r' \text{ and } h \leqslant r'] \qquad \text{Gaussians } g, h$$

Expansion in Gaussian space

$$B = \{x \in [k]^d \mid |x| \leqslant r\}$$

$$\phi(B) = \frac{\Pr_{x\sim y}[x, y \in B]}{\Pr_x[x \in B]} \approx \frac{\Pr_{g,h}[g, h \leqslant r']}{\Pr_g[g \leqslant r']} \overset{r' \to 0}{\to} 1$$



Symmetry

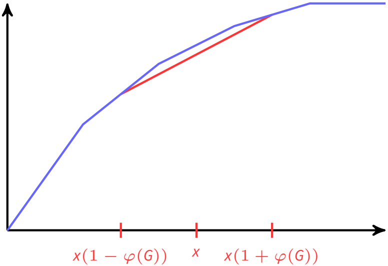
Obstacles to all local clustering algorithms
Evolving Set Process, random walk, PageRank, etc
All fall for the symmetry

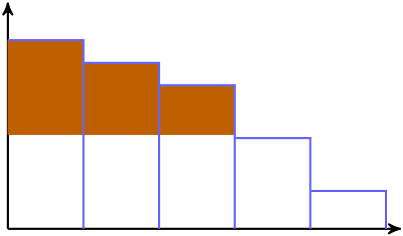
Folded noisy hypercubes are hard instances for SDP
Folding necessary to rule out sparse cuts in those instances
In our situation, we cannot fold

Our instances are easy for Lasserre/sum-of-squares

Summary

New analysis of Lovász–Simonovits curve
Faster convergence with large combinatorial gap $\varphi(G)$ or robust vertex expansion $\phi^V(G)$





Limitations of all local clustering algorithms:
Coordinate cuts vs Hamming balls under symmetry

Open problems

$\Omega(\sqrt{\phi(S^*)\log|S^*|})$ lower bound for Evolving Set Process and random walk?

Thank you