Private Release of Graph Statistics using Ladder Functions

Jun Zhang[1] Graham Cormode[2] Cecilia Procopiuc[3] Divesh Srivastava[4] Xiaokui Xiao[1]

1.Nanyang Technological University
{jzhang027, xkxiao}@ntu.edu.sg

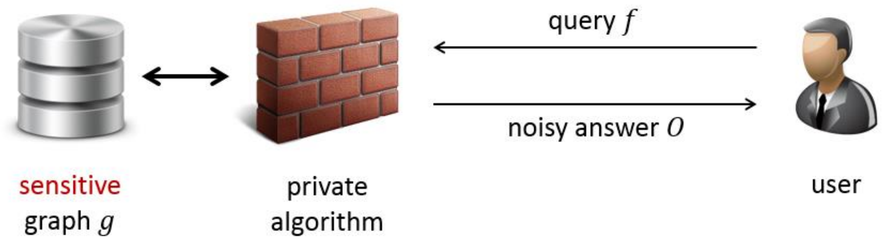
2. University of Warwick
g.cormode@warwick.ac.uk

3. Google Inc.
mpro@google.com

4. AT&T Labs -- Research
divesh@research.att.com

1. Private Release of Graph Statistics

❑ Private Data Release



Privacy: O should reveal little about any relationship in g

Utility: O should be as accurate as possible

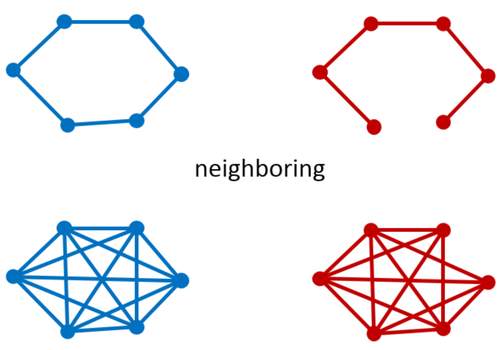
❑ Differential Privacy on Graph

Differentially private algorithm injects noise into the query answer, in order to cover the **maximum impact of a relationship (an edge)**.

query1: are nodes a and b connected?

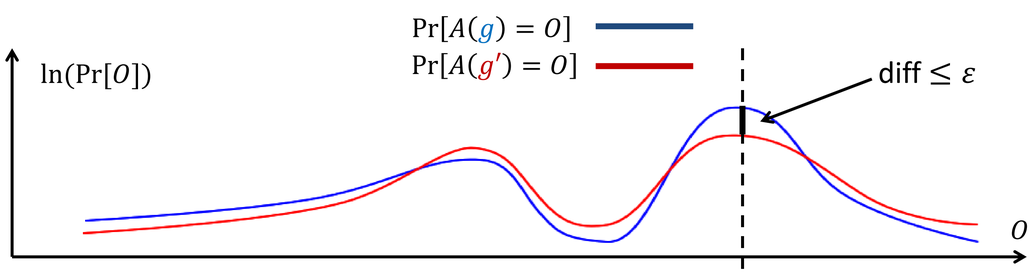
query2: how many edges?

query3: how many triangles?



2. Global and Local Sensitivity

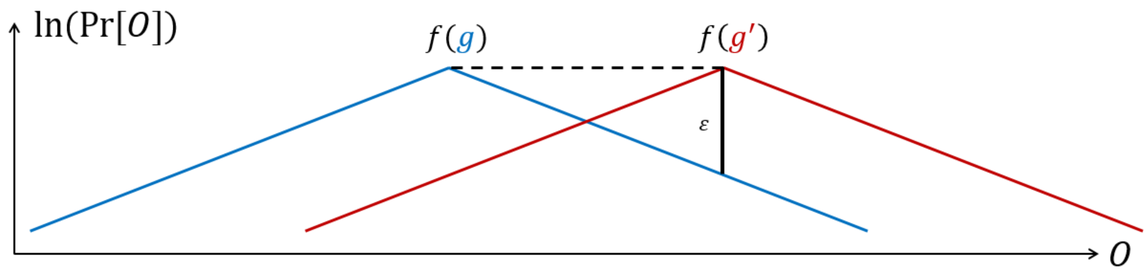
❑ Formal Definition of Differential Privacy



where g and g' are neighboring graphs that differ by **at most one** edge

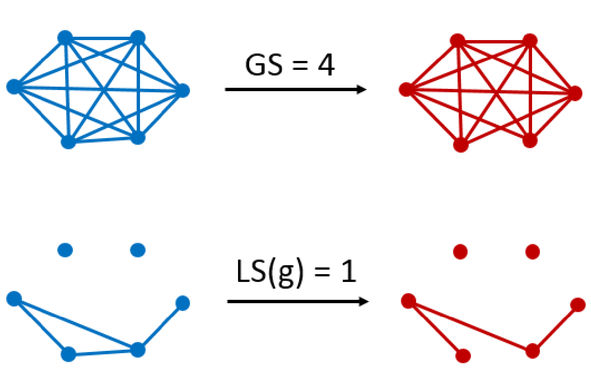
$$|\ln(\Pr[A(g) = O]) - \ln(\Pr[A(g') = O])| \leq \varepsilon$$

❑ Global Sensitivity



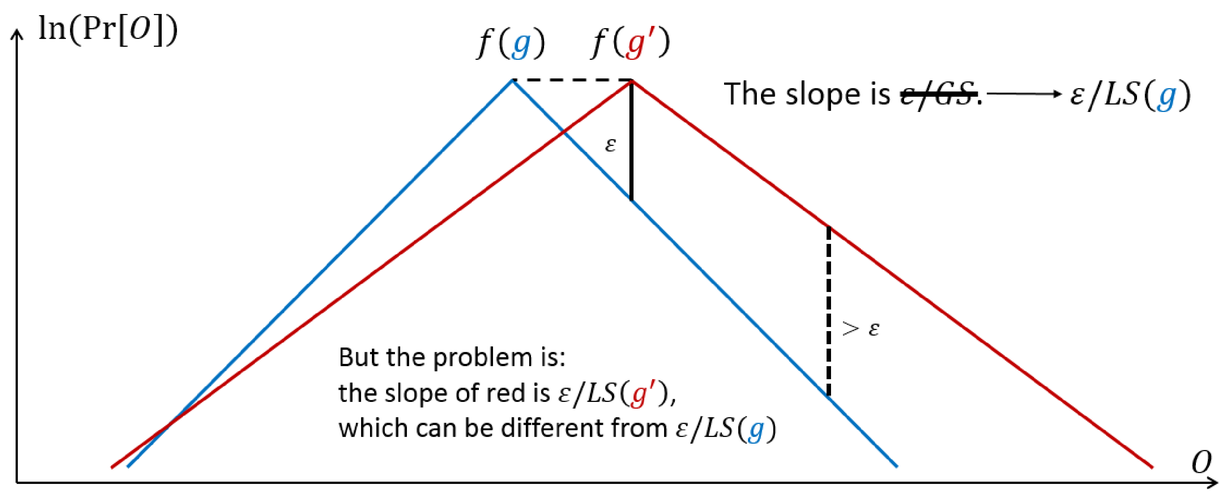
The slope is ε/GS. $GS = \max_{g,g'} |f(g) - f(g')|$

❑ Local Sensitivity



$GS = \max_{g,g'} |f(g) - f(g')|$

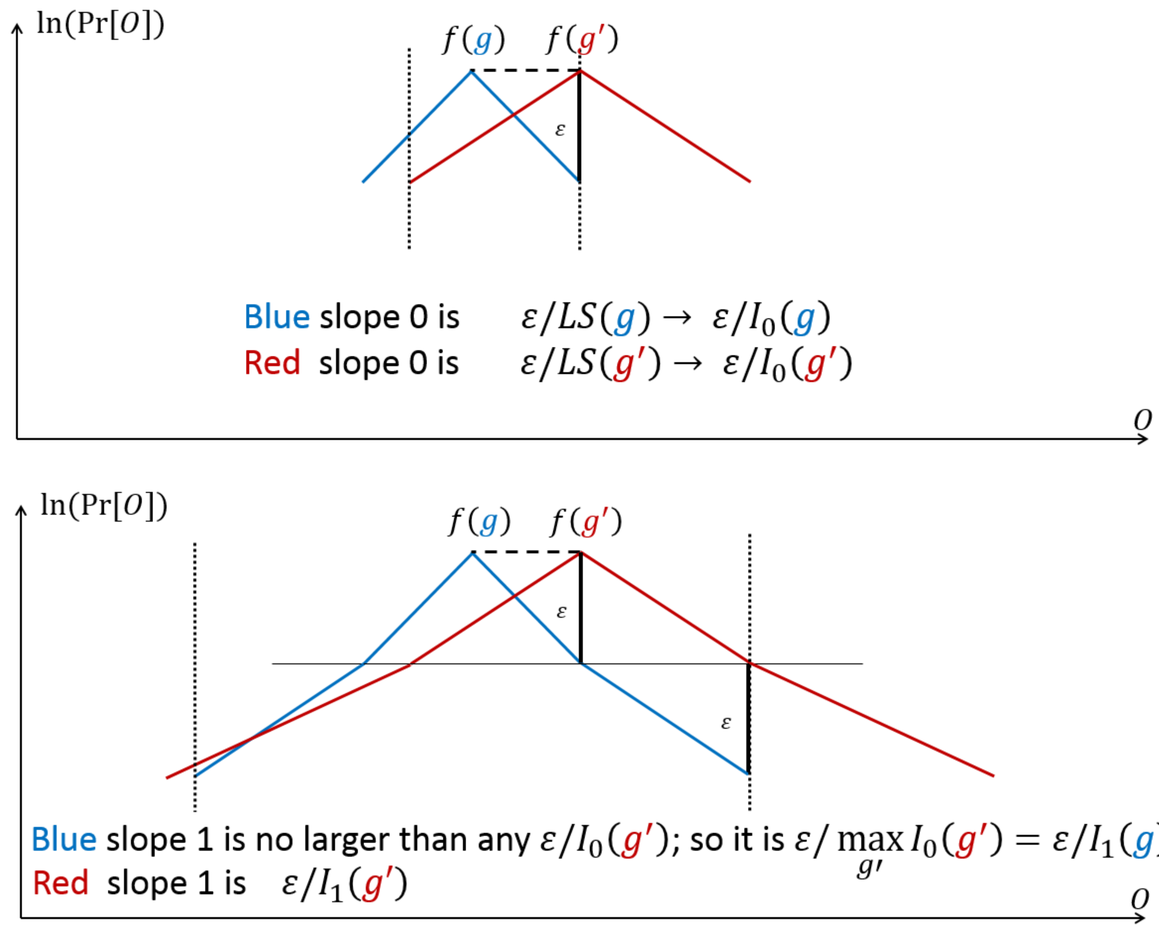
$LS(g) = \max_{g'} |f(g) - f(g')|$



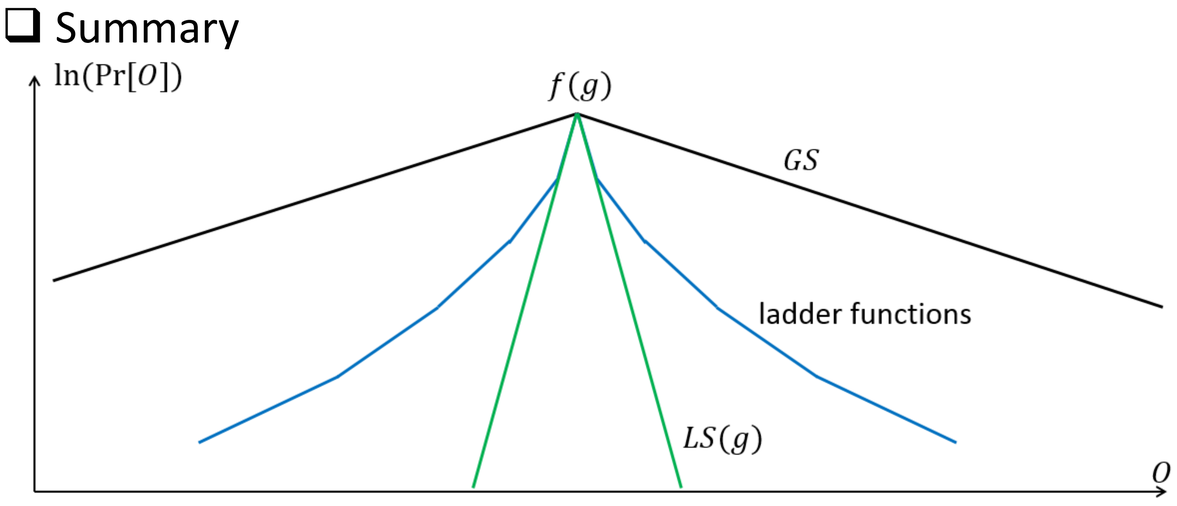
The slope is ~~ε/GS~~. $\longrightarrow \varepsilon/LS(g)$

3. Ladder Functions

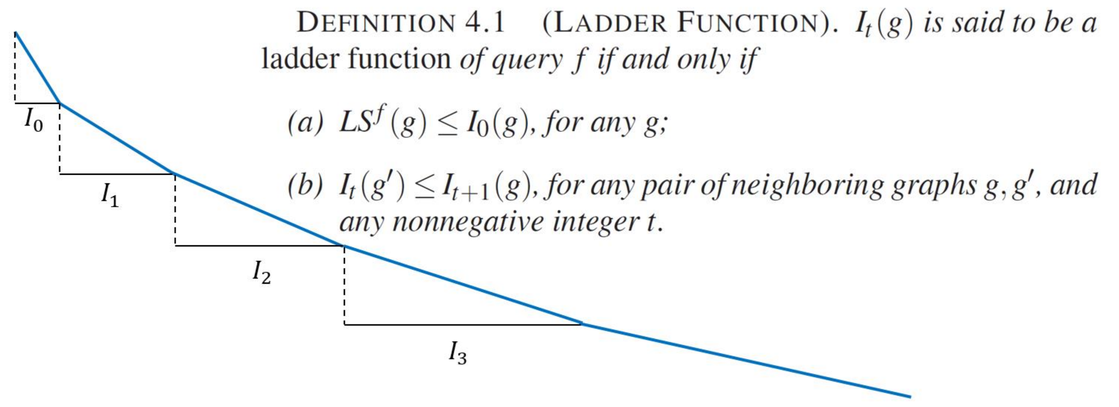
❑ Key Idea: Change Slope Gradually



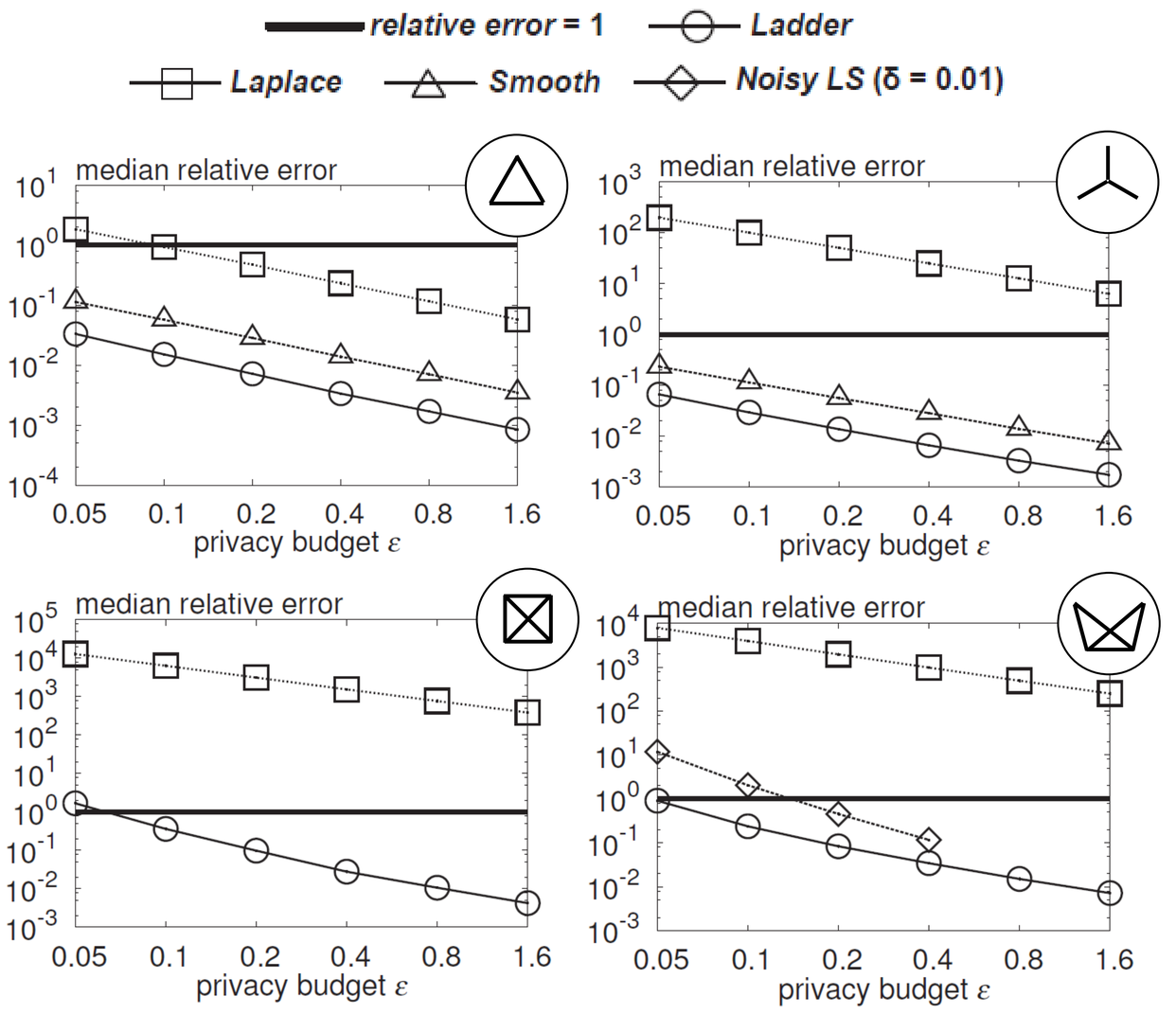
❑ Summary



❑ Formal Results



4. Experiments



NANYANG TECHNOLOGICAL UNIVERSITY

THE UNIVERSITY OF WARWICK

Google

at&t