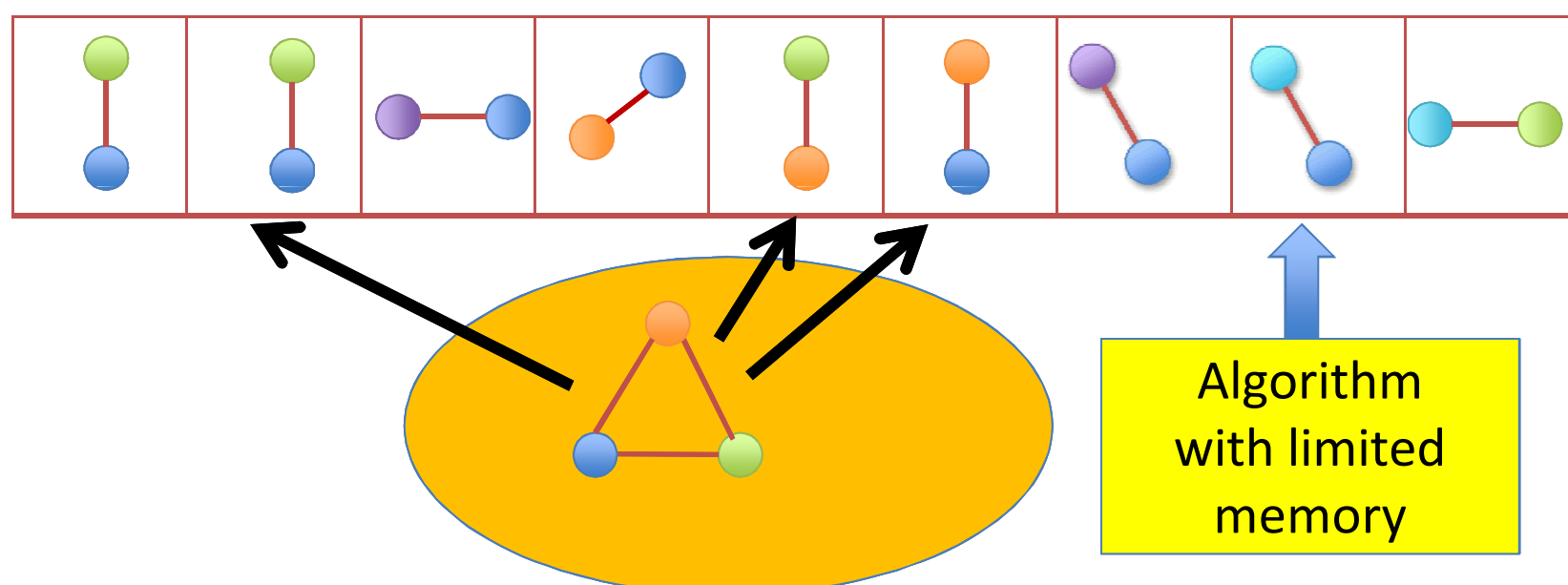
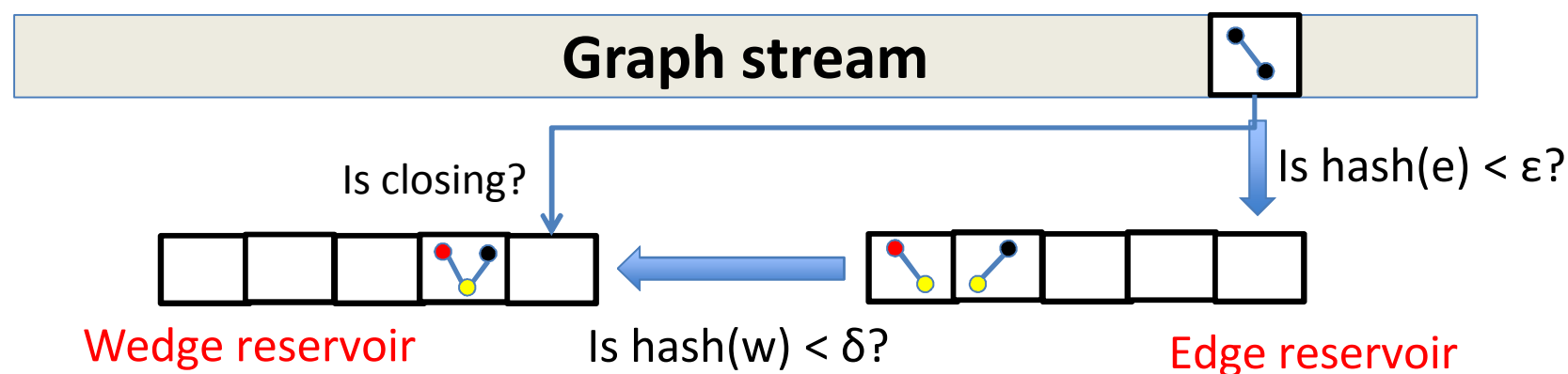
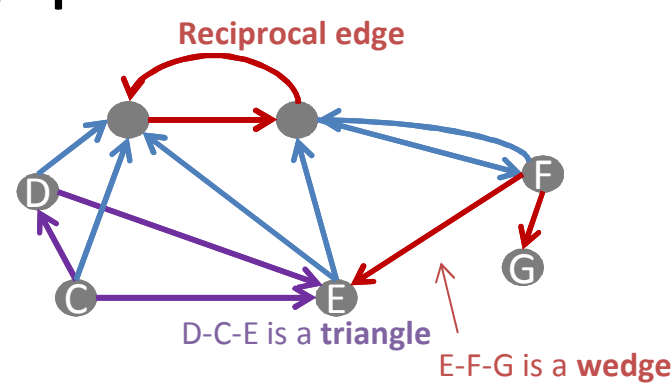


Streaming and Structural Insights

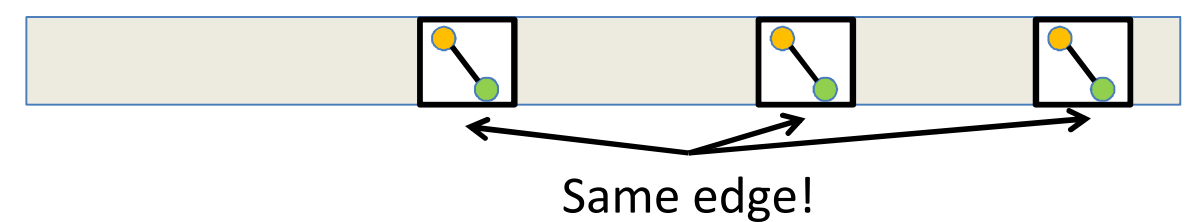
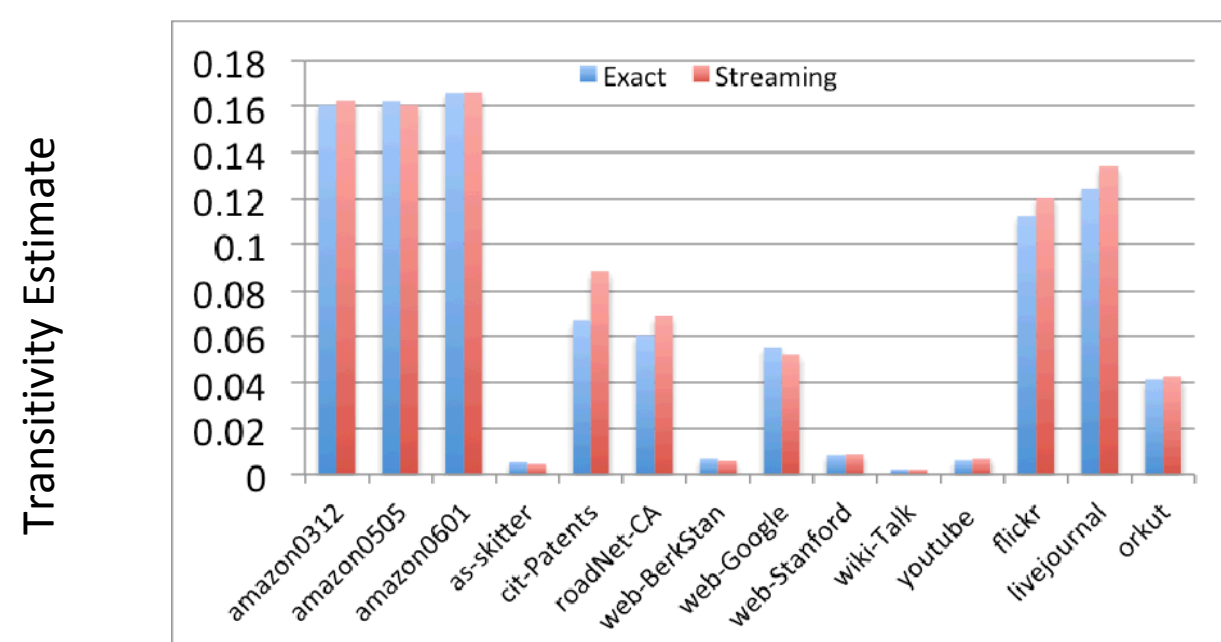


How to main structural properties of graph with limited memory footprint?

- Triangle count?
- Reciprocity?
- Degree distribution?



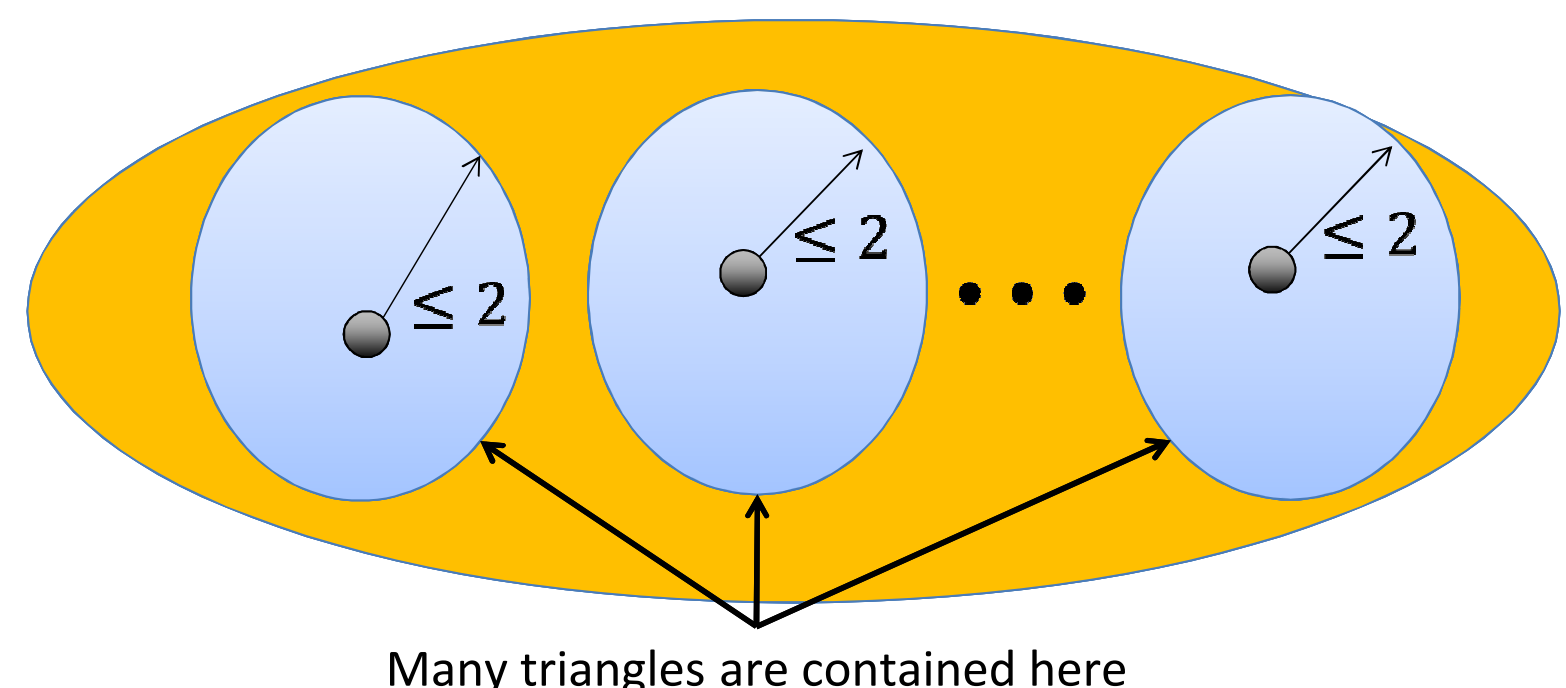
- Only space maintained are reservoirs: orders of magnitude small than stream
- All estimates: triangles, reciprocity, etc. obtained from reservoir
- Provable guarantees on low error on small space



- Real data forms multi-graphs not simple graphs!
- Often ignored in research, but major hassle in practice
- Our algorithms work for multi-graphs as well
- Use of hashing techniques (ack. to Ashish Goel) and unbiasing methods
- First triangle algorithm in literature to handle multi-graphs

From triangles to global structure

- We observe high density of triangles. Does this in of itself imply something about global structure?
- Any graph with high density of triangles satisfies special decompositions
 - Constant fraction of triangle present in dense components of radius 2
- Local structure (triangle density) \rightarrow global structure



References

- M. Jha, C. Seshadhri, and A. Pinar. *A Space Efficient Streaming Algorithm for Triangle Counting Using the Birthday Paradox*, 2013 SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), pp 589-597, 2013
- M. Jha, C. Seshadhri, and A. Pinar. *When a Graph is not so Simple: Counting Triangles in Multigraph Streams*, arXiv:1310.7665, 2013
- R. Gupta, T. Roughgarden, and C. Seshadhri. *Decompositions of Triangle-Dense Graphs*, 2014 Innovations in Theoretical Computer Science, To appear.