Applications of the Chen-Stein method to random graphs

Ayalvadi Ganesh
University of Bristol

Workshop on New Directions in Stein's method, Singapore

Outline

- Stein's method widely applied to problems in stochastic geometry and random graphs
- Discuss two applications today
  - Small world networks
  - Colouring random geometric graphs
- Motivation for many such problems comes from wireless networks

Connectivity of spatial small-world networks

Joint work with Feng Xue

Background: Spatial random graphs

- n nodes placed independently and uniformly at random on unit torus

- Random geometric graph (RGG)
  - Also known as Gilbert's disk model
  - Edge present between two nodes if they are within distance $r = r(n)$ of each other
- *k*-nearest neighbour graph (*k*-NN)
  - Edge present between two nodes if either is among the $k = k(n)$ nearest neighbours of the other

Connectivity

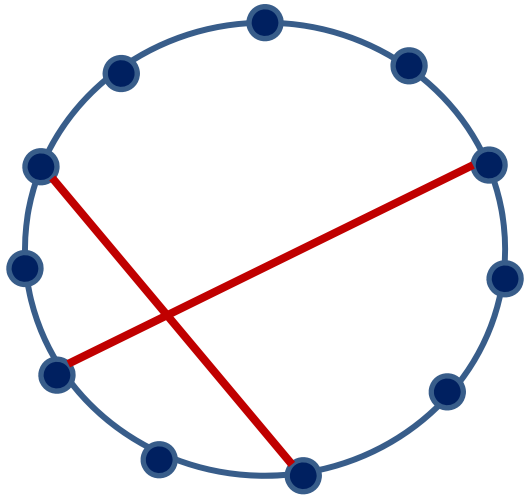
- *For what parameter values are these random graphs connected?*

- RGG (Penrose, 2003): Threshold for connectivity at $\pi r^2 n = \log n$

- k −NN graph (Balister, Bollobas, Sarkar & Walters, 2009; Falgas-Ravry & Walters, 2012): Threshold for connectivity at $\pi r^2 n = \mathrm{c} \log n$

Background: Small world networks

- Start with a lattice / ring / torus
- Augment with random *shortcuts*



- *How does diameter scale?*

Diameter scalings: Barbour & Reinert

- Model: uniformly distributed shortcuts
  - Shortcut between every pair of nodes present with probability p/n, independent of other
- Scalings: Let $p \sim n^{\alpha}$, $\alpha \in (-1,1)$
- Distance between nodes scales as:
  - a fractional power of n (< 0)
  - logarithmic in n ($= 0$)
  - constant $(\alpha > 0)$

Diameter scalings : Coppersmith, Gamarnik & Sviridenko

- Model: distance-biased shortcuts
  - Lattice on $\{1, \dots, n\}^d$
  - Shortcut (x, y) present with prob. $\propto \ |x - y|^{-s}$
- Distance between nodes scales as
  - fractional power of n $(s \geq 2d)$
  - poly-logarithmic in n $(d < s < 2d)$
  - $\frac{\log n}{\log \log n}$ $(s = d)$

Spatial small-world network models

- Start with RGG or k-NN
- Augment with random shortcuts, present with probability p between each pair of nodes, independent of everything else
- Union of spatial and Erdős-Rényi random graphs
- *For what parameter values are the graphs connected?*

Results: RGG + shortcuts model

- Suppose that $n\pi r^2 + np = \log n + c$
- Then, *P(connectivity)* $\leq e^{-e^{-c}} + O\left(\frac{\log n}{\sqrt{n}}\right)$

Idea of proof:

Number of isolated nodes is approximately Poisson distributed with parameter e^{-c}

(Chen-Stein)

Results: k-NN + shortcuts model

- Suppose that $\delta > 0$, $k/n \to 0$, and
$$(k+1)np > 2(1+\delta)\log\frac{n}{k+1}$$

- Then, with high probability, the graph is connected, and its diameter is bounded above by $7\left(\log\frac{n}{k+1}+1\right)$

Colouring random geometric graphs

Joint work with Divya Mohan and Simon Armour

Model and problem statement

- N nodes distributed independently and uniformly on the unit square
- K channels or colours available
- Each node has to be assigned a colour
- Think of $N, K \to \infty, N \gg K$

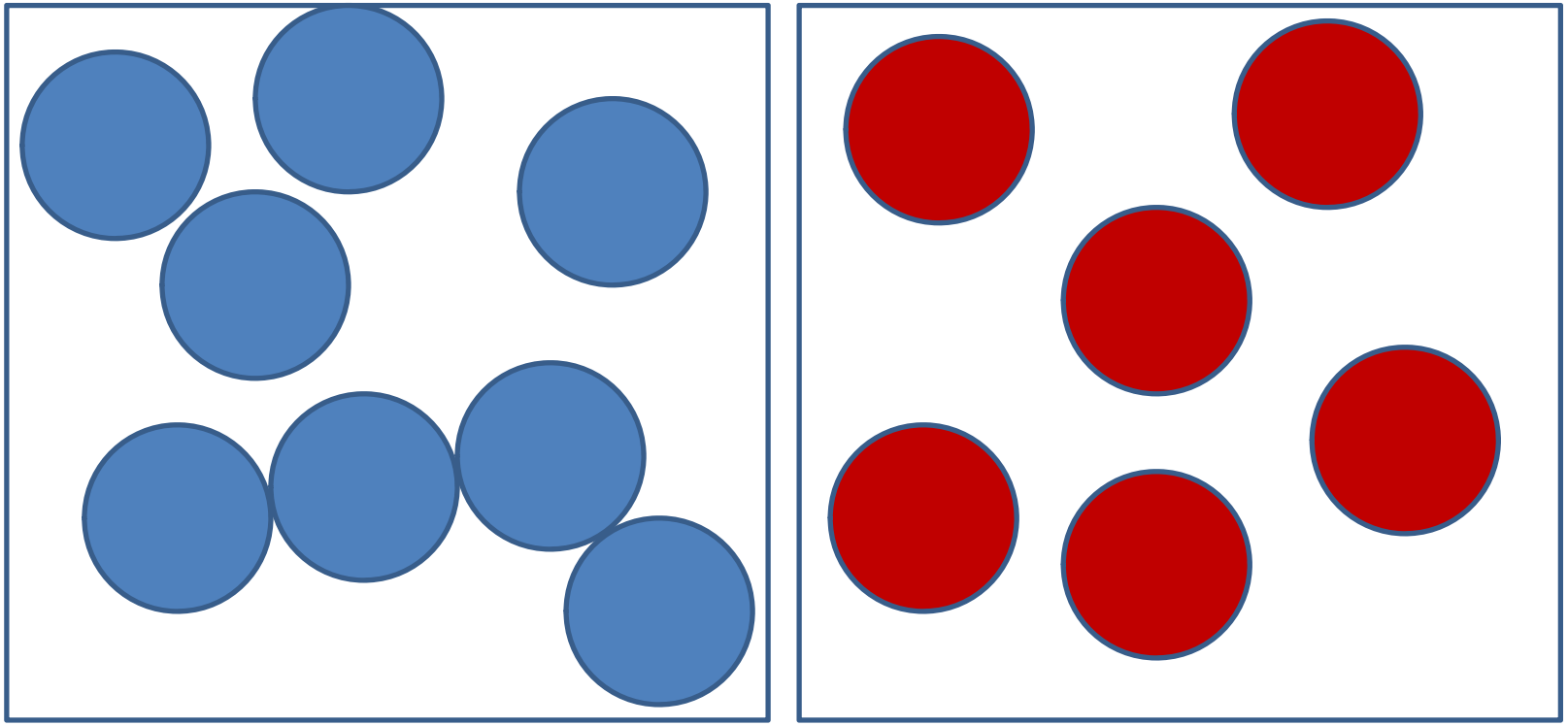
- *Objective: maximise D_{min} ,the minimum distance between two nodes with the same colour*

Background

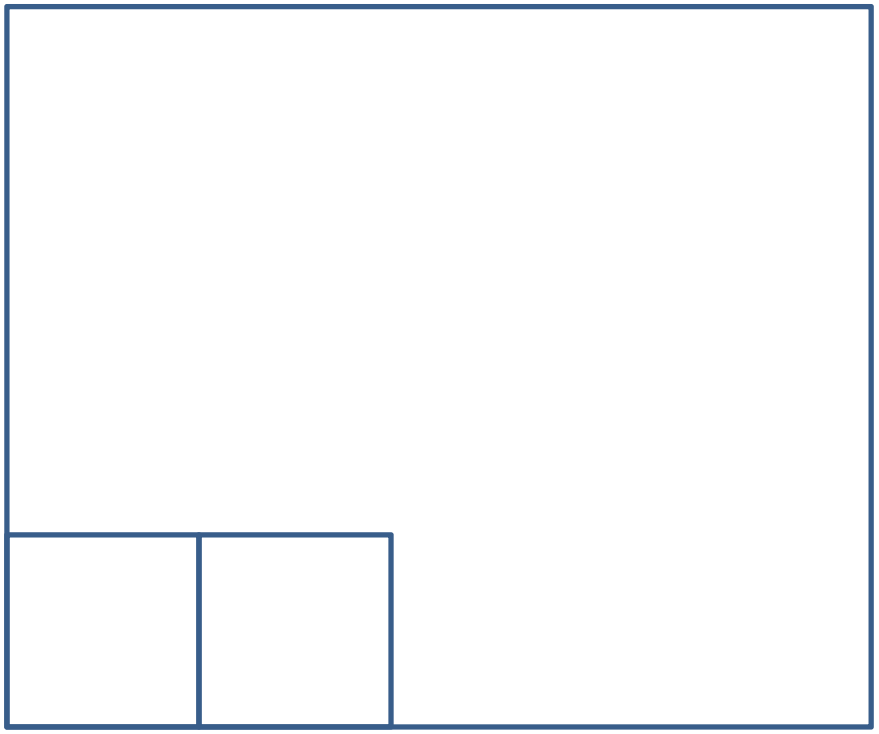
- Problem posed by Ni, Srikant and Wu (2011), who showed the following:
  - $D_{min} \leq 2\sqrt{K/\pi N}$ for any colouring algorithm and any node configuration
  - $P(D_{min} < \frac{K}{N}) \to 1$ for nodes placed and coloured independently and uniformly at random

- Big gap!
  - Is there a colouring algorithm that can do better?
  - Is there one that is decentralised?

Upper bound on D_{min} for arbitrary algorithms and node configurations



Upper bound on D_{min} for random placement and colouring



Greedy colouring algorithm

- Proposed by Ni, Srikant and Wu (2011)
- Order the nodes arbitrarily, and colour them sequentially, picking a best colour at each step
- They showed that, if $K = \Omega(\log N)$, then
  - $D_{min} = \Omega(\sqrt{K/N})$ for the greedy algorithm, and
  - $D_{min} = O(\sqrt{K/N})$ for any algorithm

Results: Random colouring

Theorem (G, Mohan & Armour):

The sequence of random variables $\frac{N}{\sqrt{K}} D_{min}$ converge in distribution to a Rayleigh random variable, i.e., for all $x > 0$,

$$P\left(D_{min} > \frac{\sqrt{K}}{N} x\right) \to e^{-x^2/2}$$

Proof idea

- Fix $x > 0$. For any two nodes u and v,

$$P(|u - v| < x \text{ and } u, v \text{ have same colour }) = \frac{\pi x^2}{K}$$

- Events for distinct node pairs aren't independent, but dependence is weak

- Number of node pairs is $\approx N^2/2$

Proof idea (continued)

- Chen-Stein method: Random number of node pairs satisfying above property approximately Poisson$\left(\frac{\pi x^2 N^2}{2K}\right)$

- $\{D_{min} > x\} \Leftrightarrow$ no such node pairs

$$P(D_{min} > x) \approx P\left(\text{Poisson}\left(\frac{\pi x^2 N^2}{2K}\right) = 0\right)$$

Node colouring is a game

- Players are nodes, actions are colours
- The payoff to a player is the negative of its distance to the nearest node with the same colour
- Pure Nash equilibrium: colouring in which no single node benefits by changing its colour
- Could have multiple Nash equilibria

Questions

- Are there decentralised dynamics for the players (in discrete or continuous time) that are guaranteed to converge to a Nash equilibrium?

- If so, how long does it take?

Greedy algorithm

- Nodes update their colours according to independent clocks
  - choosing a colour to maximise their distance to another node of the same colour
  - choice could be same as current colour
- Corresponds to *asynchronous best response dynamics* in the game

Convergence of greedy algorithm

- D_{min} is non-decreasing, and
  - is either a Nash equilibrium, or
  - can be increased by re-colouring some node
- Only finitely many possible colourings
- Must reach Nash equilibrium in finite time, under mild assumptions
- But time could be exponentially large
- We analyse performance after each node has performed at least one update step

Performance of greedy algorithm: lower bound on D_{min}

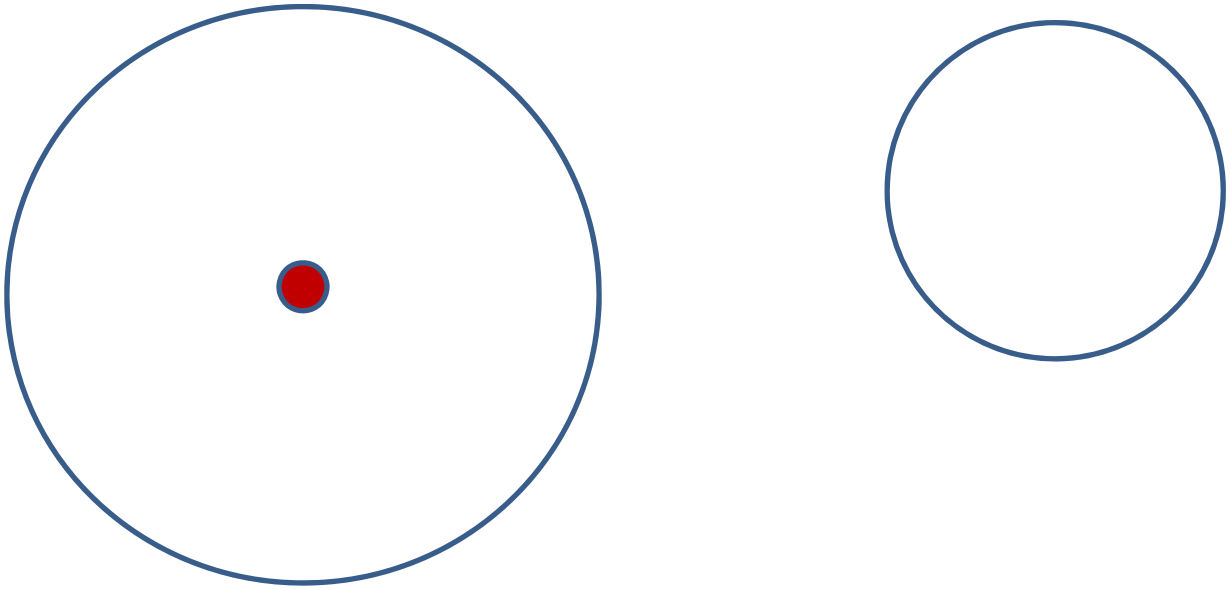
- Greedy algorithm always picks a colour different from $K-1$ nearest neighbours
- So $D_{min} < x \Rightarrow$ there is a node u such that $B(u,x)$has K or more nodes in it
- Number of nodes in $B(u,x)$ is $Bin(n, \pi x^2)$
- Large deviations for Binomial + union bound

Upper bound on D_{min}

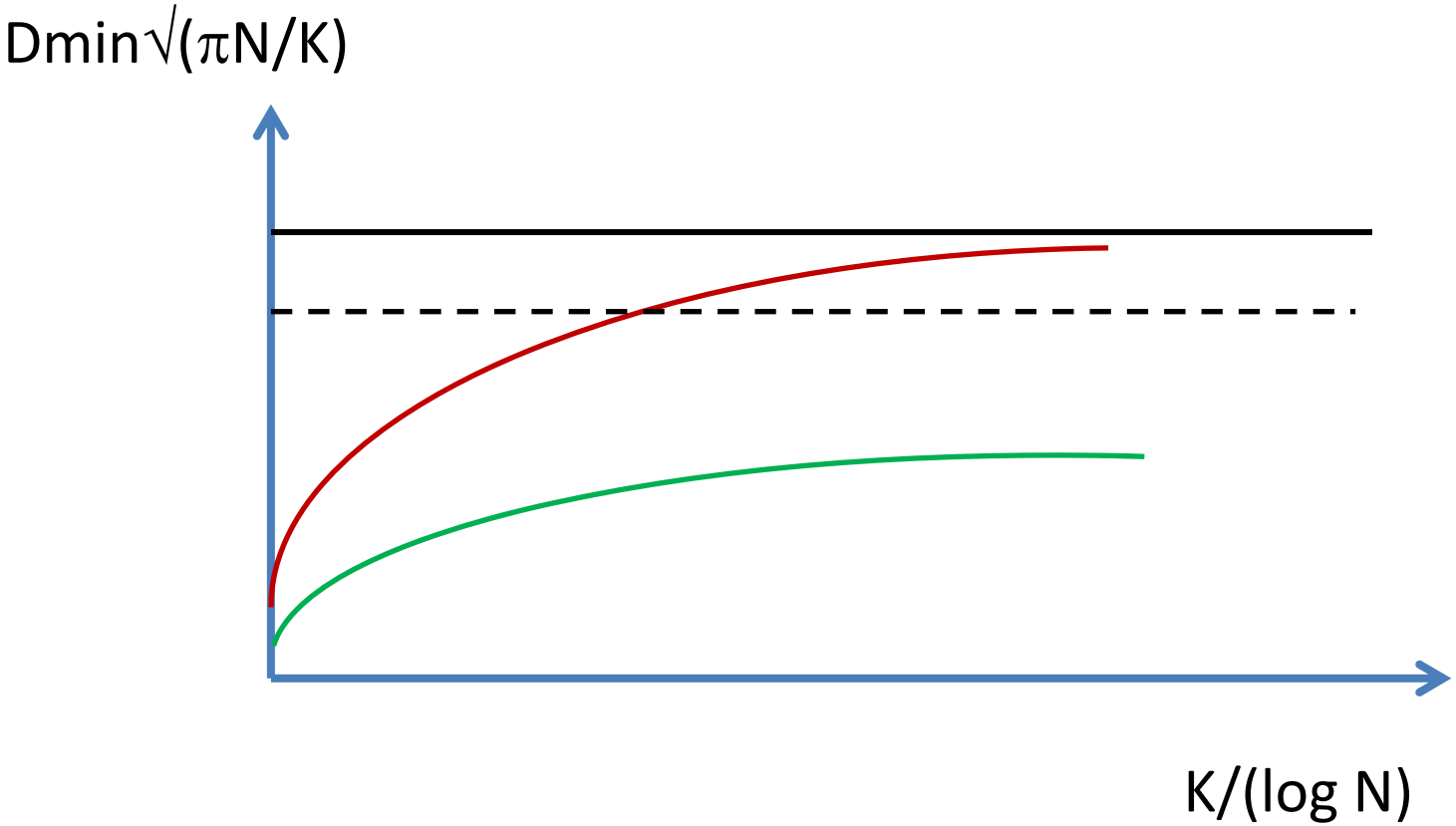
- If there is a circle of diameter x that contains $K + 1$ or more nodes, then $D_{min} < x$

- Large deviations for Binomial + second moment method

Bounds on D_{min} in pictures



Bounds on D_{min} for greedy algorithm



Bounds on D_{min}

- $G(n, r)$: geometric random graph on n nodes with threshold distance r
- For any graph G, denote
  - $\chi(G)$: chromatic number
  - $\omega(G)$: clique number
  - $\Delta(G)$: maximum degree

$$D_{min} < x \qquad (G(n, x)) > K$$

Bounds on D_{min}

- For any graph G, we have

$$\omega(G) \leq \chi(G) \leq \Delta(G) + 1$$

- Which of these is closer to $\chi(G)$?

- McDiarmid and Muller give bounds on the ratio of chromatic number to clique number