Compact Routing on Internet-Like Graphs

Dmitri Krioukov, UCSD/CAIDA
<dima@krioukov.net>

Kevin Fall, Intel Research
<kfall@intel-research.net>

Xiaowei Yang, MIT
<yxw@mit.edu>

The motivation

- BGP routing table size scalability concerns: immediate causes are well studied (multihoming, more peering, traffic engineering, address allocation, etc.), but…
- Typical "solution": aggregate using (multiple) levels of *hierarchical network partitioning* in the Kleinrock-Kamoun style, but…
  - This scheme does not work for densely connected networks, so:

The research problem

- **What are the *fundamental* scalability limits (in terms of routing table size) for routing?**
- Recall, *stretch* is defined as:
  - S = max[(hop count using routing scheme) / (shortest path in graph)]
- Fact: modern "compact routing" schemes can guarantee small routing table sizes. The price is increased maximum stretch, but…
- Low (~1) stretch may well be a requirement for Internet routing, so: **What is the average stretch produced by these schemes on Internet-like [*scale-free*] topologies?**

Stretch/local memory Results

- Trivial shortest path $\Rightarrow$ $O(n \log n)$
- $1 \leq s < 1.4 \Rightarrow \Omega(n \log n)$ *(Gavoille/Perennes 96)*
- $1.4 \leq s < \mathbf{3} \Rightarrow \Omega(n)$ *(Gavoille/Genegler 01)*
- $\mathbf{3} \leq s < 5 \Rightarrow O(n^{1/2} \log n)$ *(Eilam/Gavoille/Peleg 98)*
- Thorup & Zwick (TZ01): $s=\mathbf{3}$, $\Omega(n^{1/2} \log^{1/2} n)$
  - Improves over $s=\mathbf{3}$, $O(n^{2/3} \log^{4/3} n)$ *(Cowen 99)*
  - Nearly "optimal," up to logarithmic factor
  - The basis for our studies
  - (uses custom node labels; not a dynamic scheme)

Our Approach

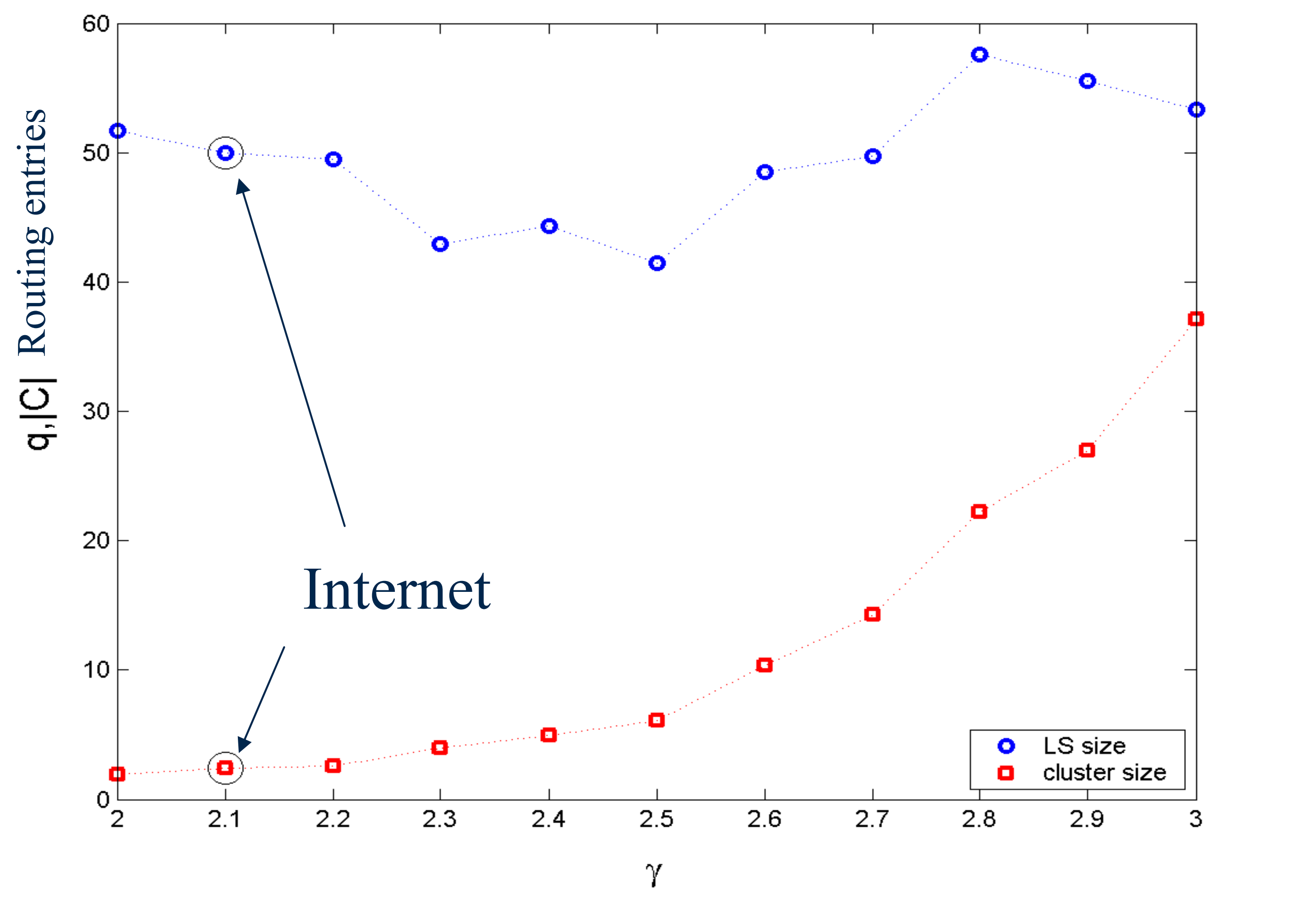
- Analysis: evaluate average TZ stretch as a function of the first two moments of an (assumed Gaussian) distance distribution in a graph
- Simulations: develop a TZ simulator and use it on uncorrelated random power-law graphs (generated by PLRG) with node degree *k* ($P_k \sim k^{-\gamma}$), and on classical random graphs

Our Results

Findings (analysis + simulations agree)

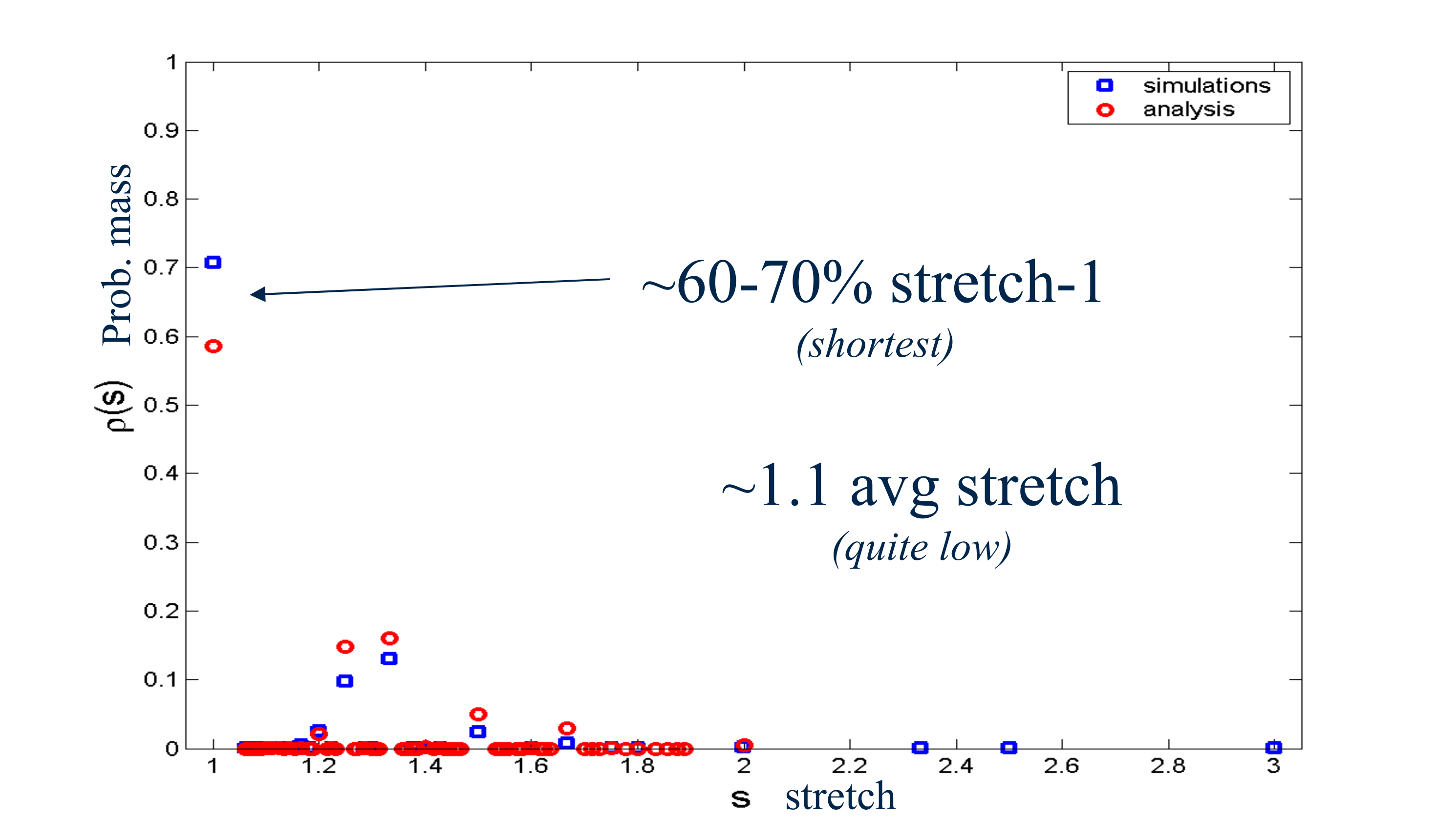
- Routing table sizes are *well below* the theoretical upper bounds (52 vs. 2187 for n=10,000)
- Average stretch:
  - is low (~1.1, ~70% paths are shortest [stretch-1])
  - does not depend on the power-law exponent γ
  - *decreases* with n (i.e. weak neg. correlation on n)
- Remarkably, the average stretch function has a *unique critical point* and the Internet is located in its close neighborhood

Routing table sizes are small

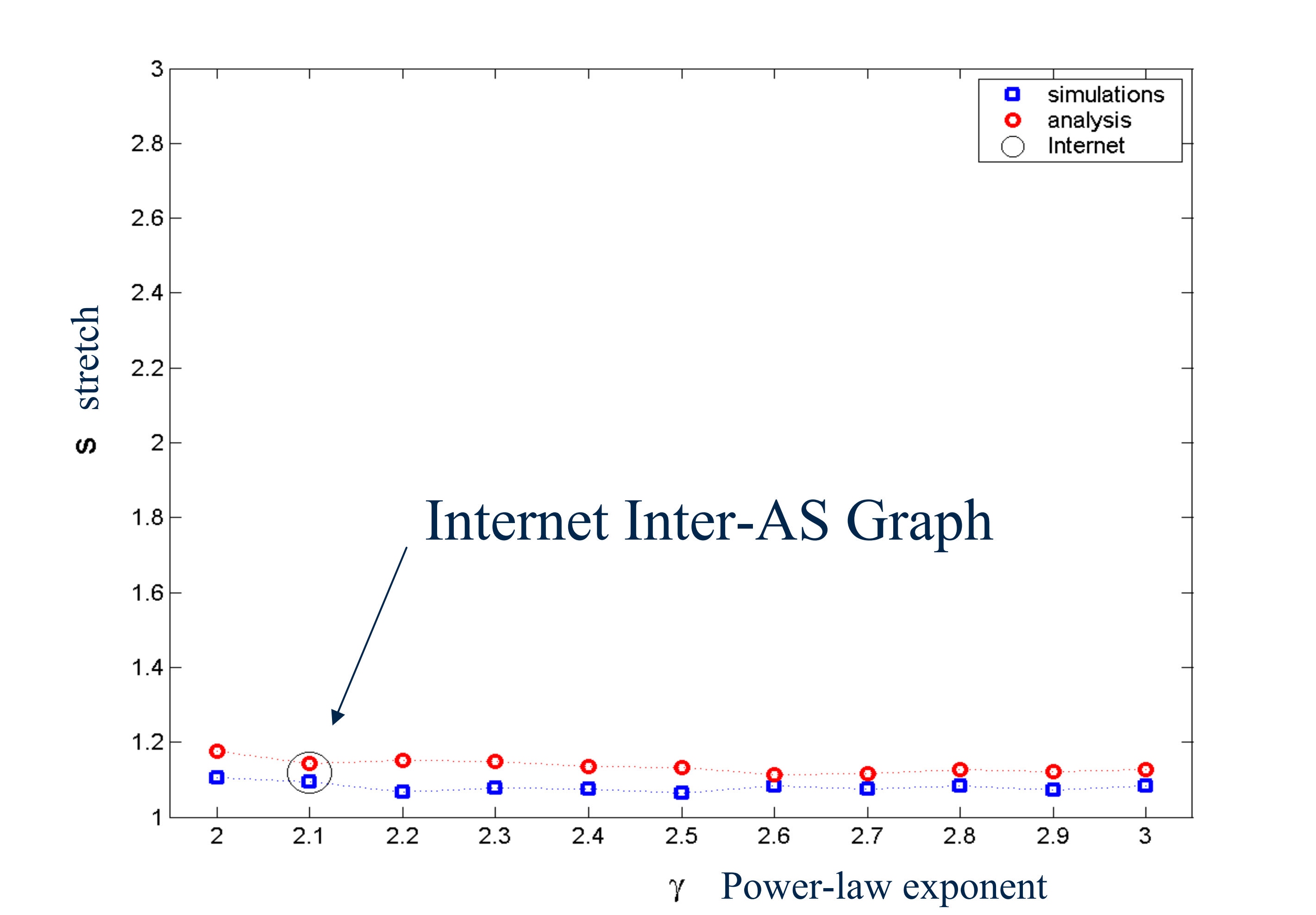


Stretch distribution / average

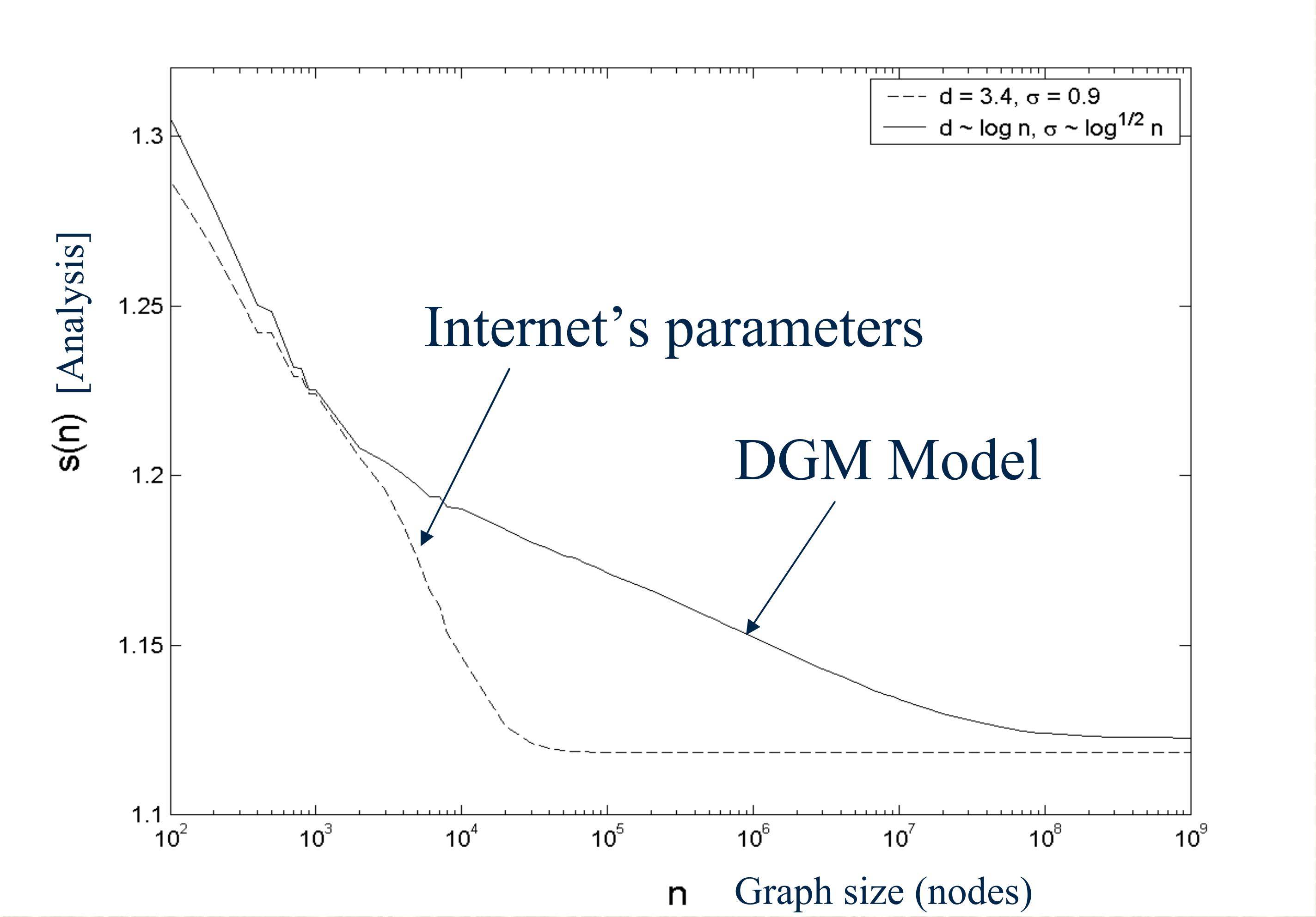
average s = 1.09 (simulations, 10k nodes) □
average s = 1.14 (analysis, DGM model) o



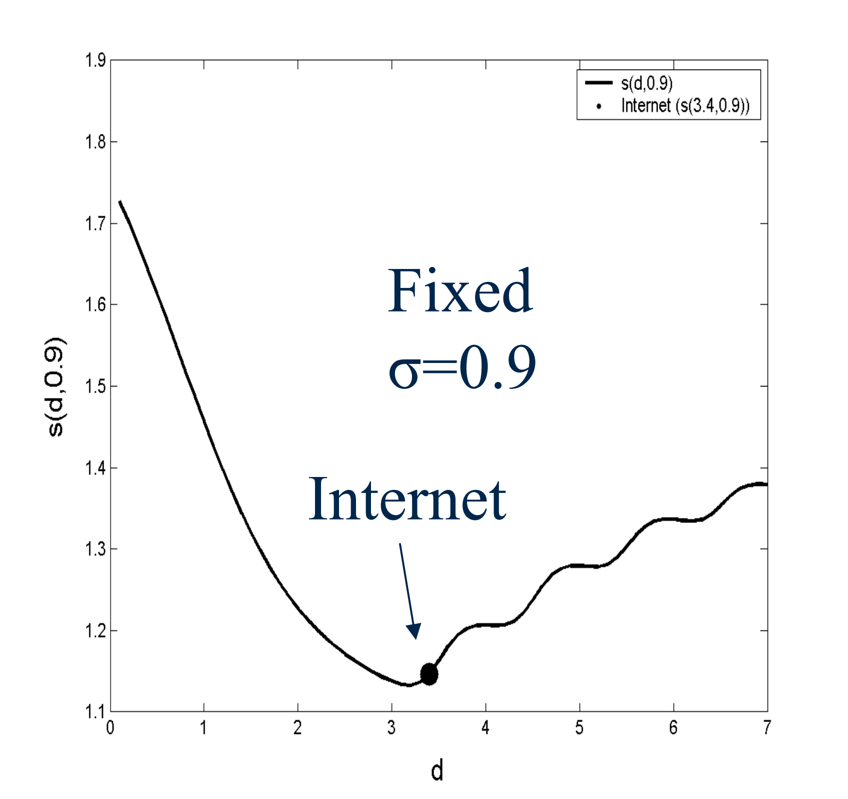
Independence of γ (n:10000)

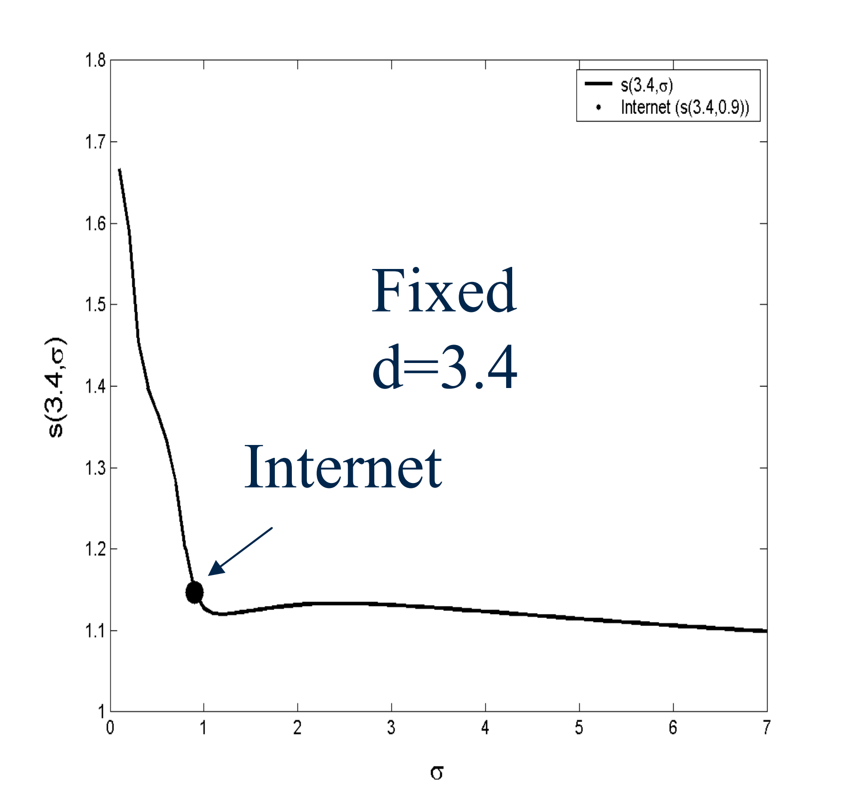


S Decreases with network size

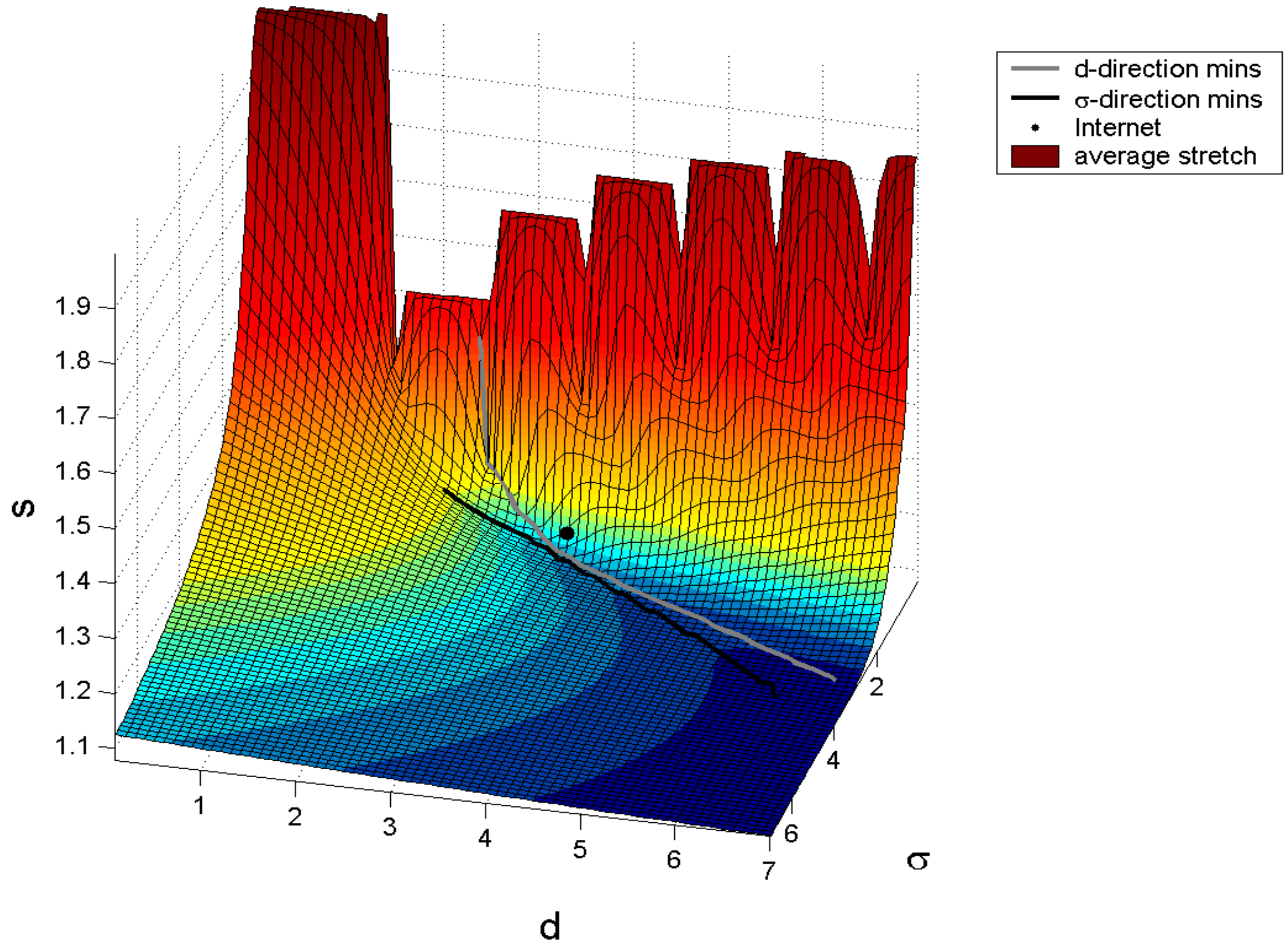


Avg Stretch Dependency on d,σ: Internet point ~ minimizes stretch

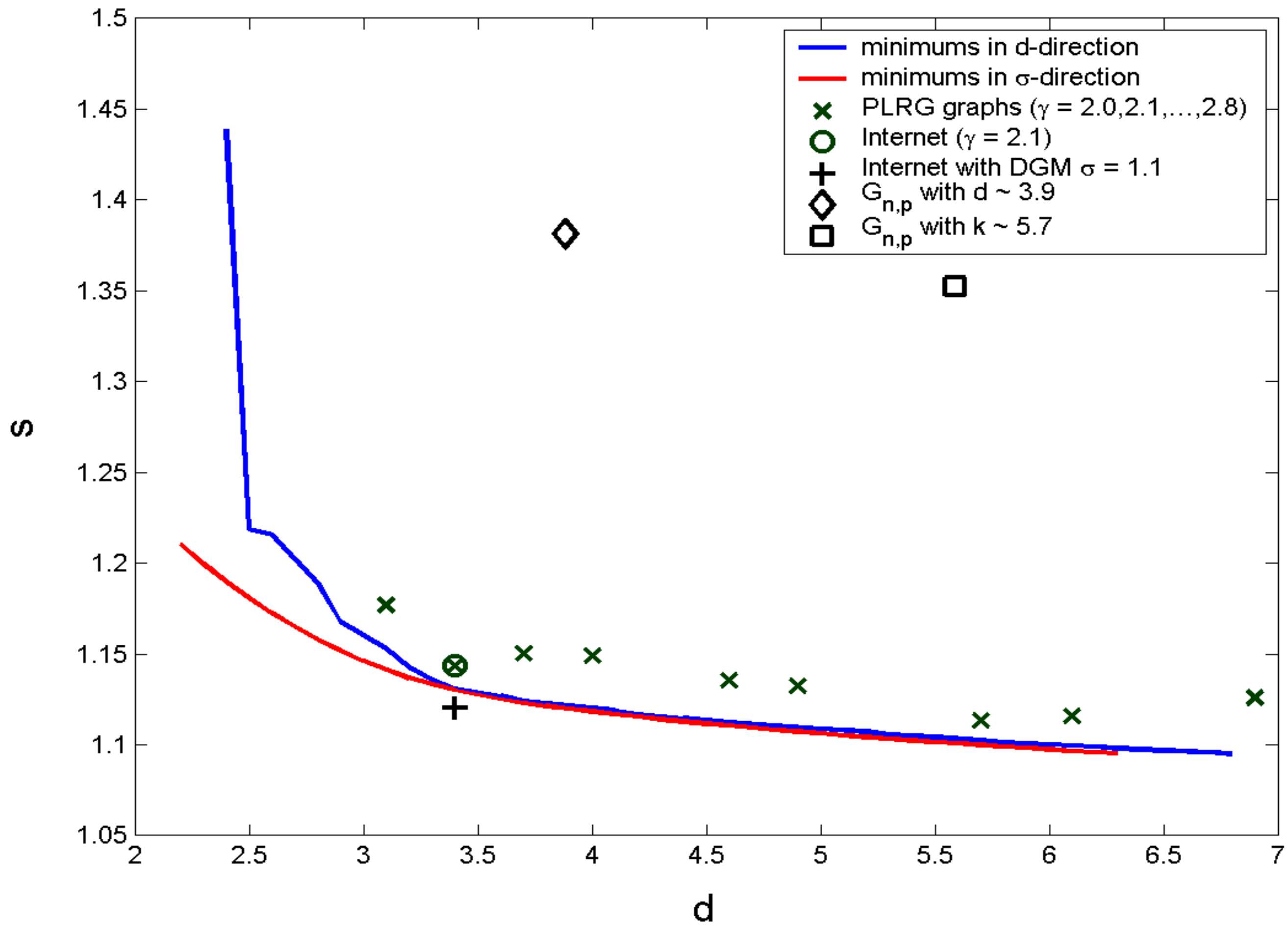




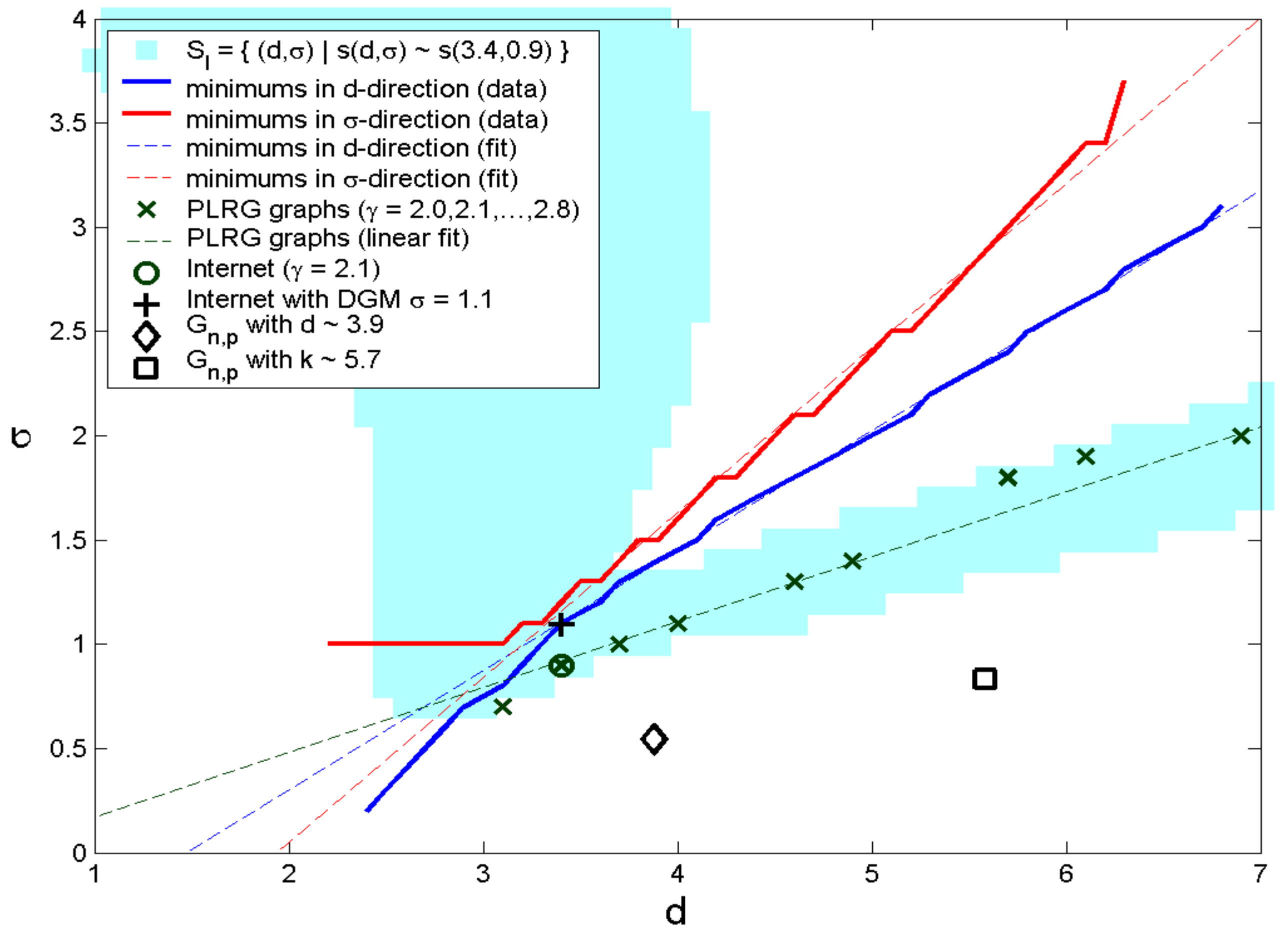
TZ Stretch function (3d)



Stretch function (d-s plane)



Stretch function (d-σ plane)



Summary

- Fundamental limits to routing table size scalability
  - Compact routing (TZ scheme):
    - + *bounded stretch, small table sizes*
    - + *appears to work very well on scale-free graphs*
    - *- not yet a dynamic routing scheme, not stretch-1*
- Internet point ~ quasi-stationary point
  - Stretch function may reveal drivers of Internet topology evolution
  - Need better understanding to know why

Thank you!