

Réunion ALPAGE, Bordeaux, 31 Jan - 01 Fév 2008

# Checking the Validity of the Augmented Graph Model

Pierre Fraigniaud

CNRS

University Paris Diderot



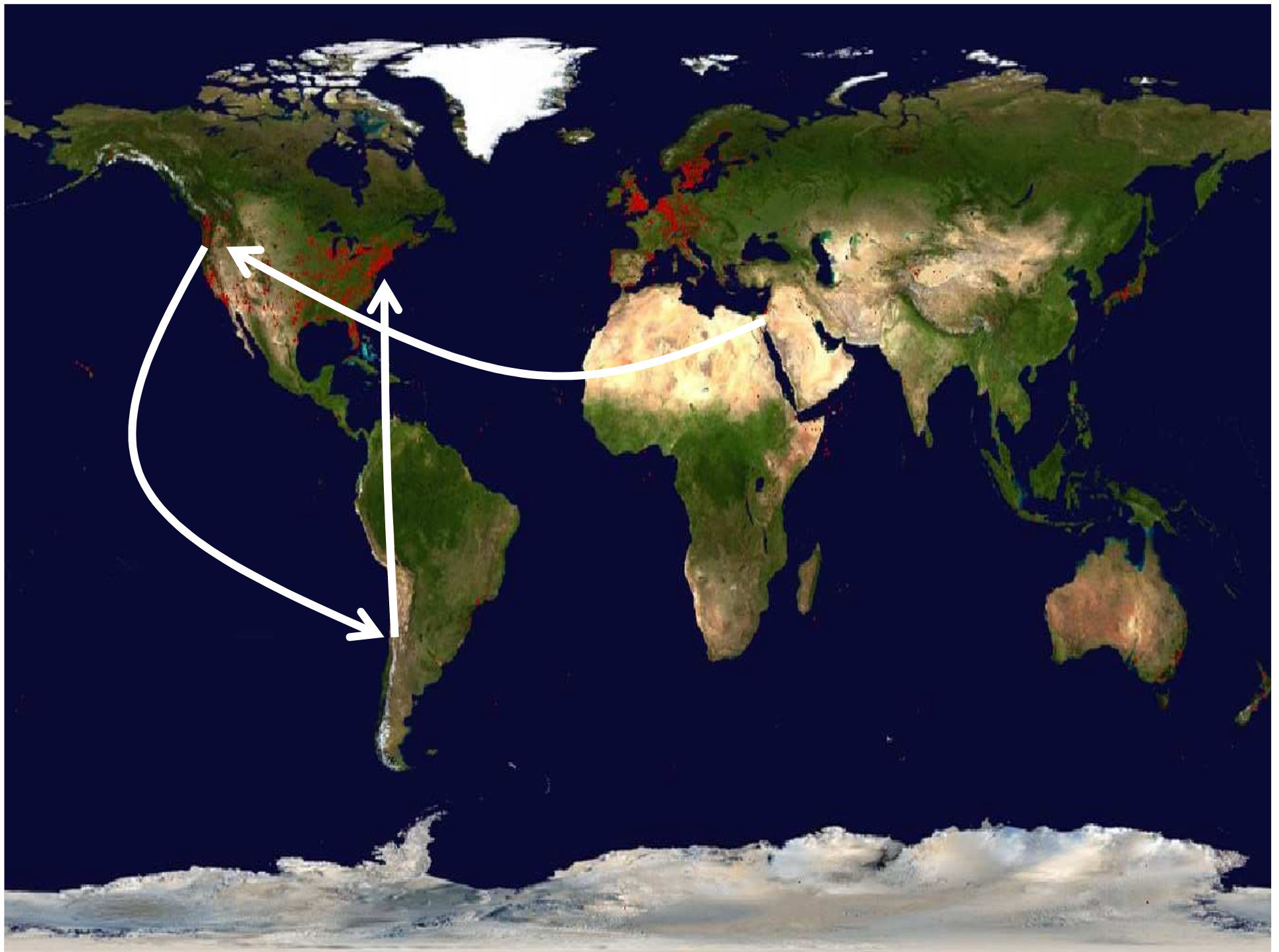
# Milgram Experiment (60's)



Name: Paul Johnson

Professional occupation: Broker

Place of living: Boston, Massachusetts



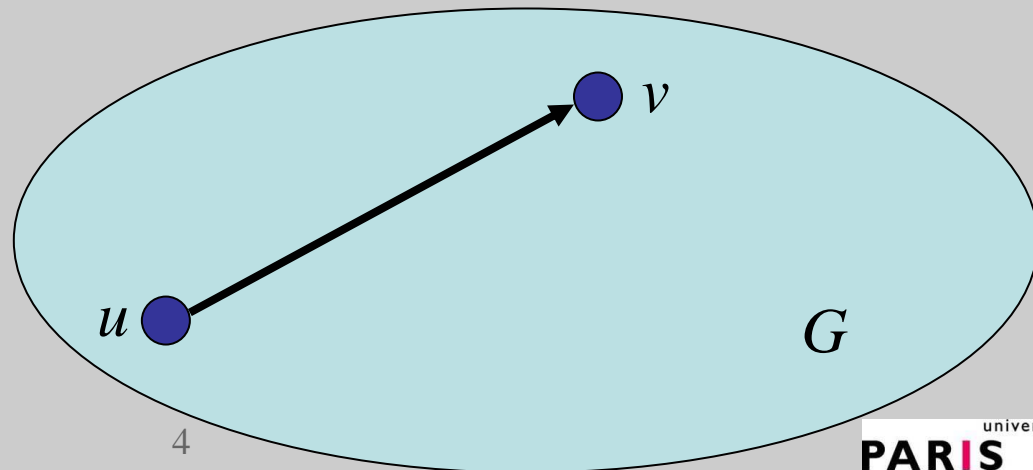
# Model $(G, \varphi)$

- A graph  $G$ , called “base graph”
- A collection of probability distributions:

$$\varphi = \{\varphi_u : u \in V(G)\}$$

called “augmenting distribution”

$$\varphi_u(v) = \Pr(u \rightarrow v)$$



# Greedy diameter of $(G, \varphi)$

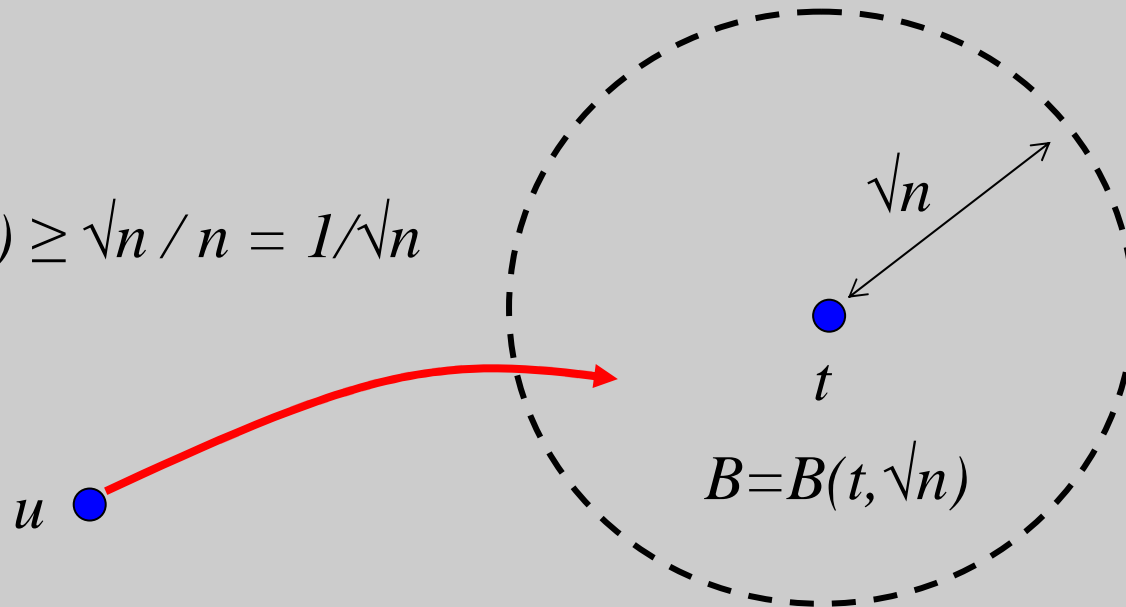
- Greedy routing:
  - Every intermediate node selects its neighbor (possibly its long range contact) that is closer to the target in  $G$ , and forwards to it.
- Greedy diameter:
  - $E_{\varphi}(s, t)$  = expected #steps of greedy routing from  $s$  to  $t$  in  $G$  augmented by  $\varphi$
  - $gd(G, \varphi) = \max_{(s, t) \in V(G) \times V(G)} E_{\varphi}(s, t)$
- Remark:  $gd(G, \varphi) \geq \text{diameter}(G, \varphi)$

# Upper bound

## Theorem [Peleg 2005]

The uniform augmenting scheme  $U$  satisfies  $gd(G, U) \leq O(\sqrt{n})$  for any  $G$

$$Pr(u \rightarrow B) \geq \sqrt{n} / n = 1/\sqrt{n}$$



# Lower bound

Theorem [F., Lebhar, Lotker, 2006]

There are graphs  $G$  such that, for any augmenting distribution  $\varphi$ ,

$$gd(G, \varphi) \geq \Omega(n^{1/\sqrt{\log(n)}})$$

# Ball growth

## Definition

A graph  $G$  has ball growth  $\leq \rho$  if for any node  $x$  and any radius  $r$ ,  $|B(x, 2r)| \leq \rho |B(x, r)|$

## Theorem

[Duchon, Hanusse, Lebhar, Schabanel, 2005, 2006]

For any graph  $G$  of ball growth  $\leq \rho$  there exists  $\varphi$  such that  $gd(G, \varphi) \leq \text{polylog } n$

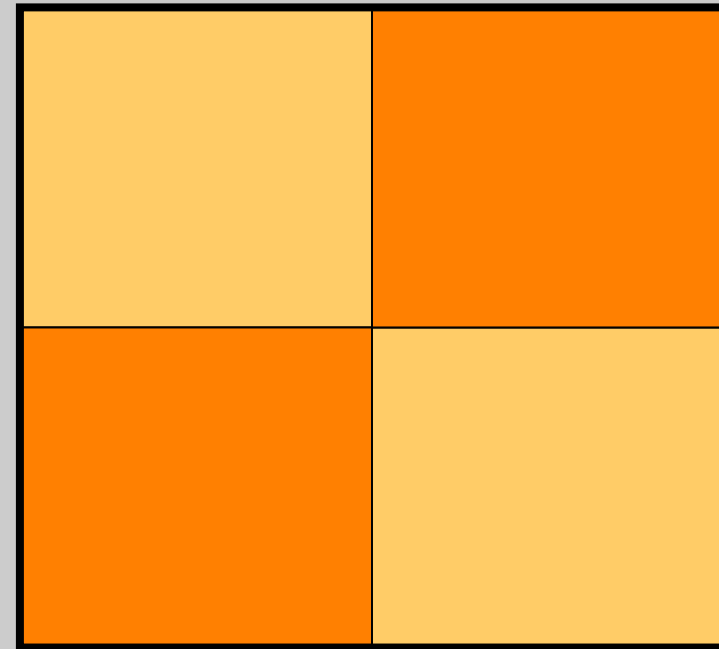
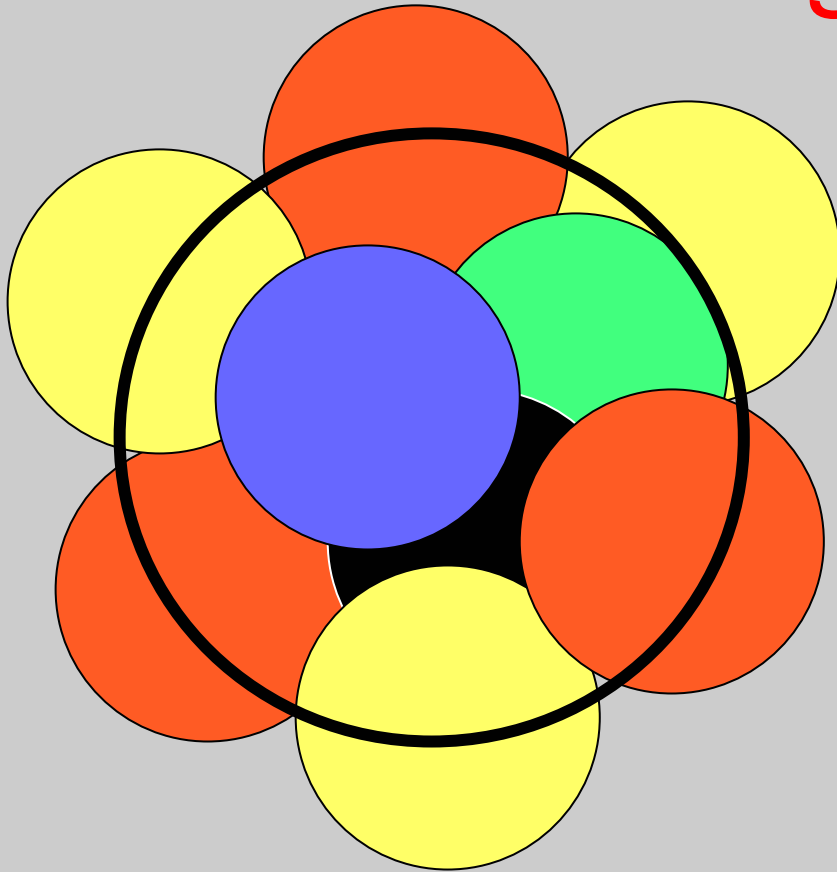
## Proof.

*Density-based* distribution:

If  $\text{dist}_G(u, v) = r$  then  $\varphi_u(v) \approx 1/|B(u, r)|$



# Doubling dimension



# Doubling dimension

## Definition

A graph  $G$  has doubling dimension  $\leq d$  if every  $B(x, 2r)$  can be covered by  $\leq 2^d$  balls of radius  $r$

## Theorem [Slivkins, 2005]

For any graph  $G$  of doubling dimension  $O(\log \log n)$  there exists  $\varphi$  such that  $gd(G, \varphi) \leq \text{polylog } n$

# Extracting the long range links

Let  $H \in (G, \varphi)$ , i.e.,  $E(H) = E(G) \cup F$

Objective: extracting  $F$  from  $H$ .

*Maximum Likelihood Method (MLM)*

Compute the edge-set  $S \subseteq E(H)$  of cardinality  $n$  that maximizes

$Pr(H \mid S \text{ is the set of long range links})$

**Difficulties**

1. Exponential number of sets  $S$
2. Requires the knowledge of  $\varphi$

# Local Maximum Likelihood Method (LMLM)

- For  $H \in (G, \varphi)$ , with  $E(H) = E(G) \cup F$ , LMLM decides whether  $e \in F$  depending on
$$Pr(H \mid e \in F)$$
  - Augmenting distribution  $\varphi$ :
    - [Chung, Lu, 2004] [Andersen, Chung, Lu, 2006]  
**Power law distribution**
    - [F., Lebar, Lotker, 2007]  
**Density-based distribution**
- If  $dist_G(u, v) = r$  then  $\varphi_u(v) \approx 1/|B(u, r)|$

Difficulty: locality!

# Clustering hypothesis on $G$

- For every  $e = (u, v) \in E(G)$ , there exist at least  $k$  disjoint paths of length at most  $L$  between  $u$  and  $v$  [Chung, Lu, 2004]
- For every  $e = (u, v) \in E(G)$ , there at least  $k$  units of flow can be pushed from  $u$  to  $v$  along paths of length at most  $L$  [Andersen, Chung, Lu, 2006]
- Every  $e \in E(G)$  belongs to at least  $\Omega(\log(n)/\log\log(n))$  triangles [F., Lebhar, Lotker, 2007]

# Extraction algorithm

Given  $H \in (G, \varphi)$ , with  $E(H) = E(G) \cup F$

- Density based augmenting distribution  $\varphi$
- Base graph  $G$  with
  - bounded ball growth or bounded doubling dimension
  - clustering  $\Omega(\log(n)/\log\log(n))$

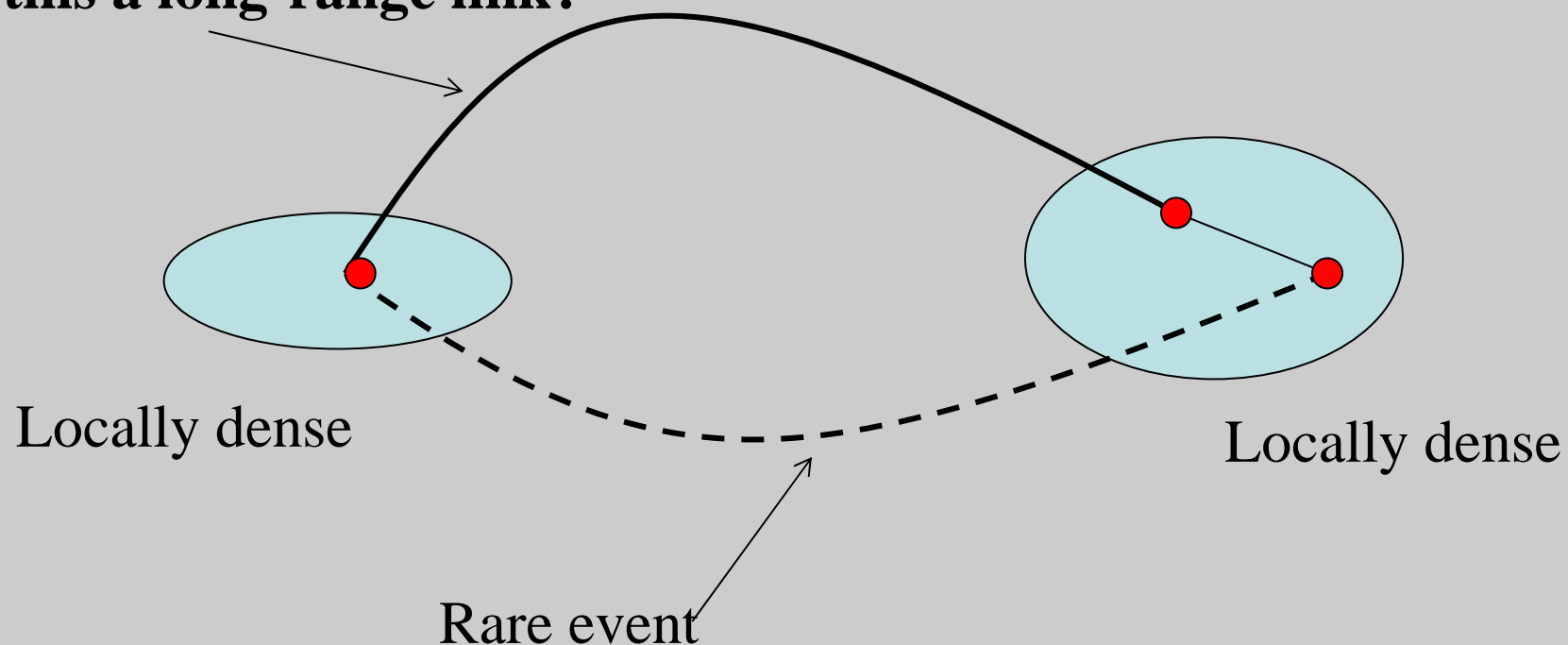
Theorem [F., Lebhar, Lotker, 2007]

There is a LMLM algorithm that returns  $F' \subseteq F$  where:

- W.h.p.,  $H - F'$  contains at most  $\text{polylog } n$  edges of stretch larger than  $\text{polylog } n$ , and
- Greedy routing using the map  $H - F'$  performs in  $\text{polylog } n$  expected number of steps

# Idea of the proof

Is this a long-range link?



# Clustering Coefficient Revisited

- Theorem [F., Lebhar, Lotker, 2007]

For any  $0 < \varepsilon < 1/5$ , any MLM algorithm for recovering the base graph  $C_{2n+1}$  in  $H \in (C_{2n+1}, h)$  fails in the detection of an expected number  $\Omega(n^{5\varepsilon} / \log n)$  of long range links of stretch  $\Omega(n^{1/5 - \varepsilon})$ .



A satellite image of Earth from space, showing the Middle East, Africa, and parts of Europe and Asia. The image is set against a black background. A blue speech bubble with a black outline is positioned in the upper right quadrant, containing the text "Thank you!".

***Thank you!***