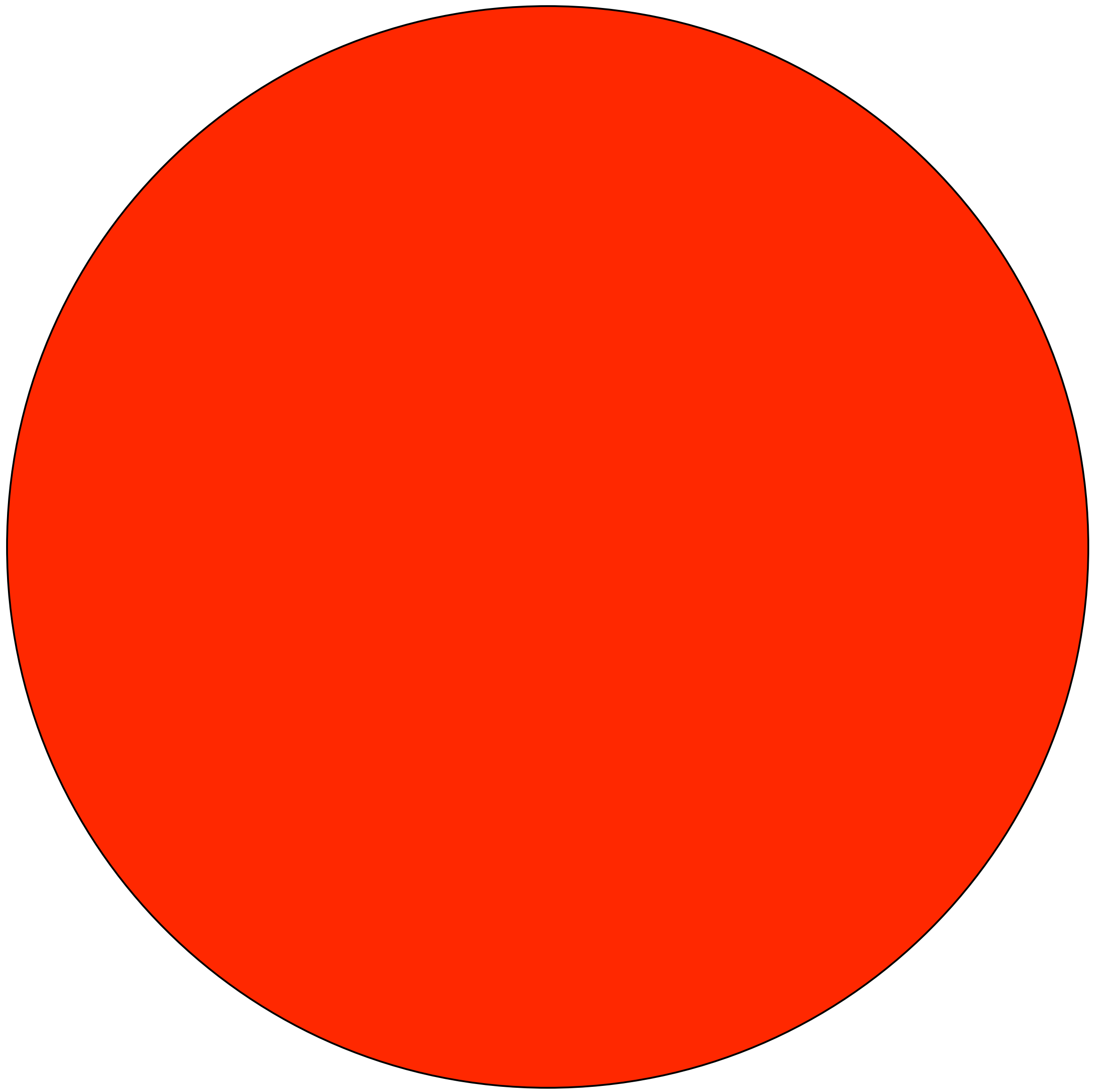
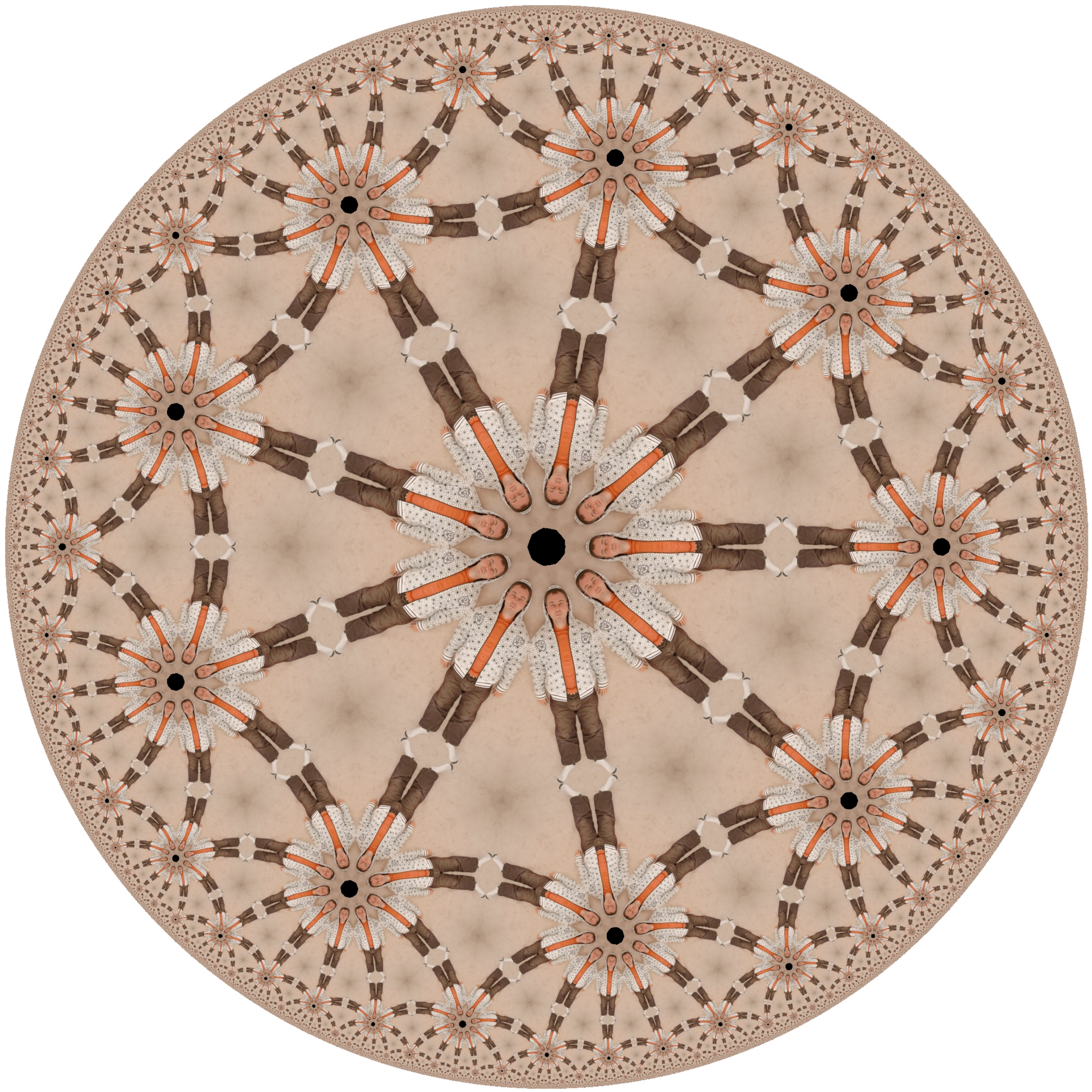
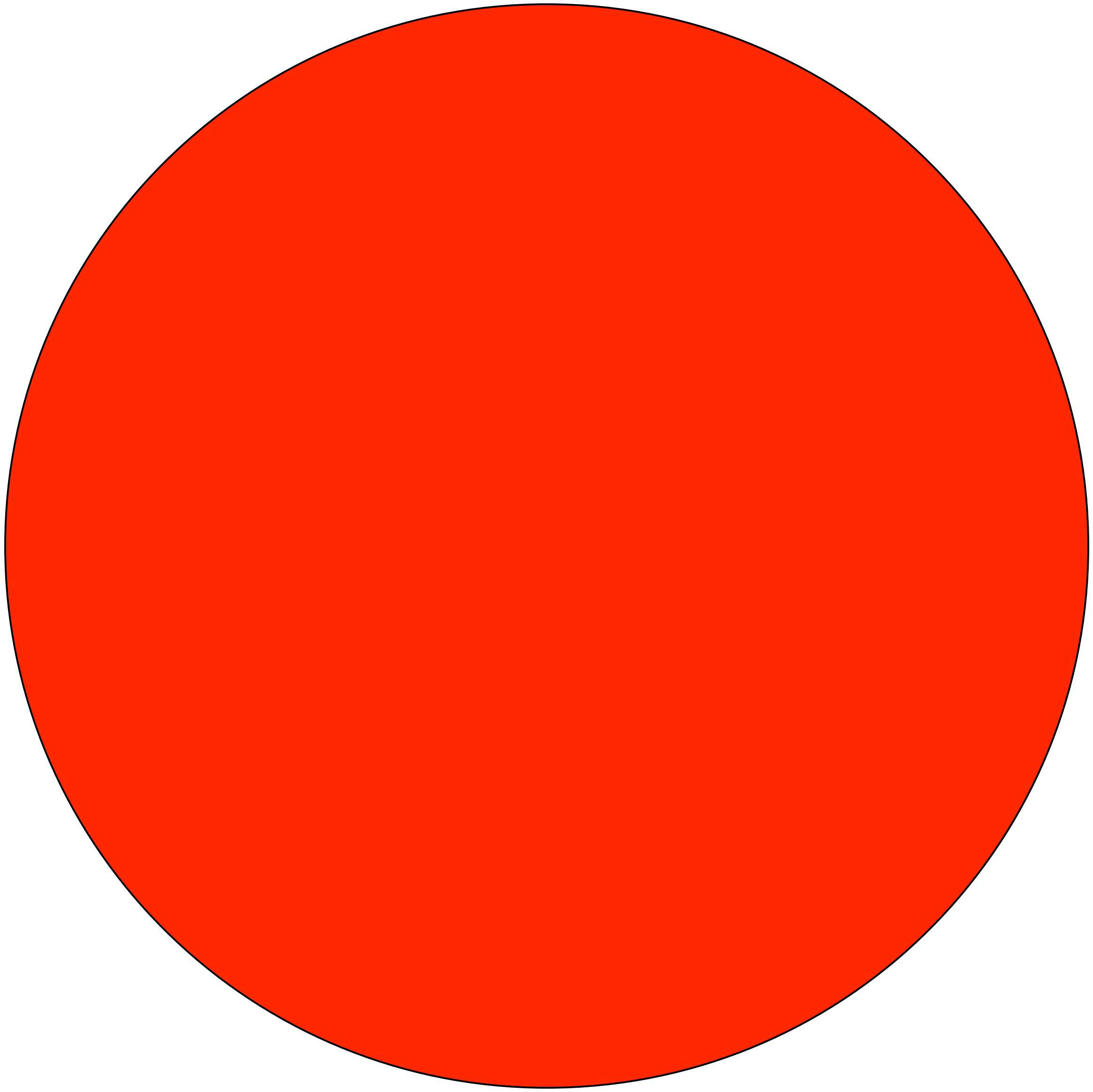
Hyperbolic mapping of complex networks

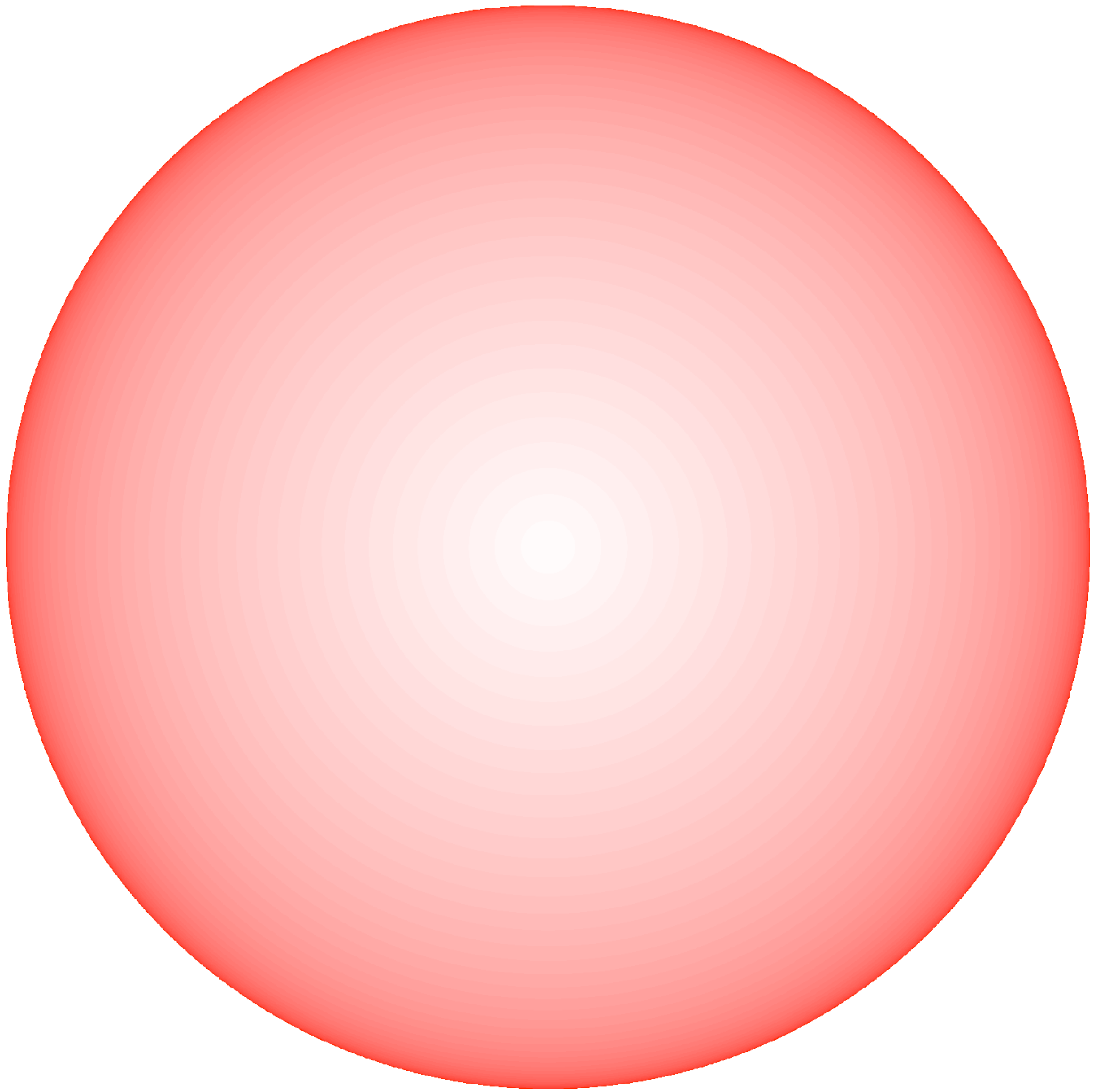
Dmitri Krioukov
CAIDA/UCSD
F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá

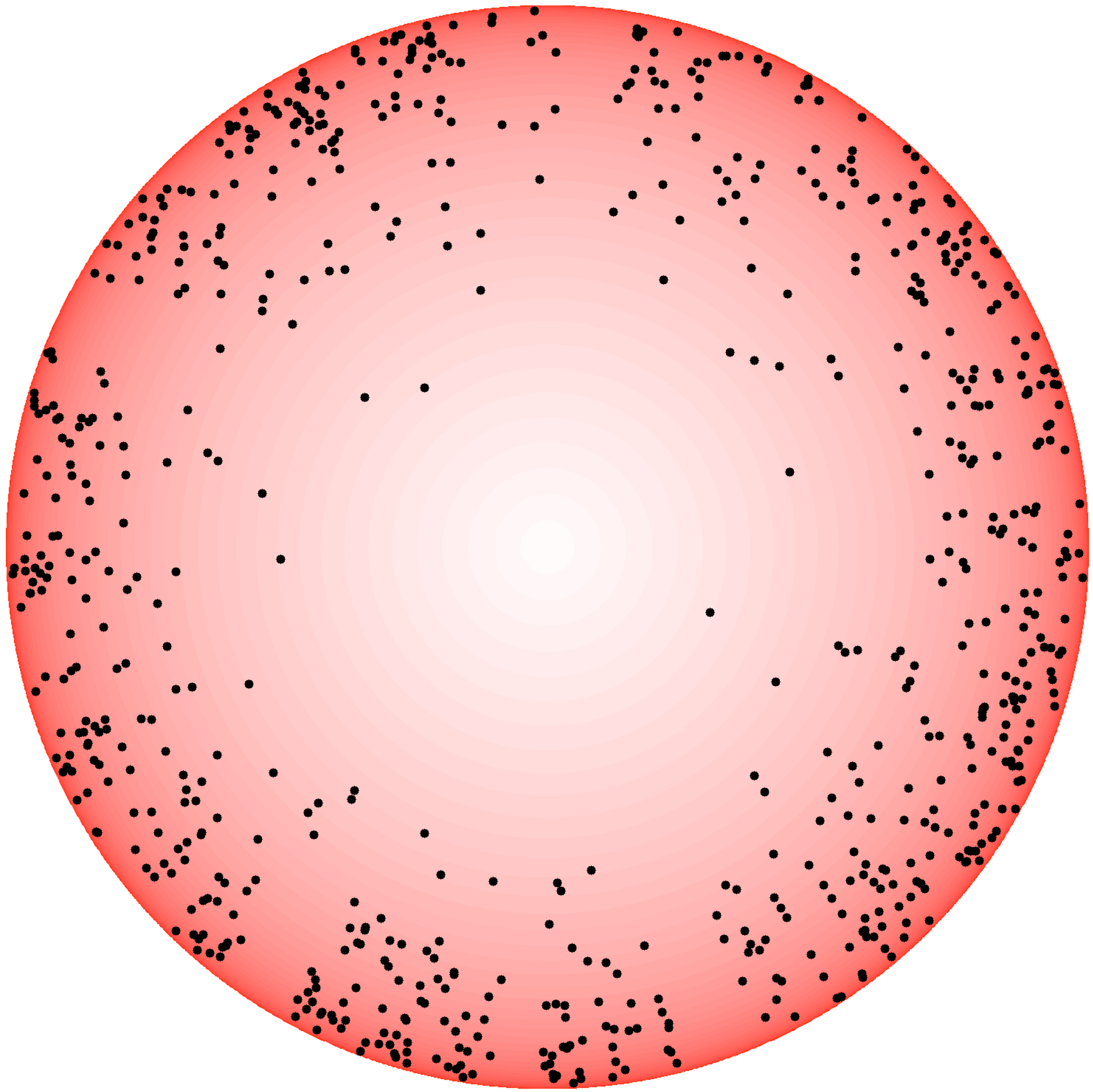
MMDS 2010
Stanford, June 2010

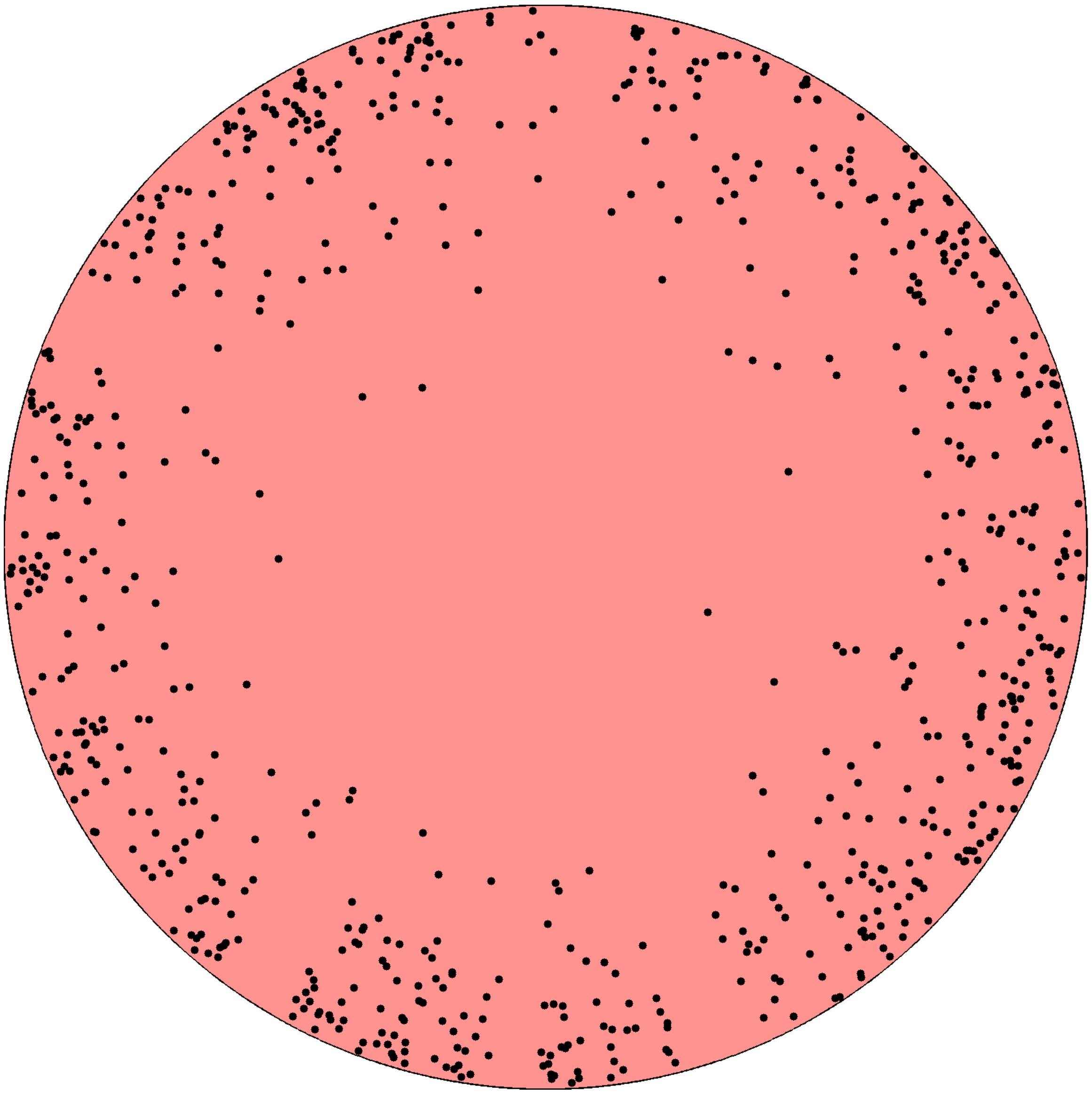


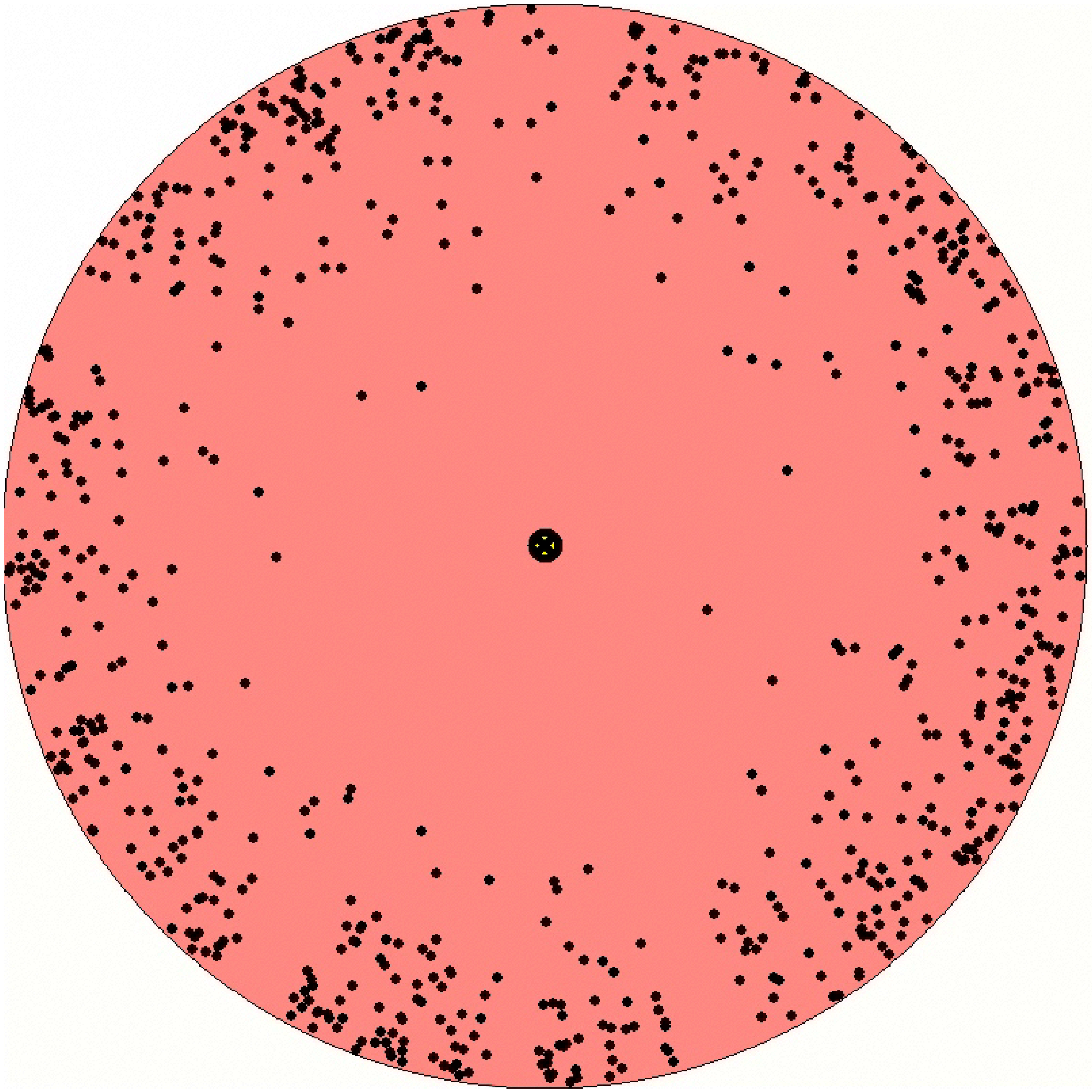


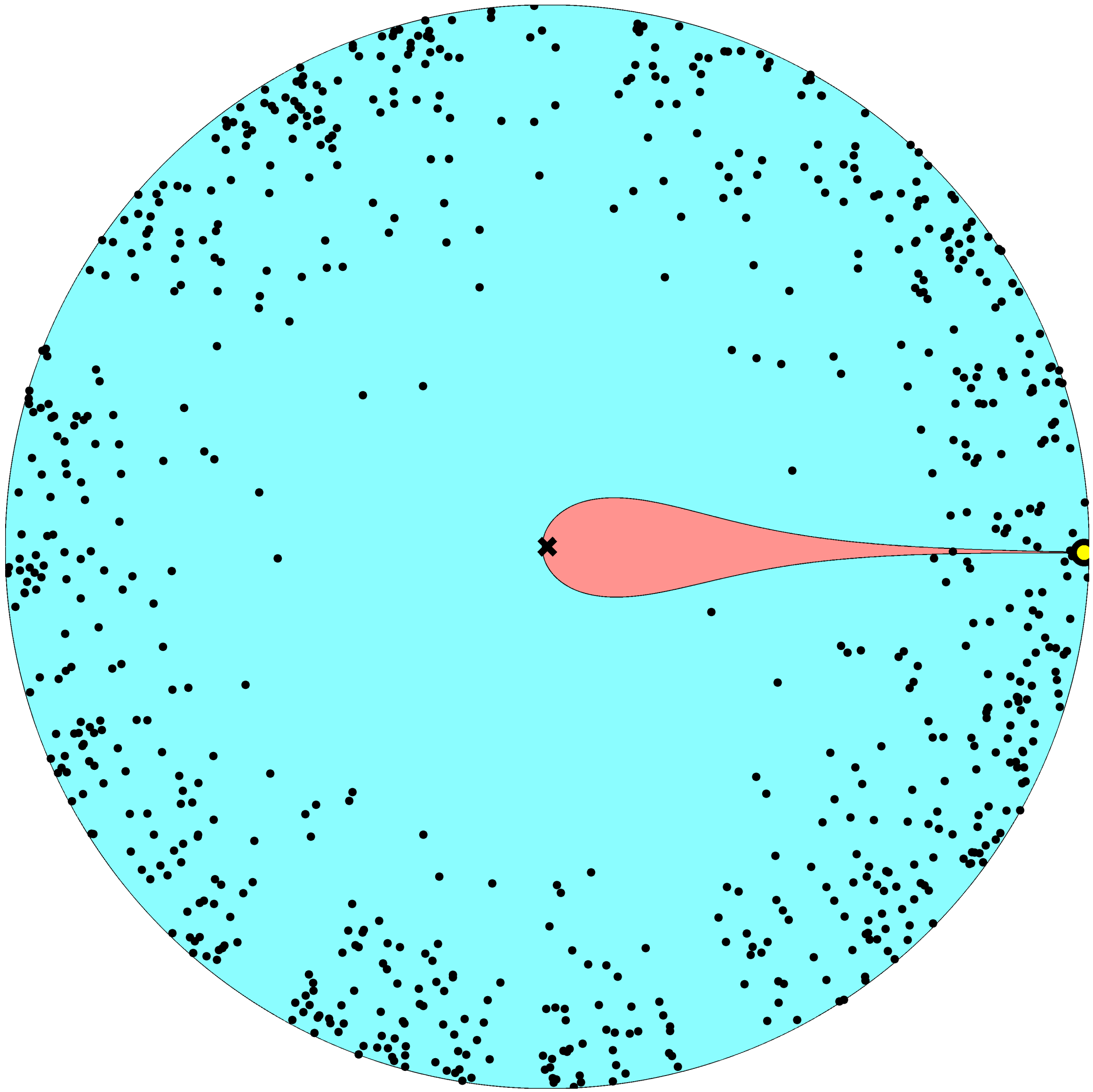


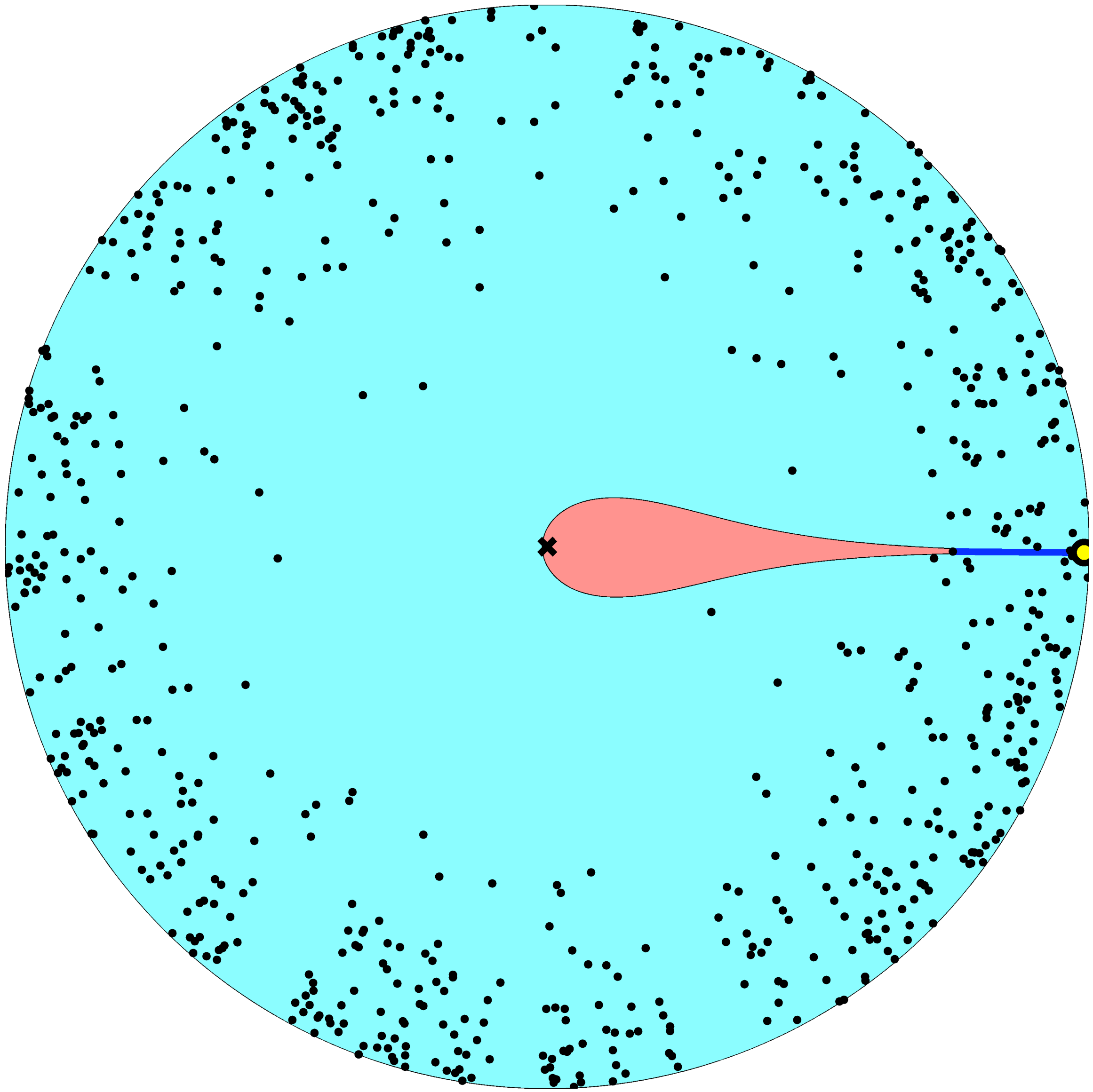


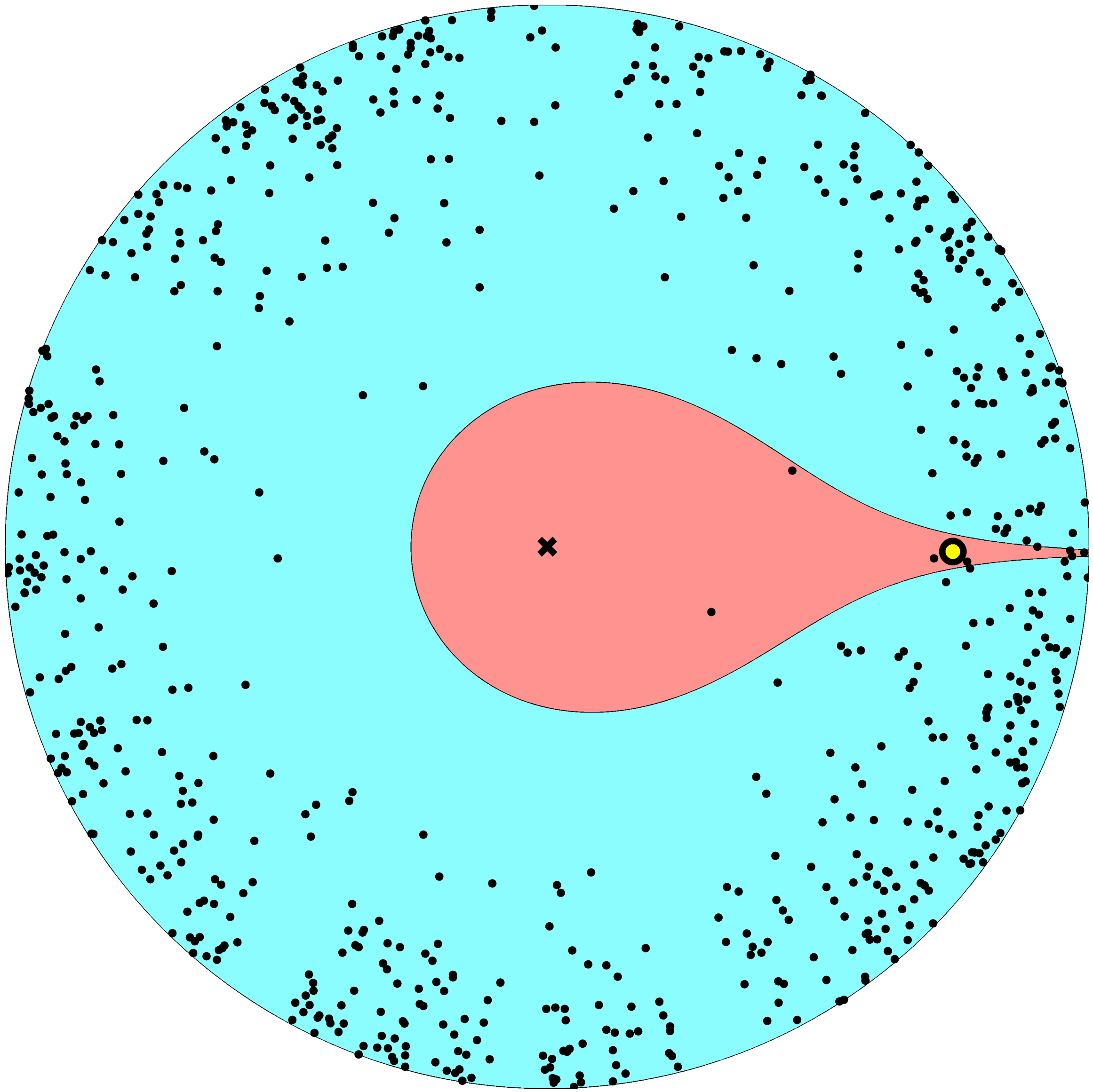


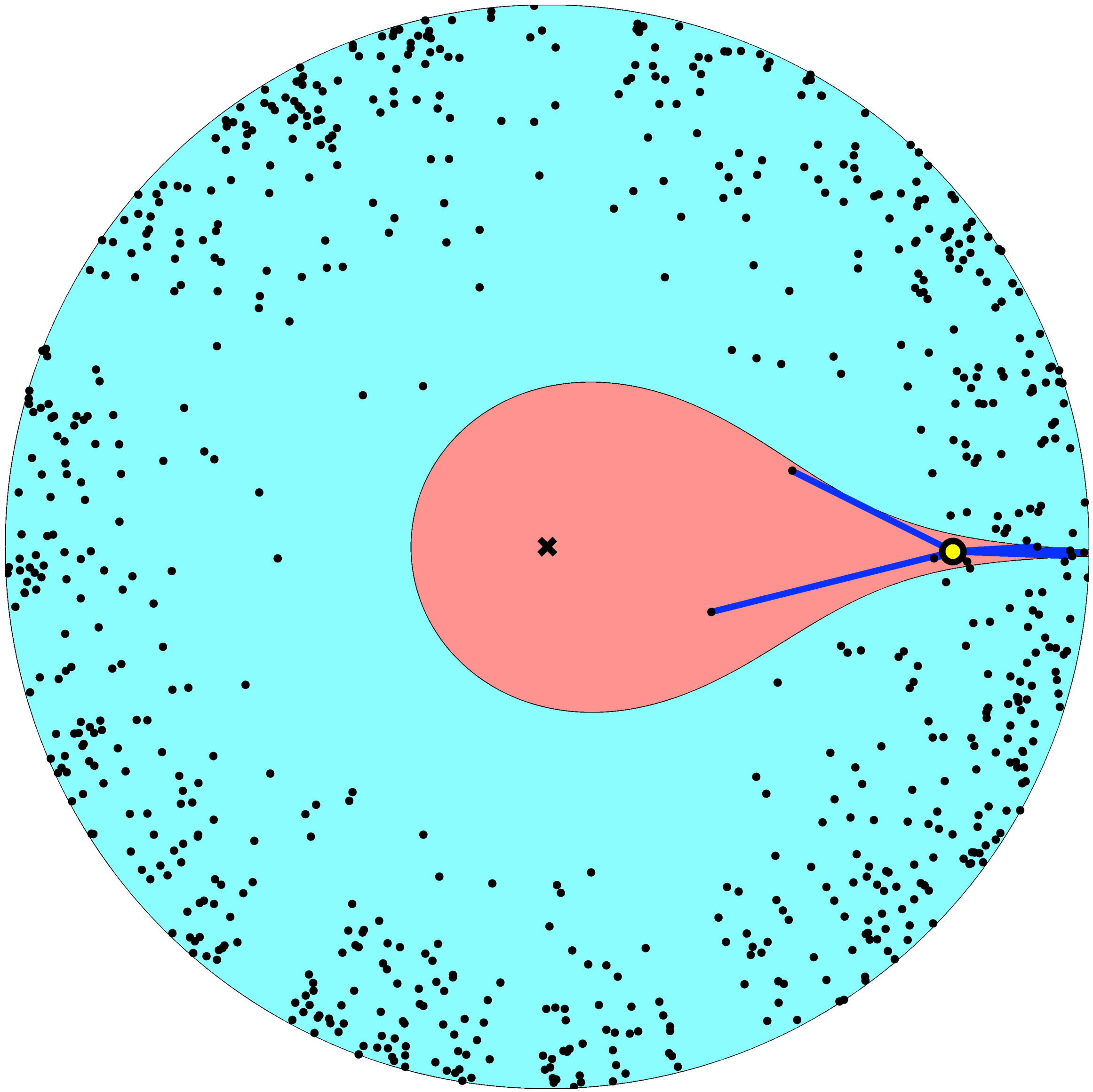


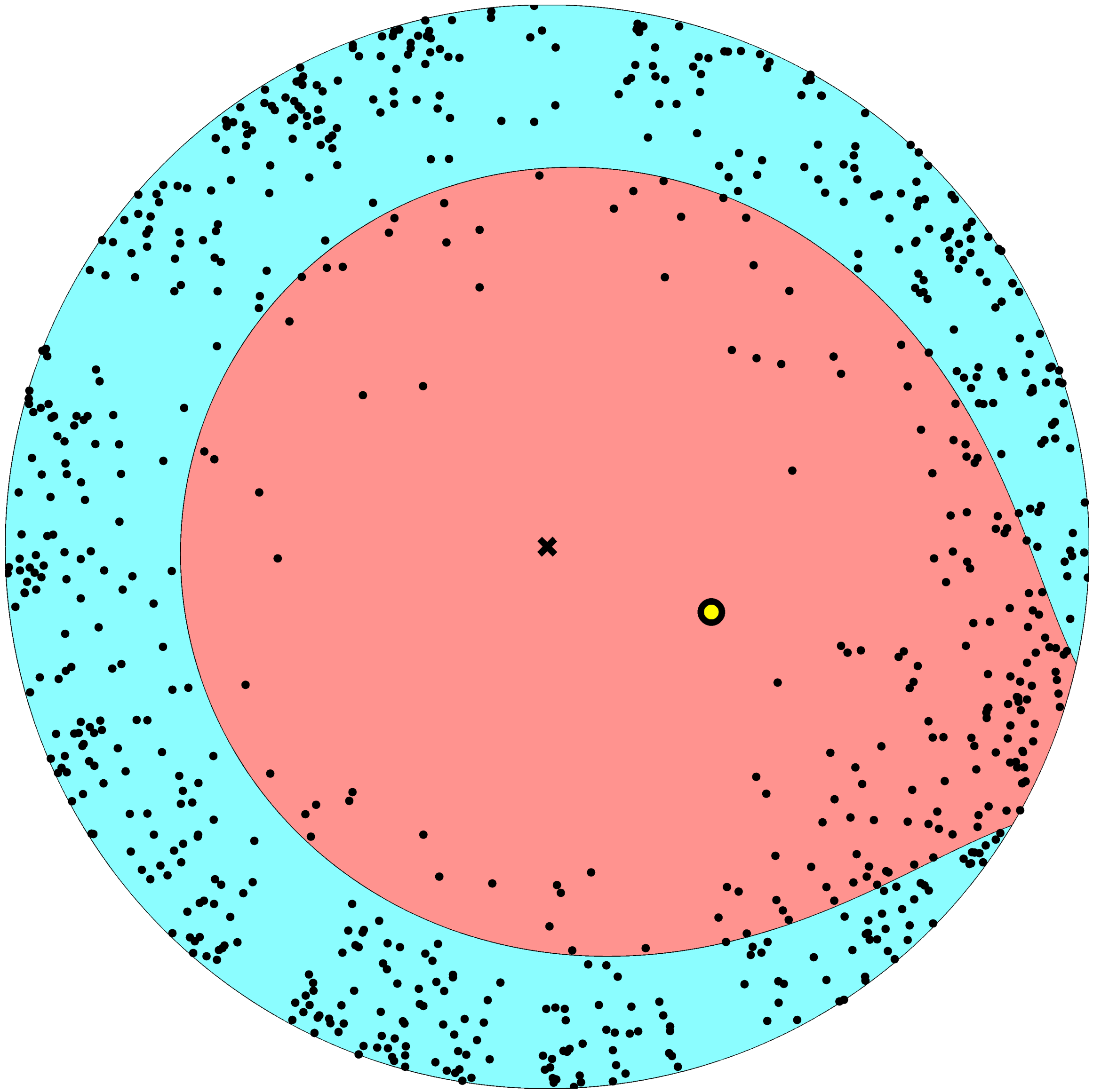


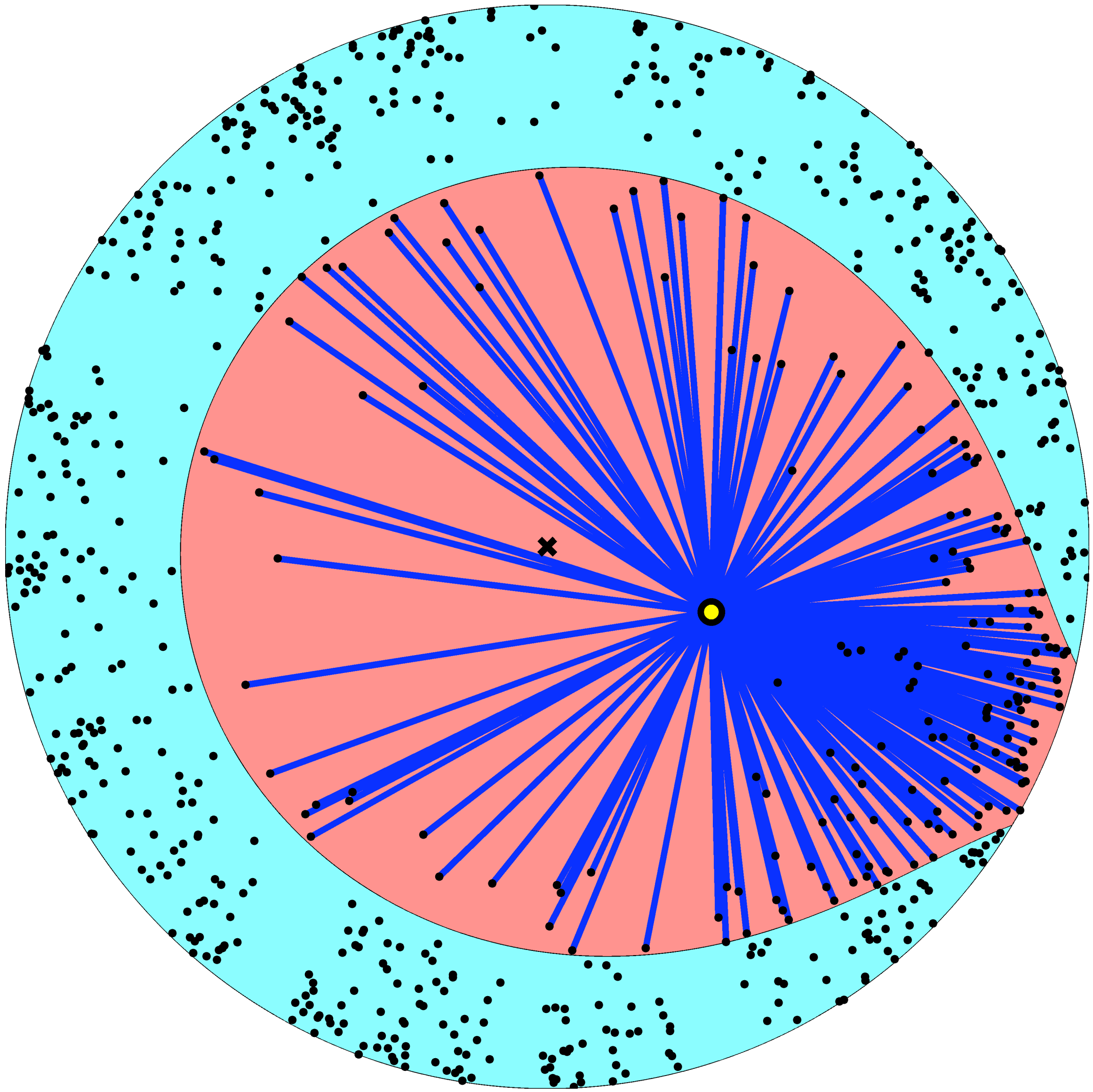












Node density

$$\rho(r) \sim e^{r-R}$$

R – disk radius

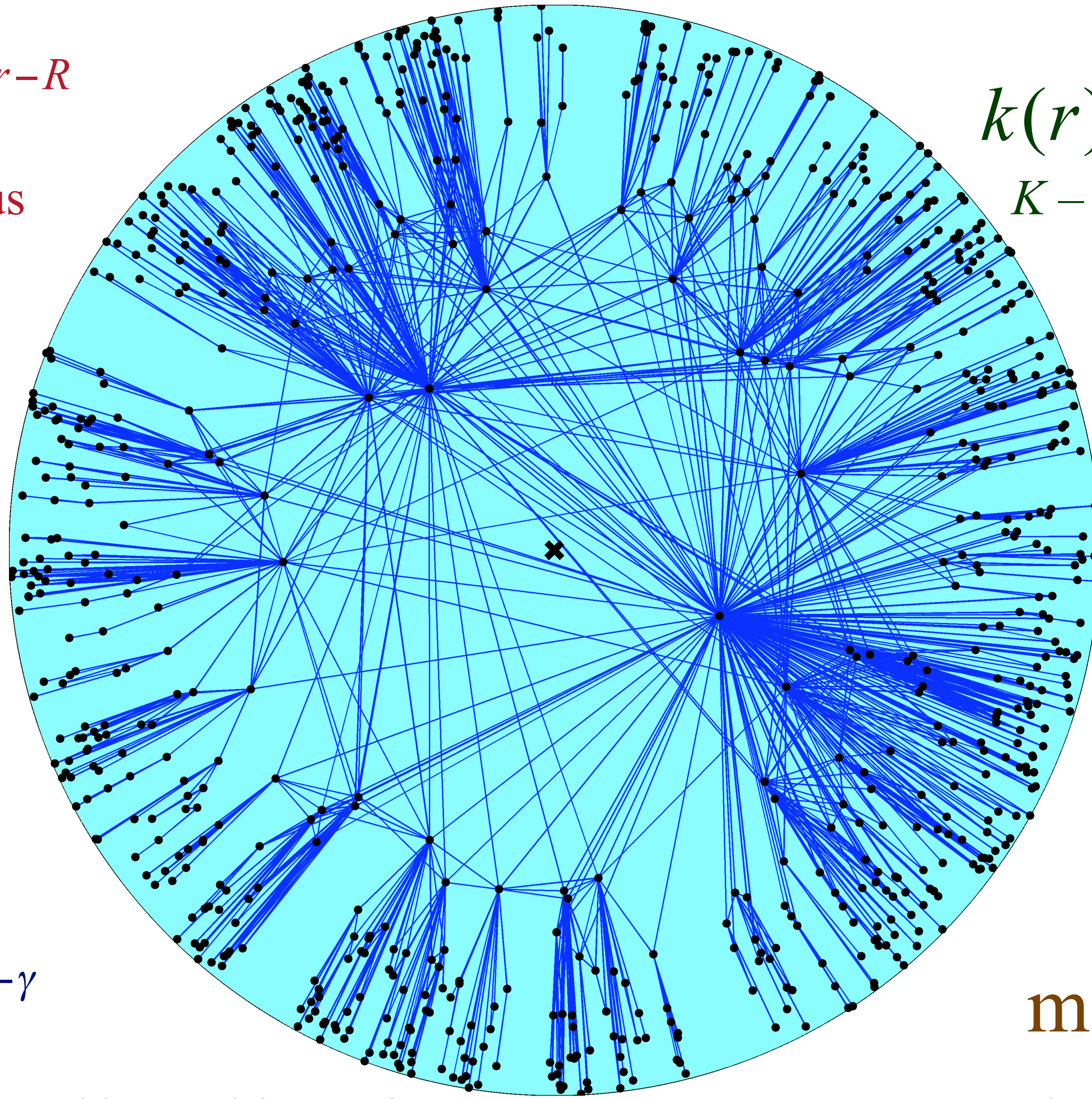
$r \in [0, R]$

Node degree

$$k(r) \sim e^{\frac{\zeta}{2}(R-r)}$$

K – disk curvature

$\zeta = \sqrt{-K}$



$$\gamma = \frac{2}{\zeta} + 1$$

$$P(k) \sim k^{-\gamma}$$

Degree distribution

$$c(k) \sim k^{-1}$$

maximized

Clustering

Connection probability as the Fermi-Dirac distribution

$$p(x) = \frac{1}{e^{\frac{\zeta}{2}\left(\frac{x-R}{T}\right)} + 1} \xrightarrow{T \to 0} \Theta(R - x)$$

- connection probability $p(x)$ – Fermi-Dirac distribution
- hyperbolic distance x – energy of links/fermions
- disk radius R – chemical potential
- two times inverse sqrt of curvature $2/\zeta$ – Boltzmann constant
- parameter T – temperature

Chemical potential *R* is a solution of

$$M = \binom{N}{2} \int g(x)\, p(x)\, dx$$

- number of links M – number of particles
- number of node pairs $N(N-1)/2$ – number of energy states
- distance distribution $g(x)$ – degeneracy of state x
- connection probability $p(x)$ – Fermi-Dirac distribution

Cold regime $0 \leq T < 1$

- Chemical potential $R=(2/\zeta)\ln(N/\nu)$
  - Constant ν controls the average node degree
- Clustering decreases from its maximum at $T=0$ to zero at $T=1$
- Power law exponent γ does not depend on T, $\gamma=(2/\zeta)+1$

Phase transition $T=1$

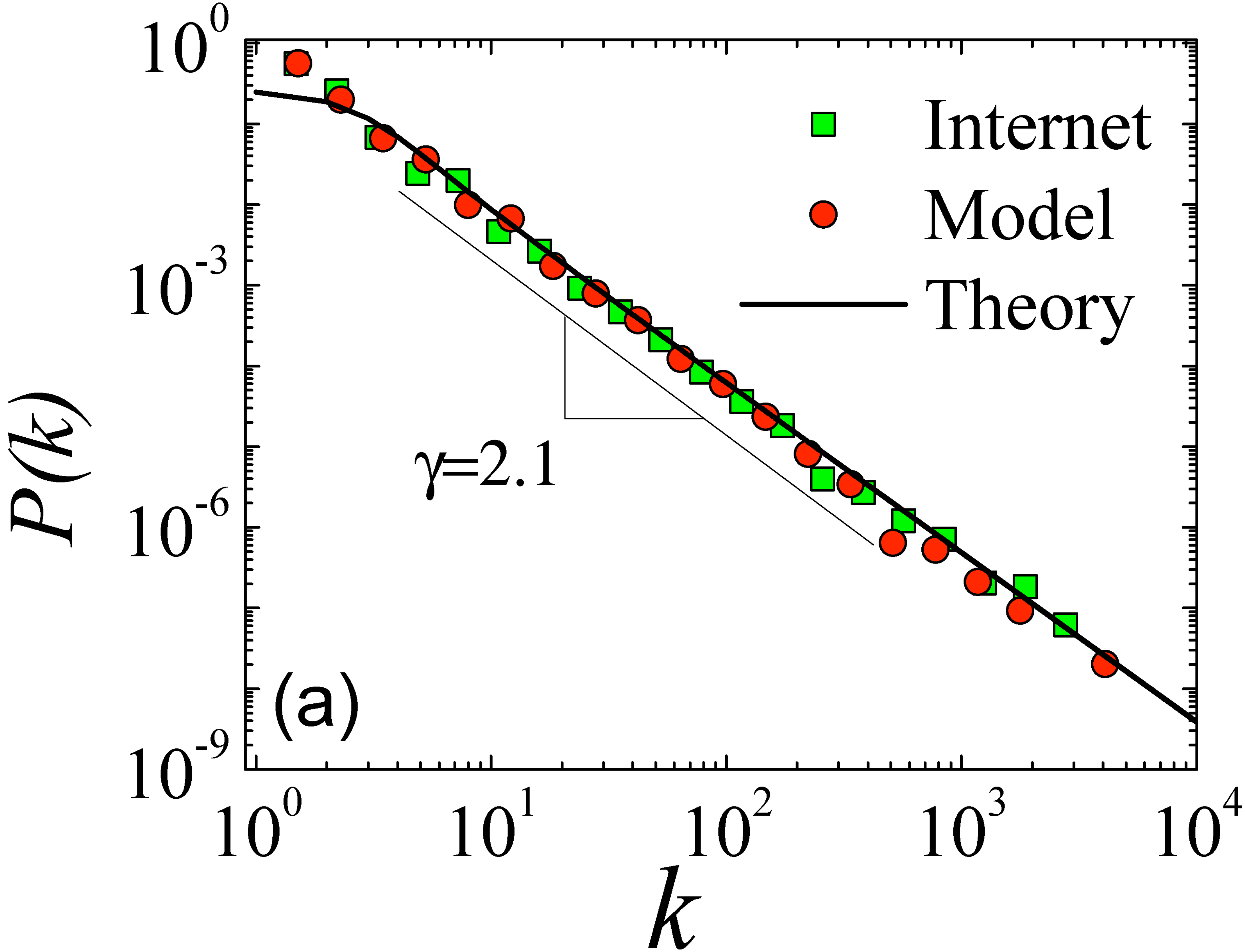
- Chemical potential R diverges as $-\ln(|T-1|)$

Hot regime $T>1$

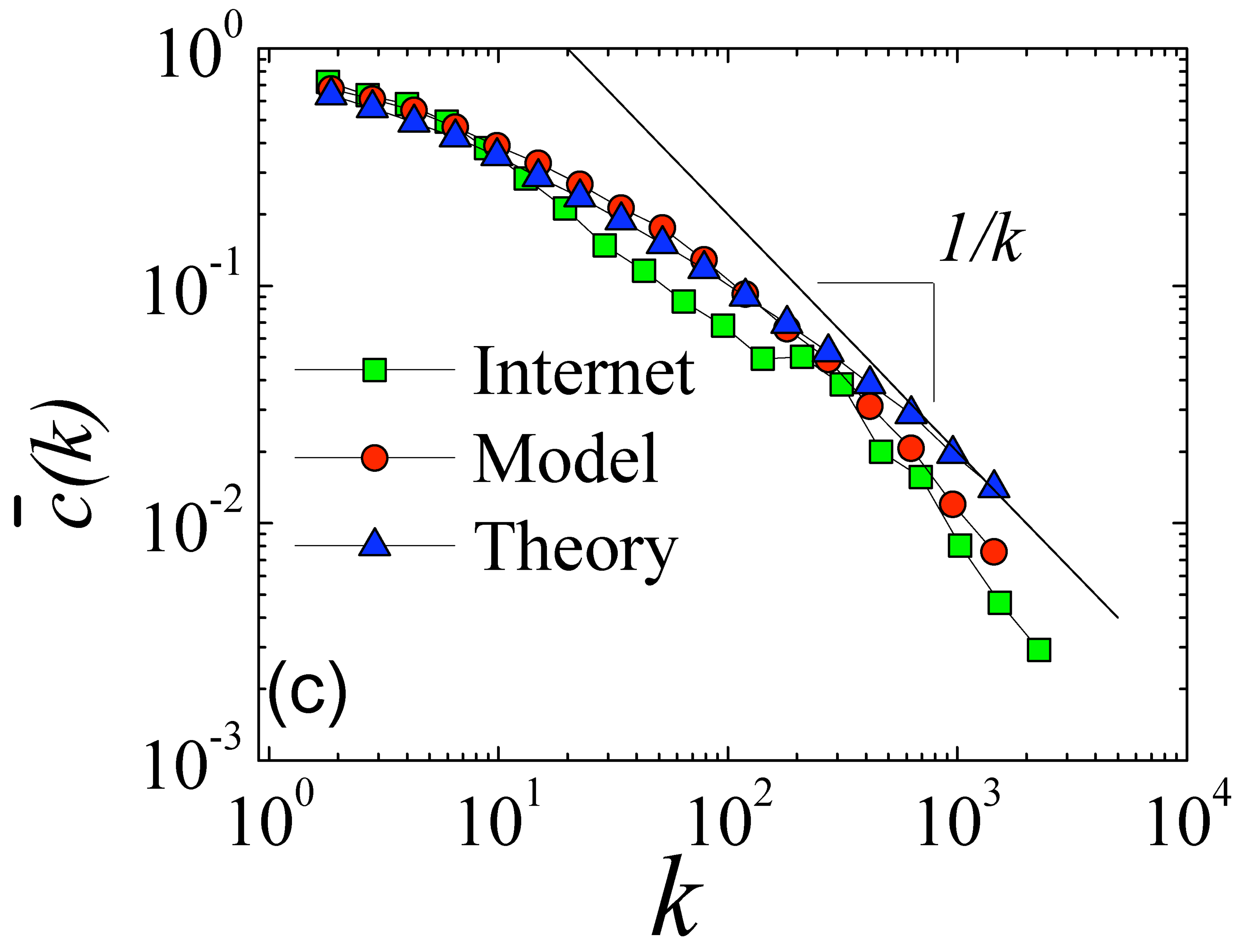
- Chemical potential $R=T(2/\zeta)\ln(N/\nu)$
- Clustering is zero
- Power law exponent γ does depend on T, $\gamma=T(2/\zeta)+1$

Two famous limits at $T \to \infty$

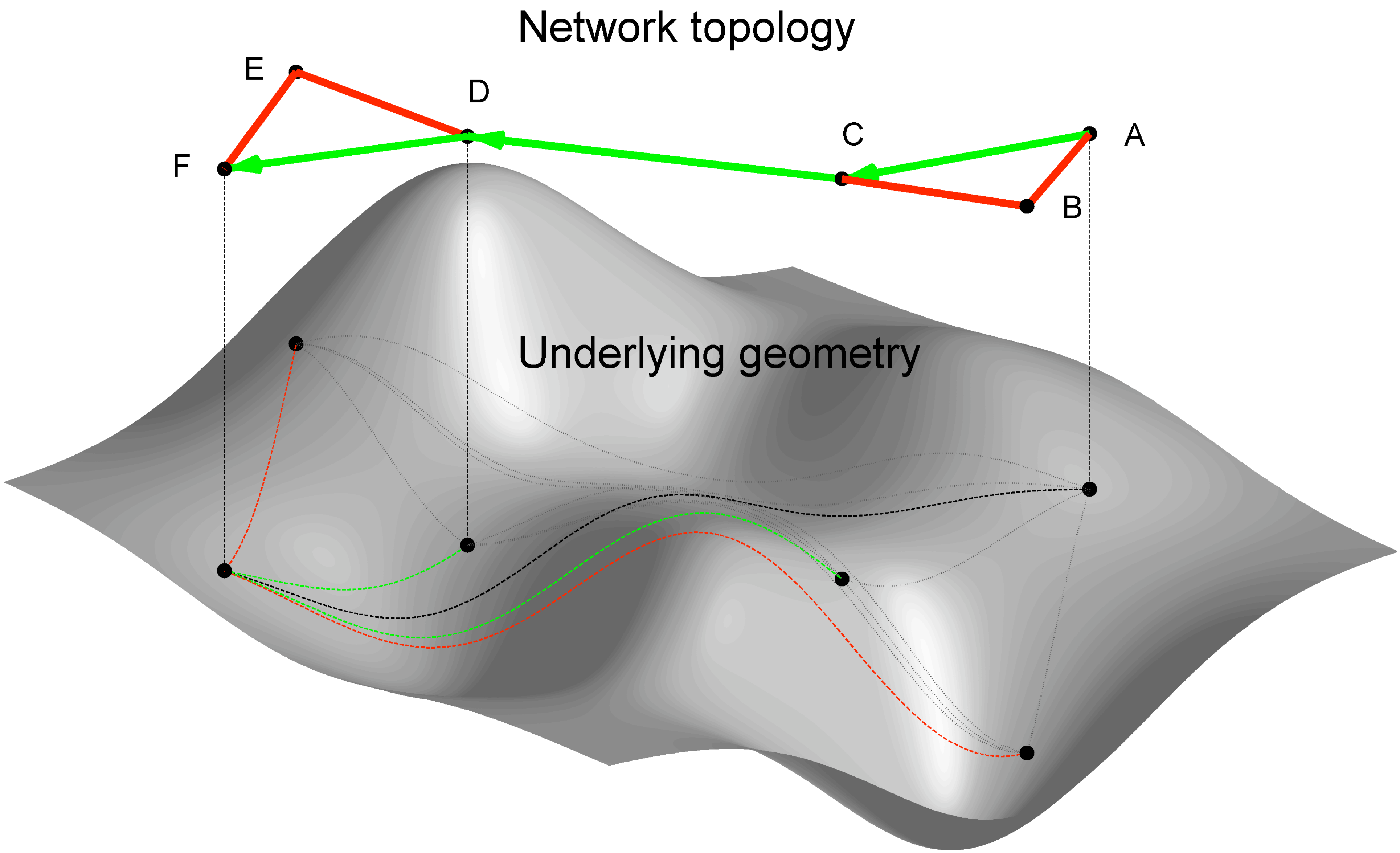
- Classical random graphs
  (random graphs with given average degree):
  $T \to \infty$; ζ fixed
- Configuration model
  (random graphs with given expected degrees):
  $T \to \infty$; $\zeta \to \infty$; T/ζ fixed



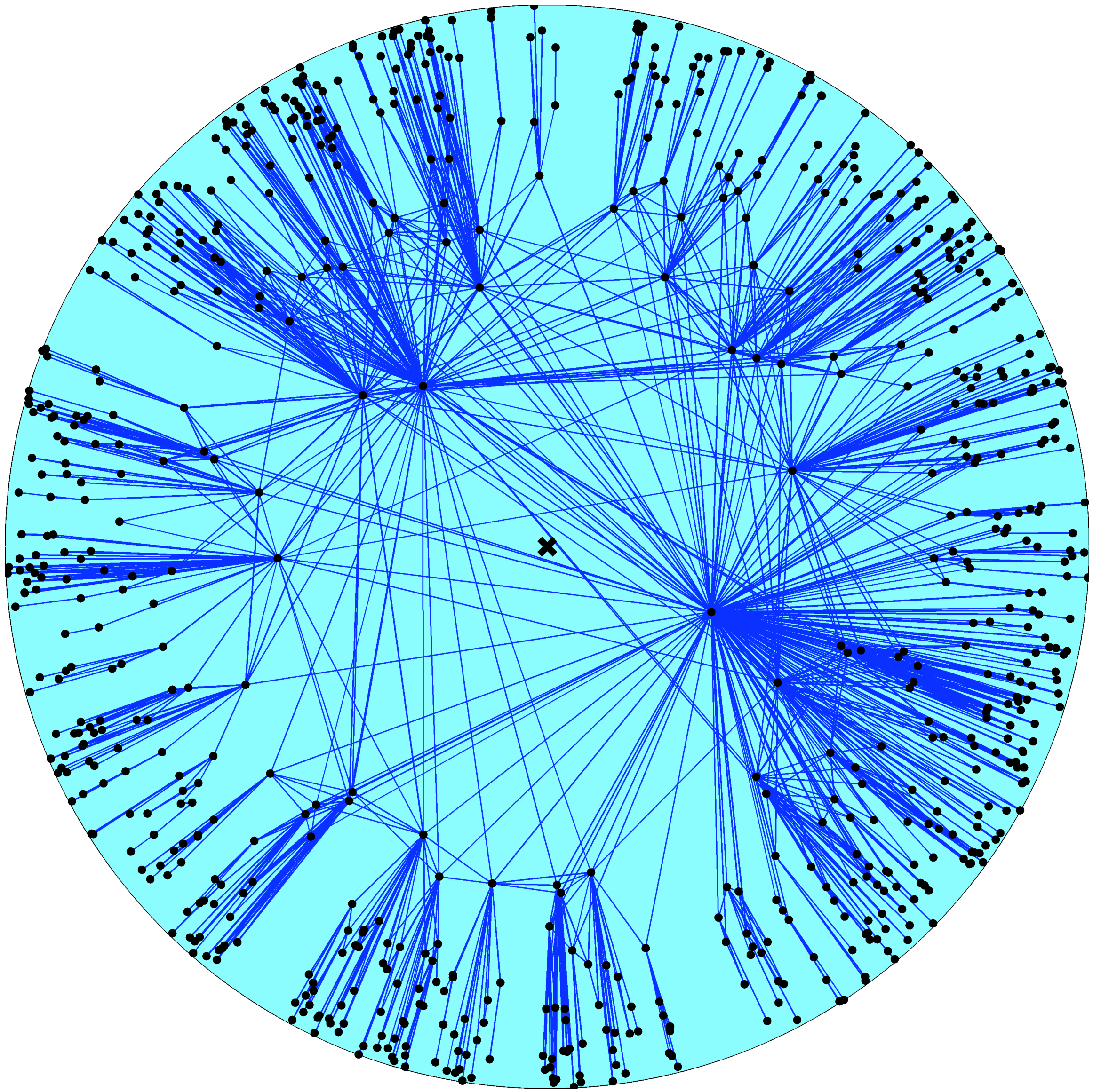
Internet
Model
Theory
γ=2.1
(a)
P(k)
k
10^0
10^-3
10^-6
10^-9
10^0
10^1
10^2
10^3
10^4

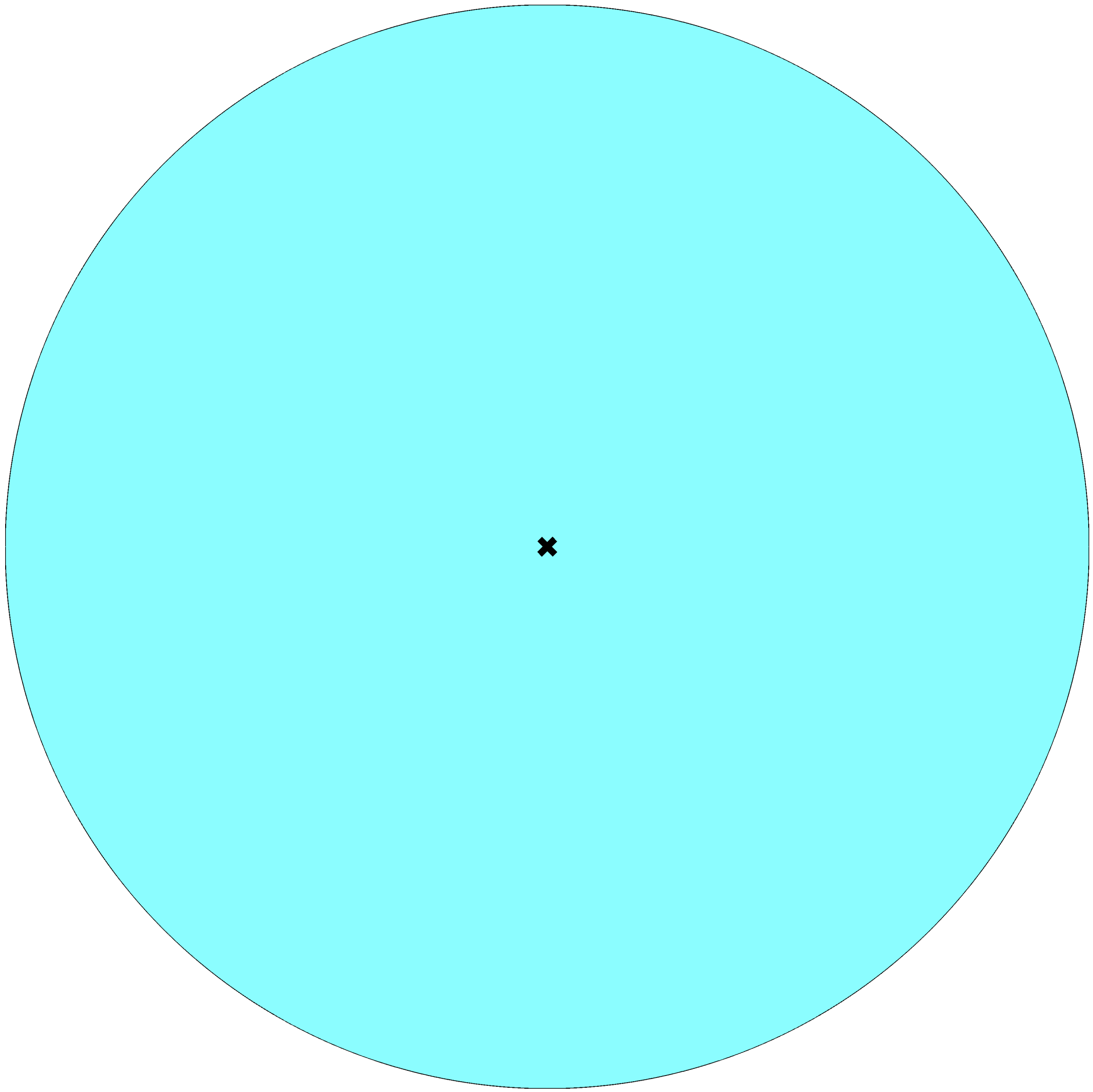


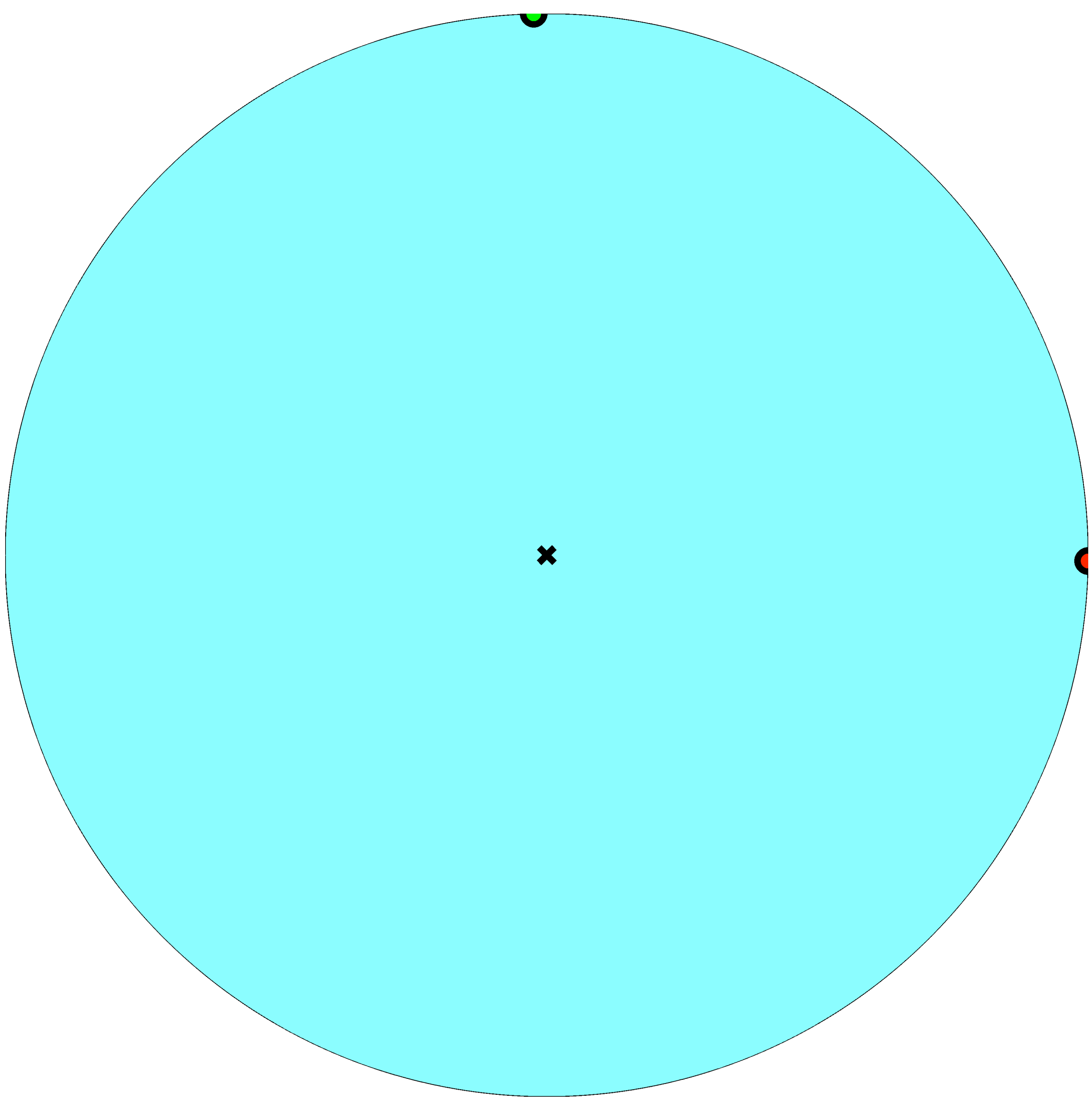
10^0
10^{-1}
10^{-2}
10^{-3}
10^0
10^1
10^2
10^3
10^4
$\bar{c}(k)$
k
$1/k$
Internet
Model
Theory
(c)

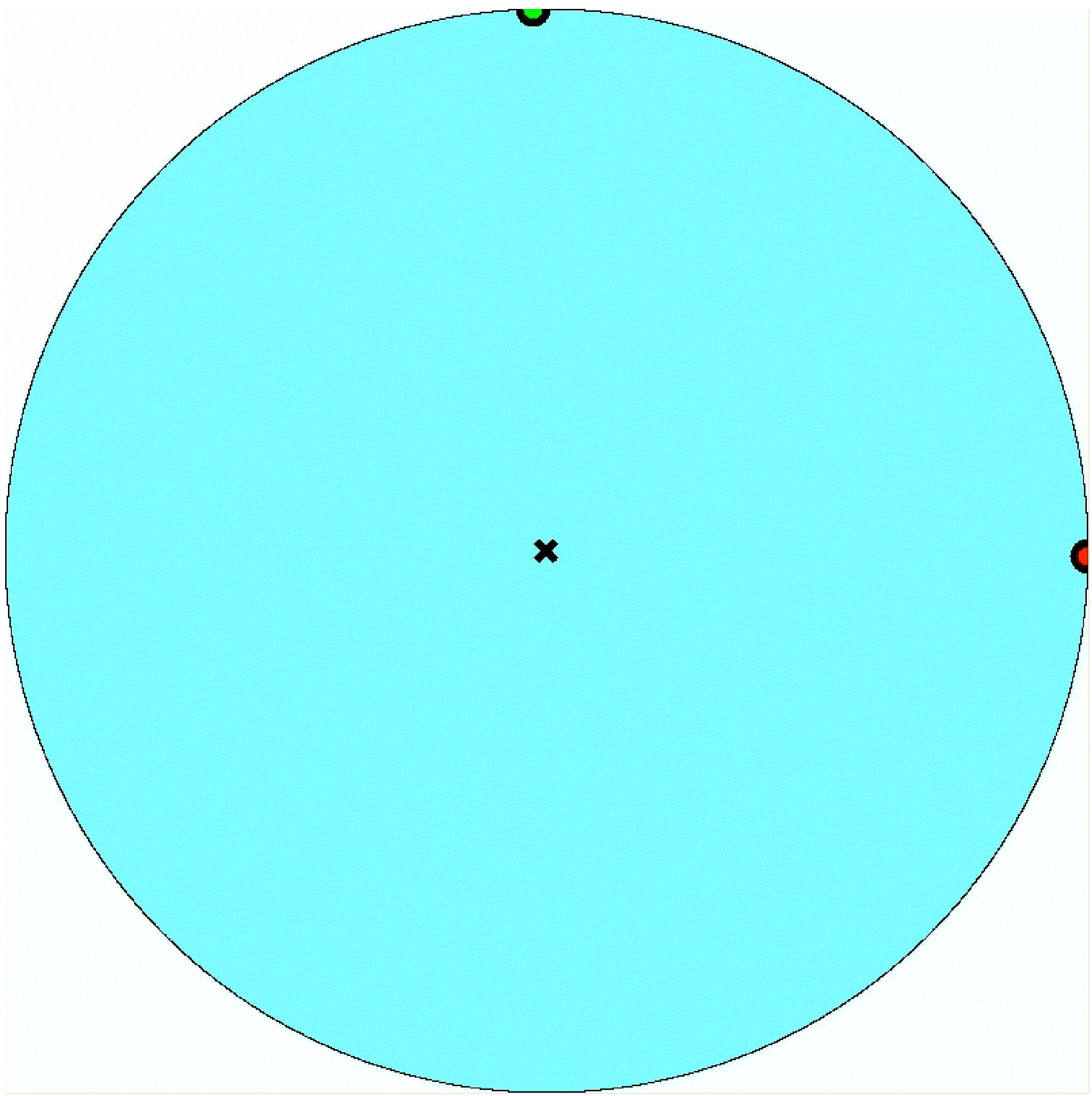


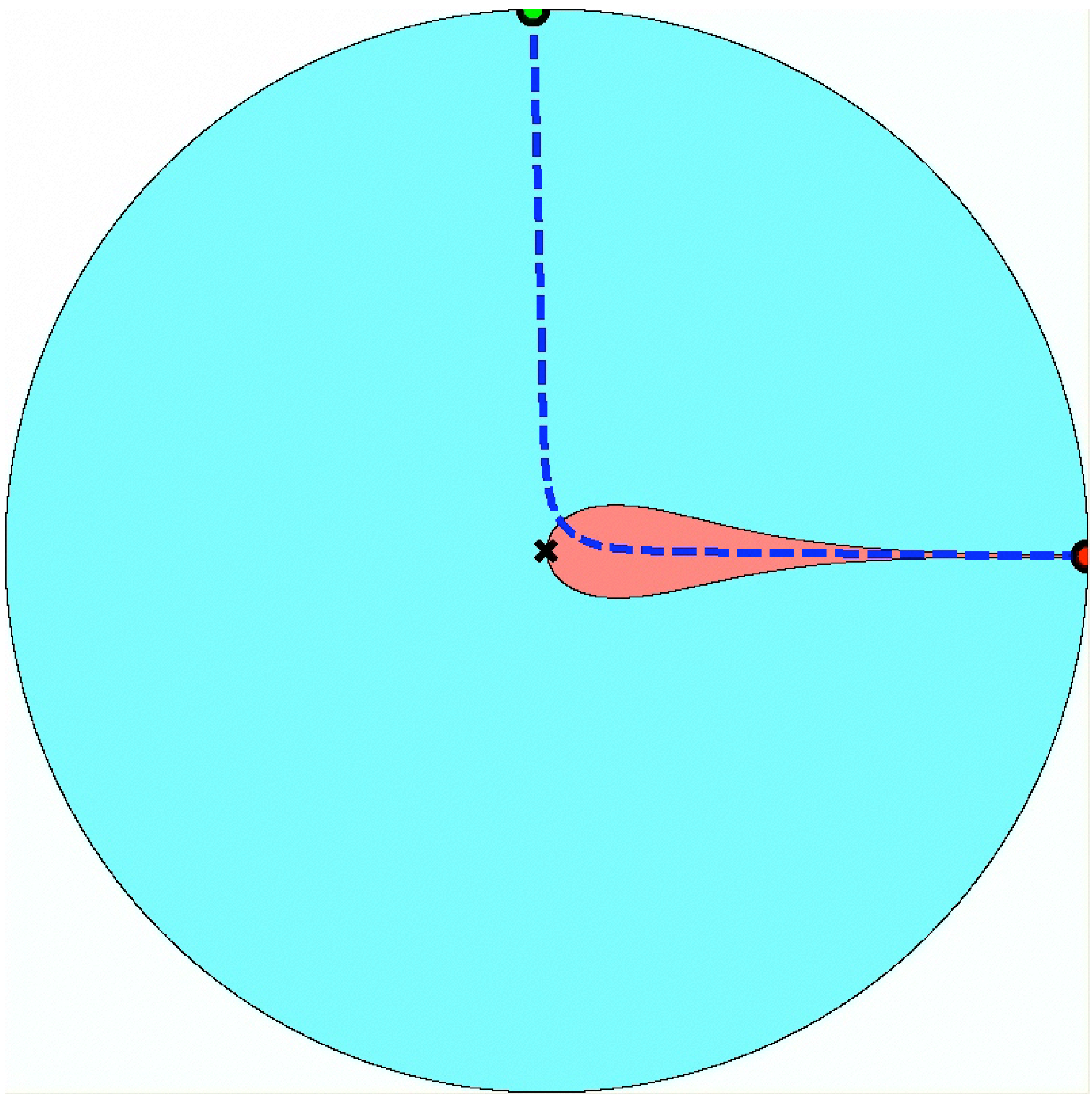
Network topology
E
D
C
A
F
B
Underlying geometry

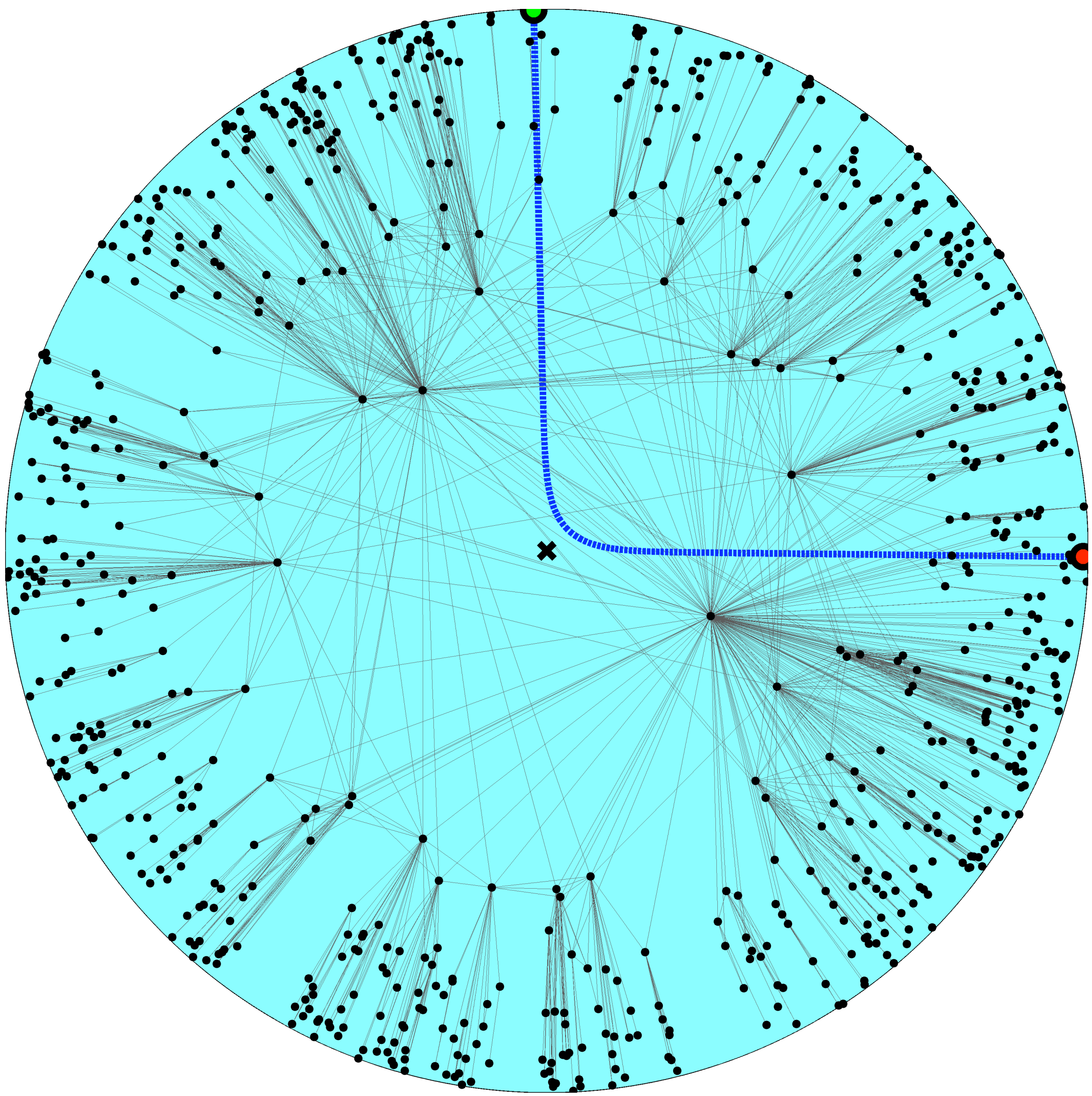


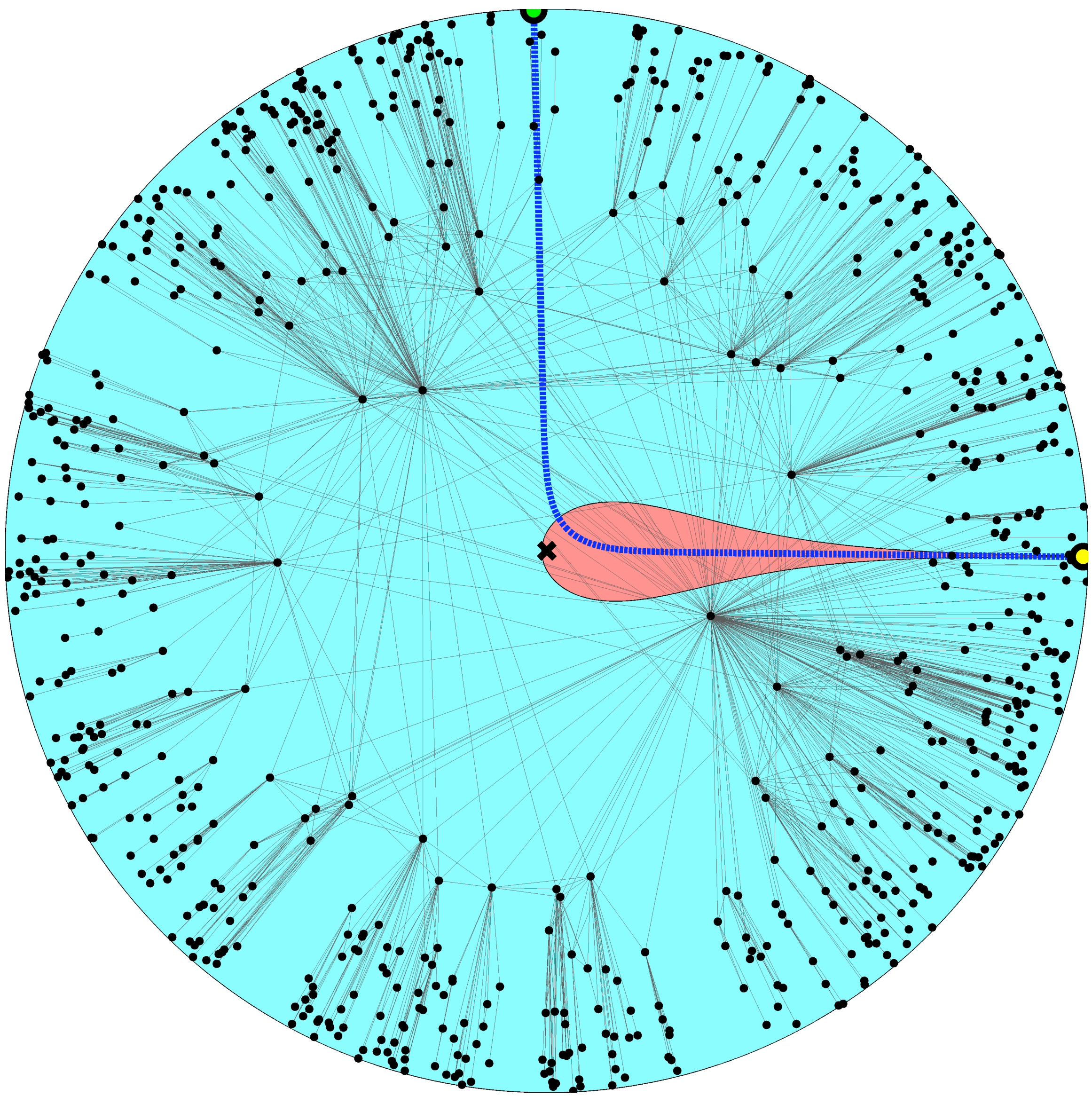


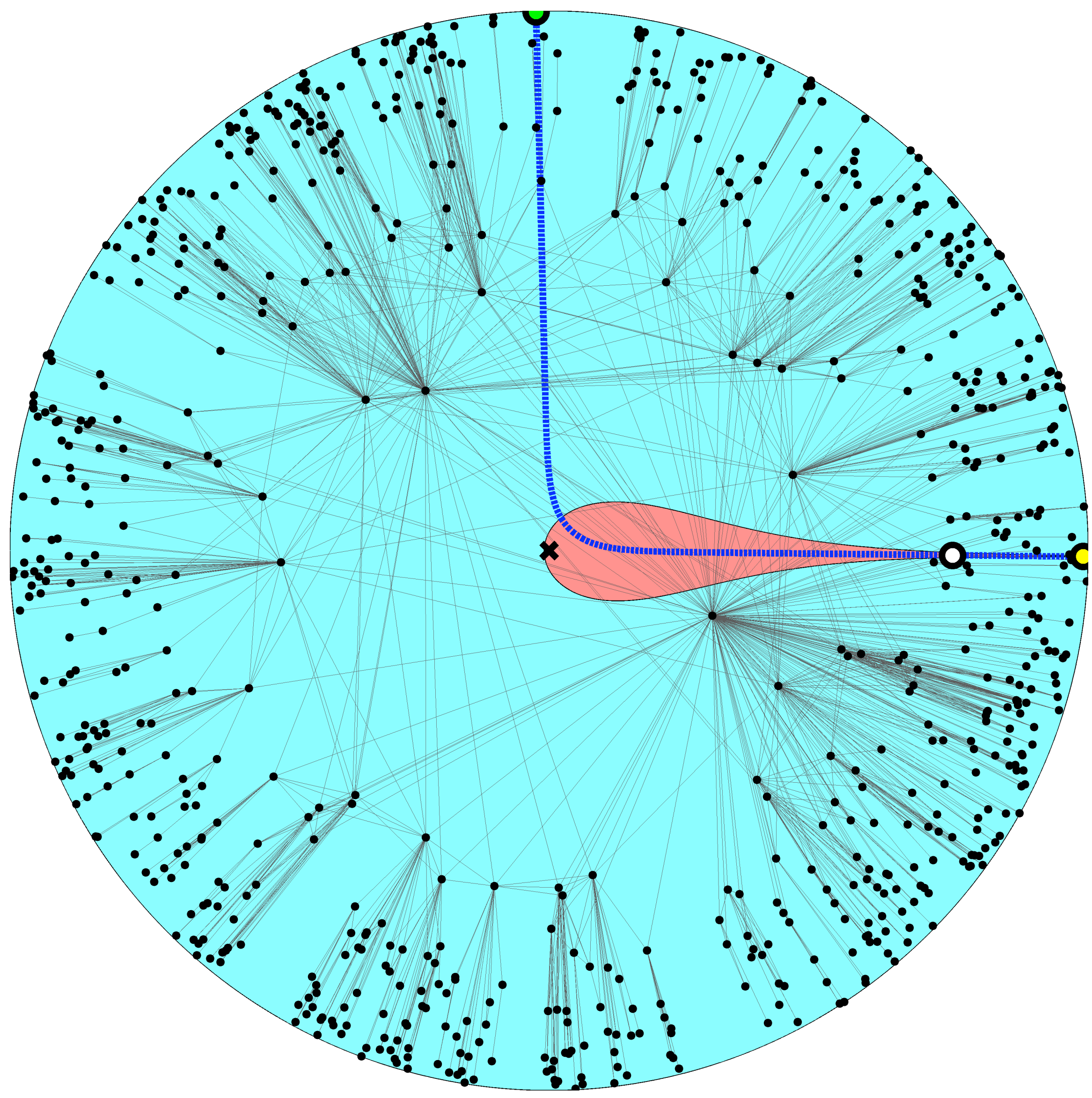


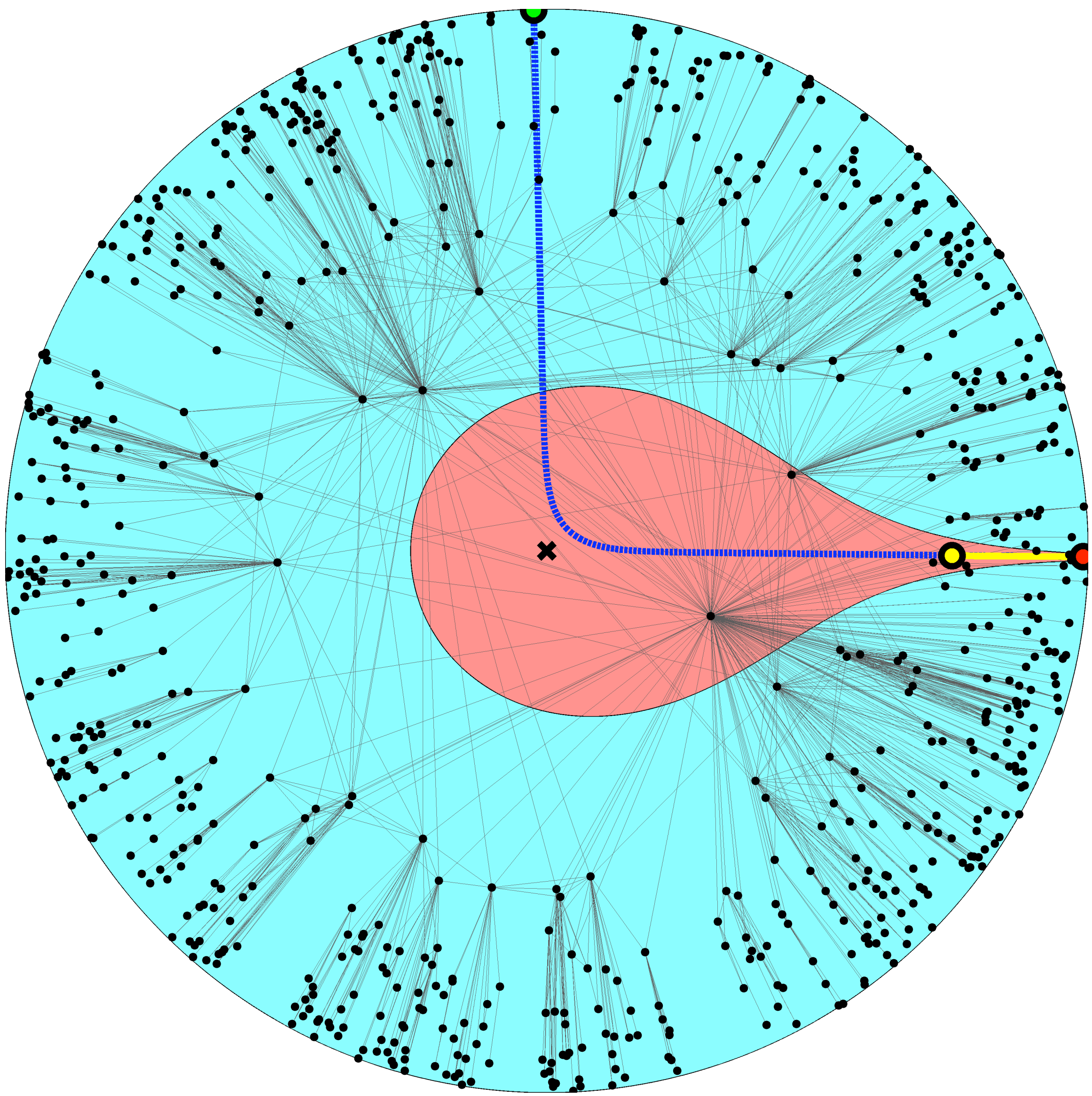


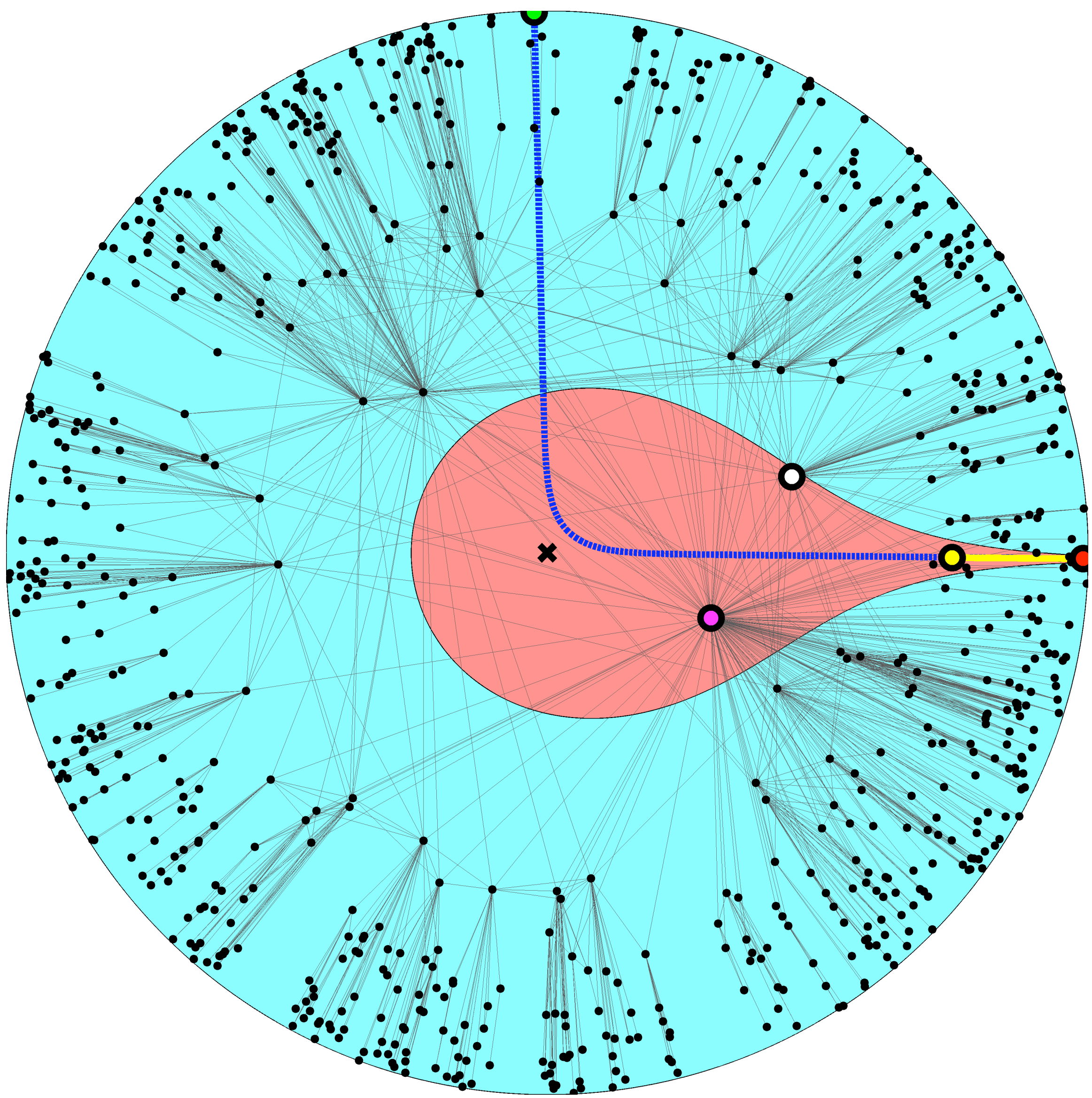


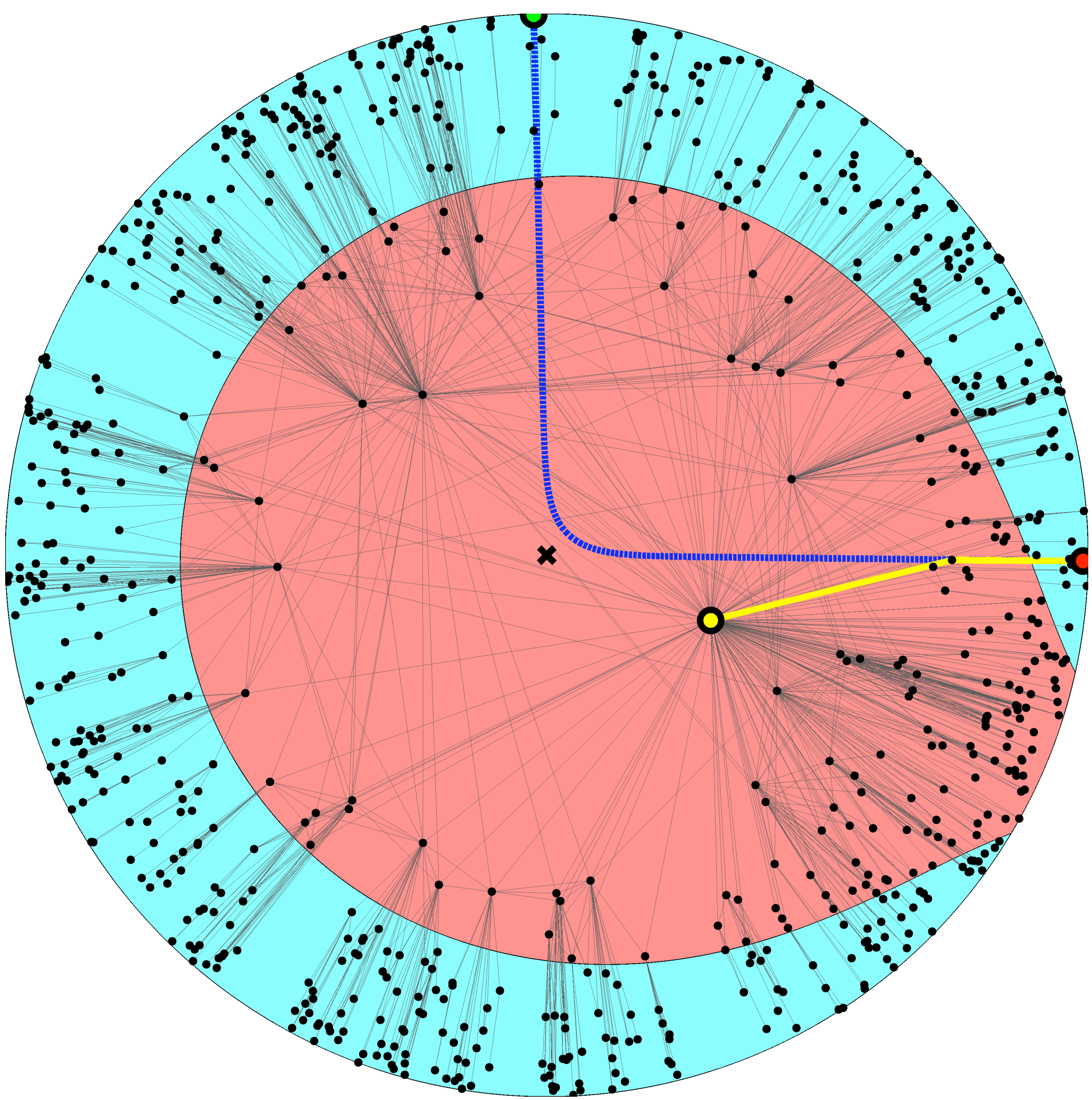


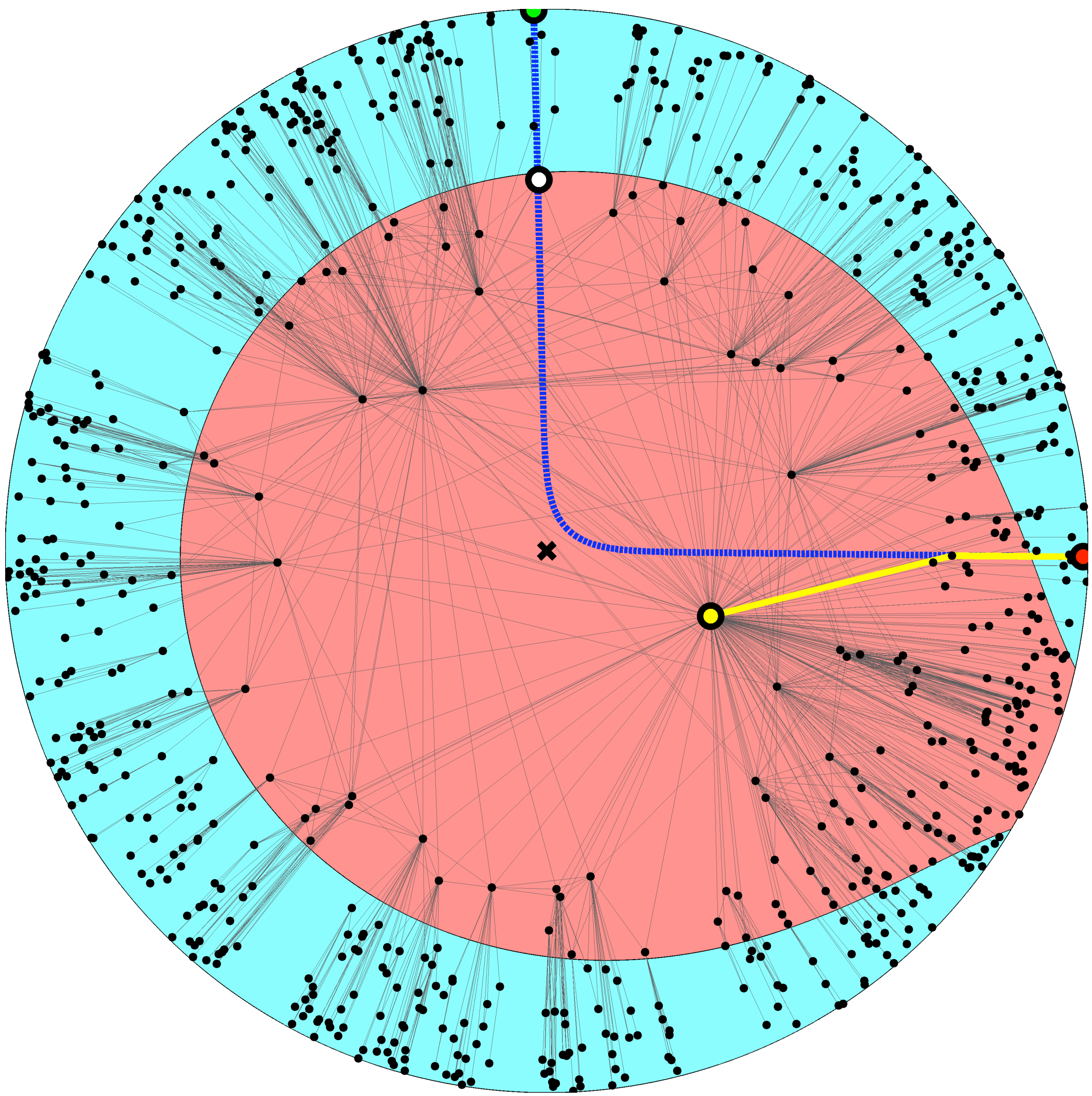


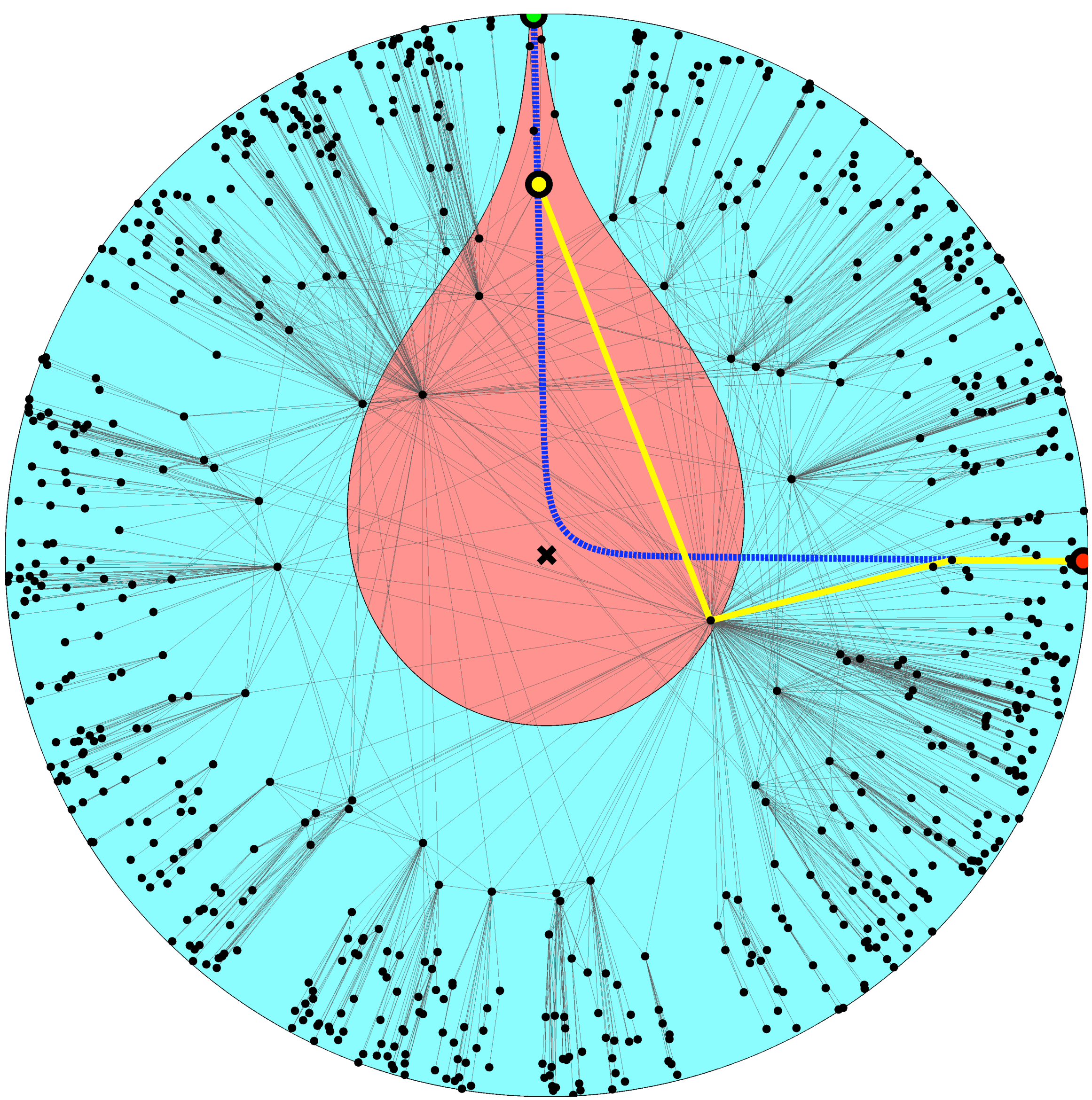


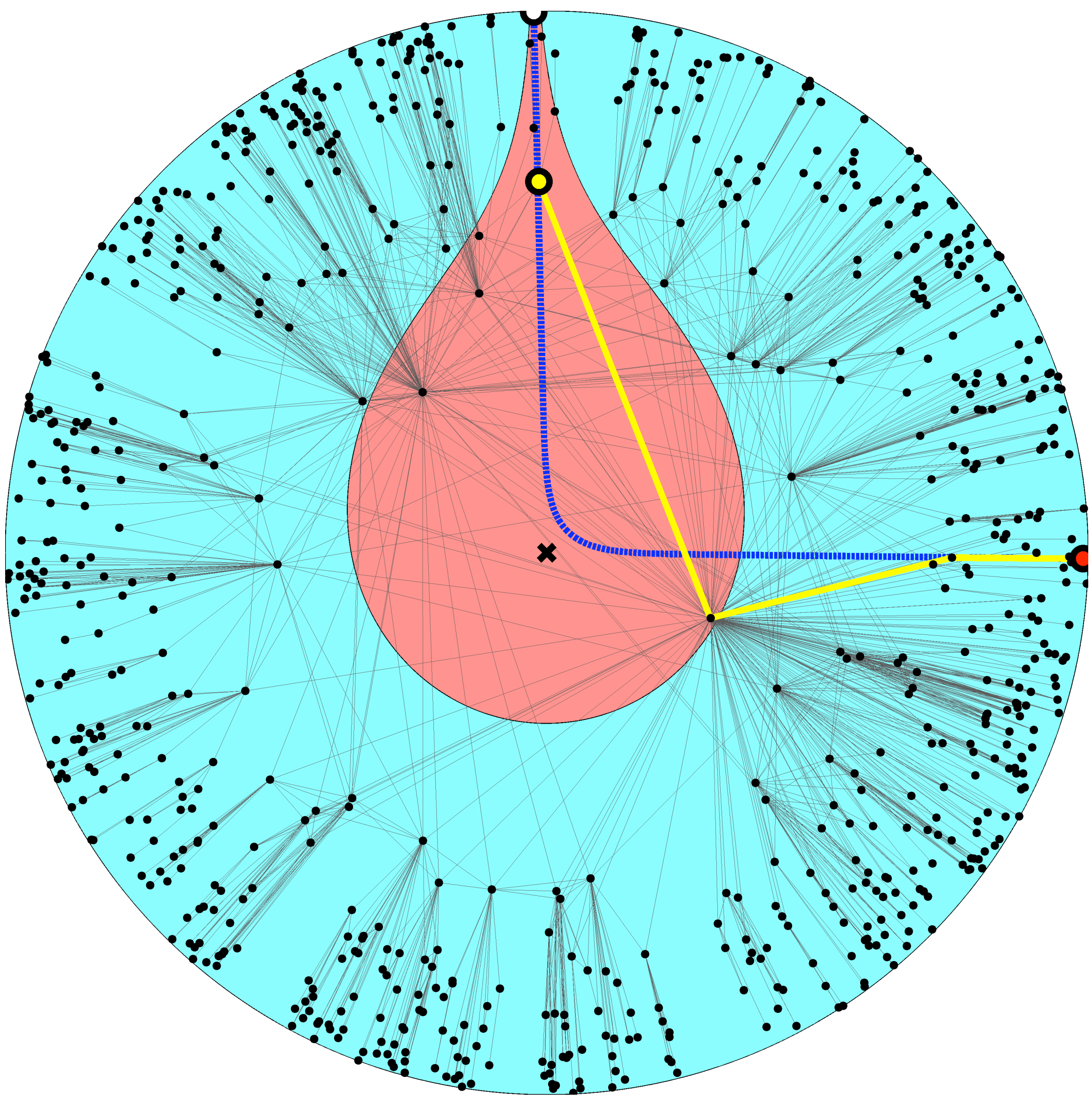


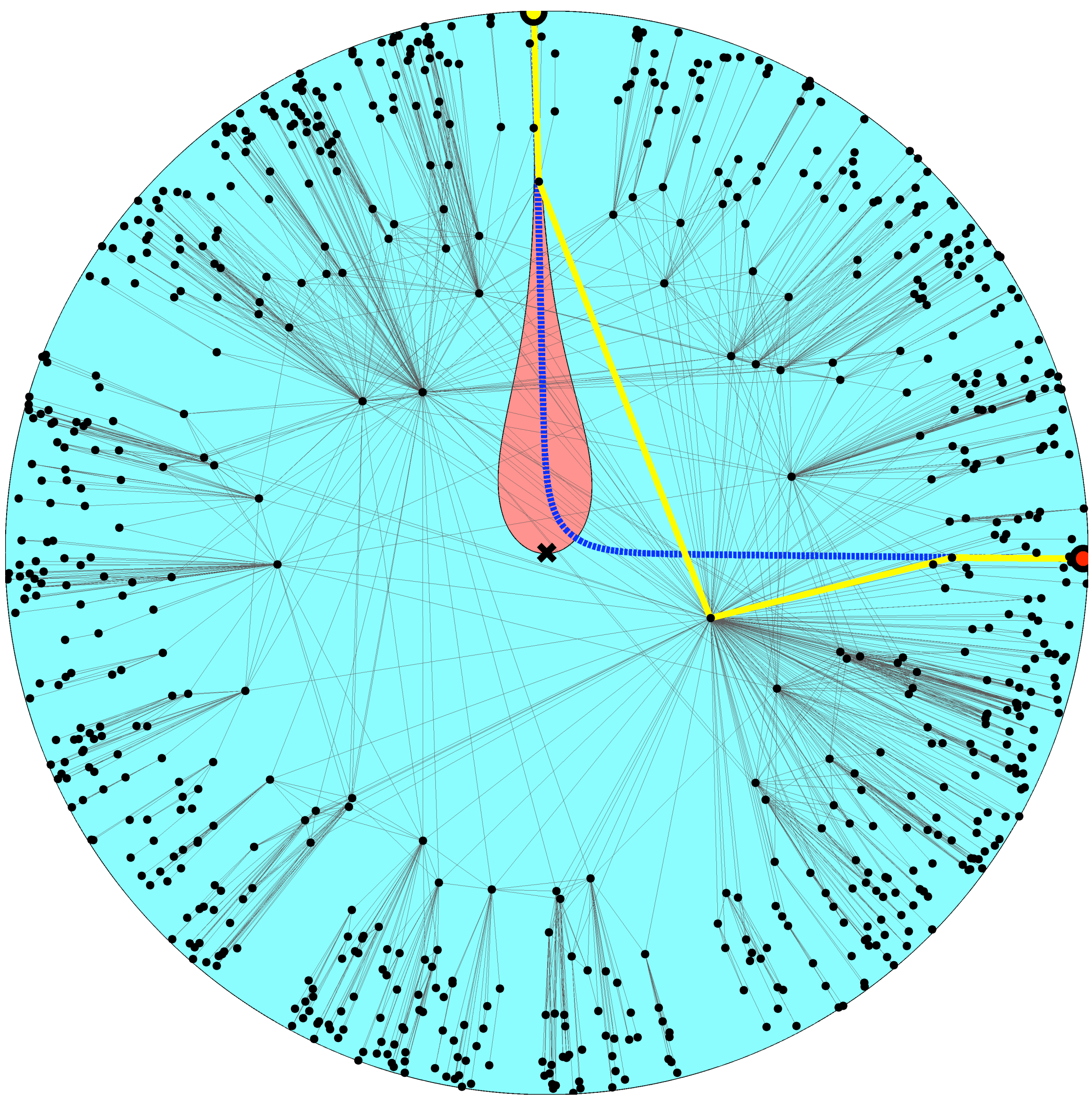


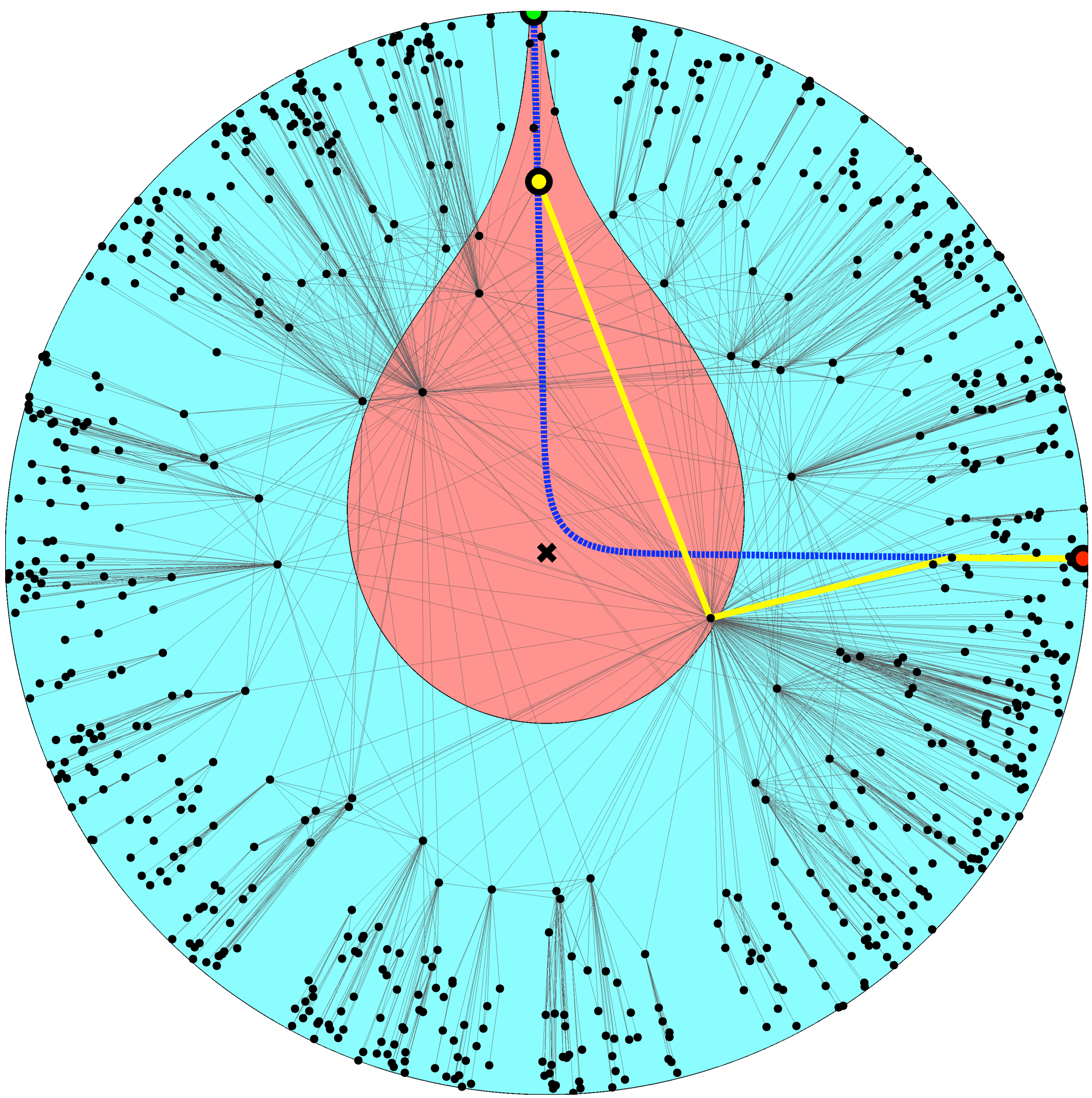


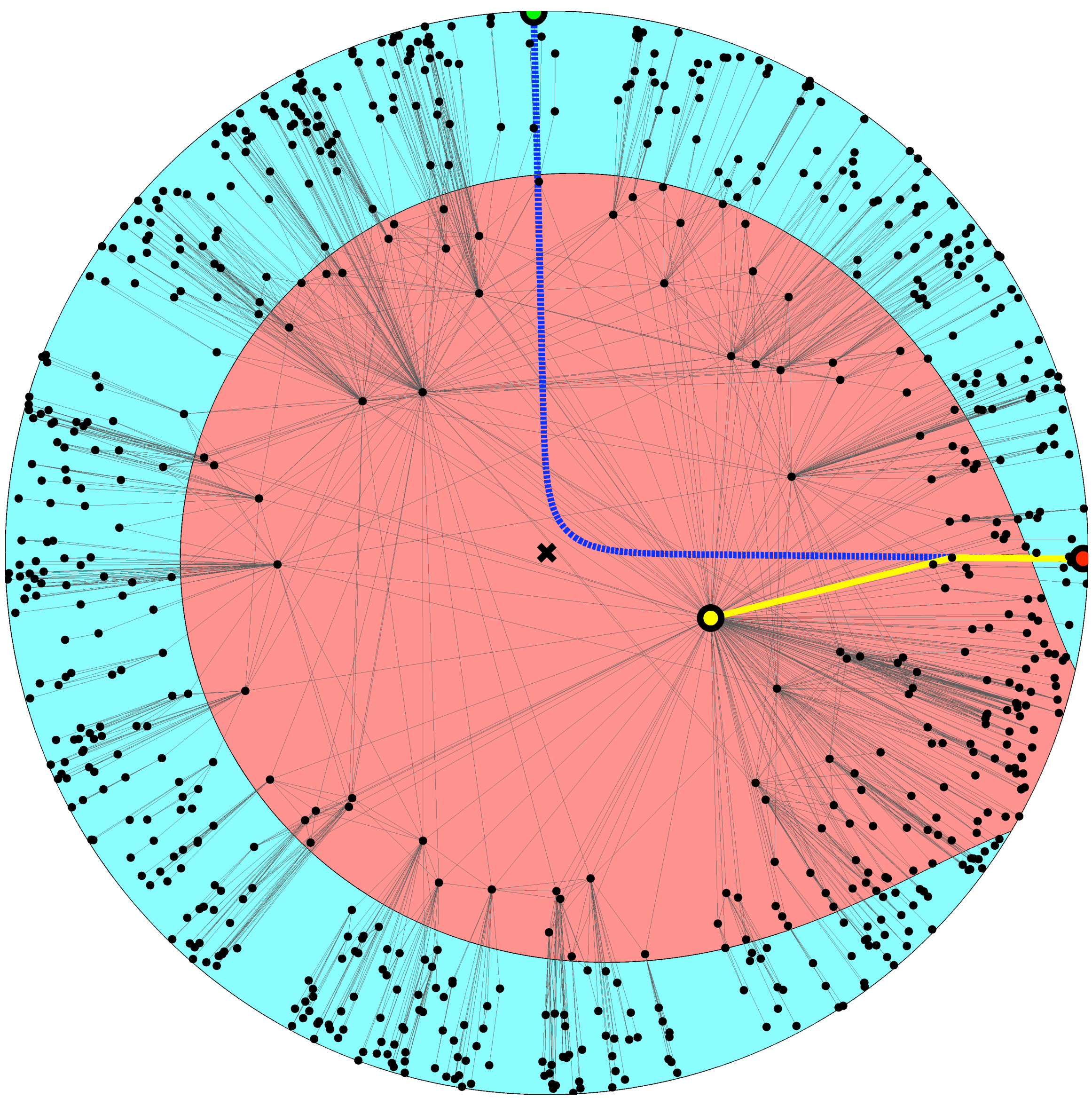


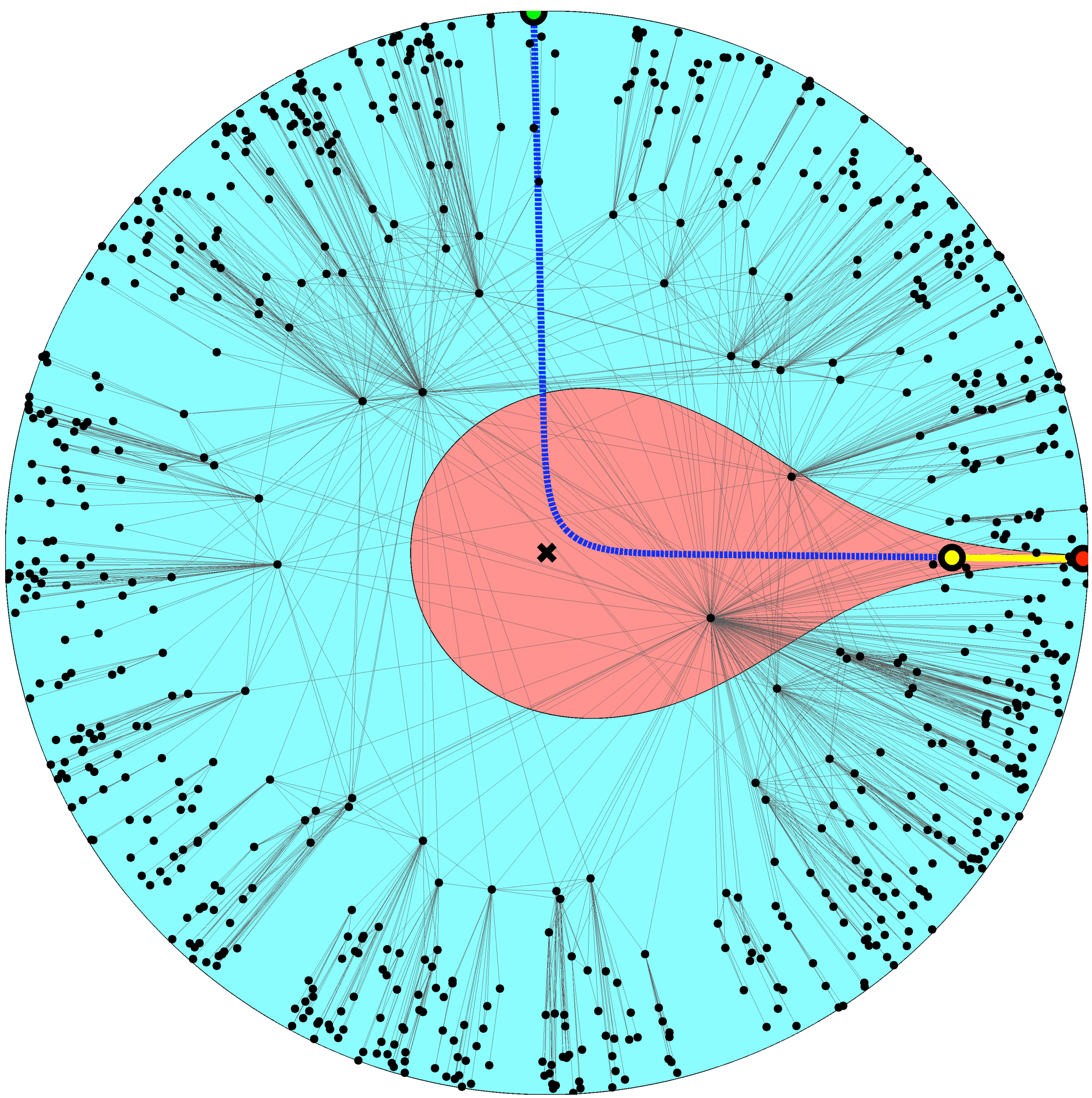


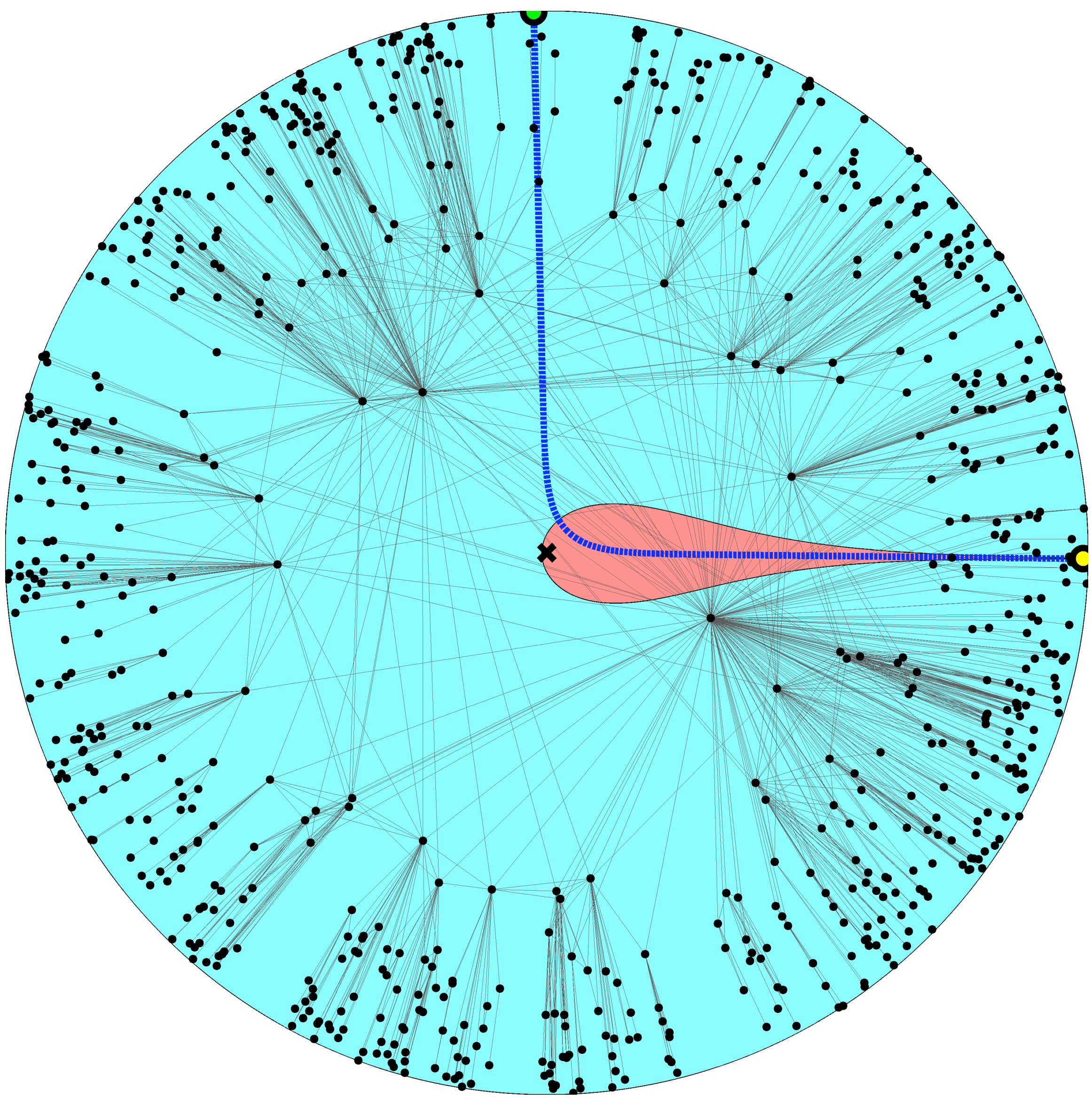


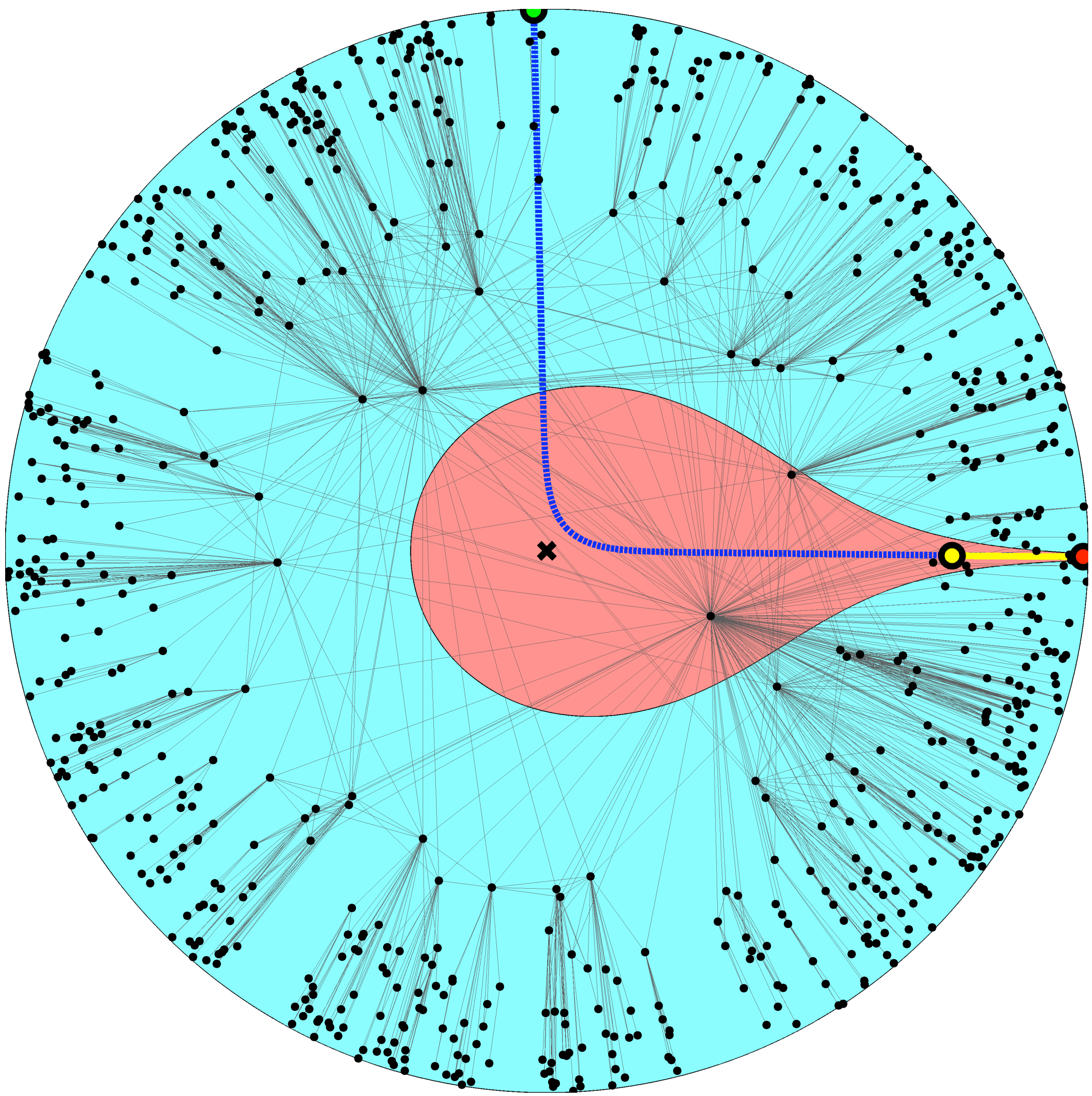


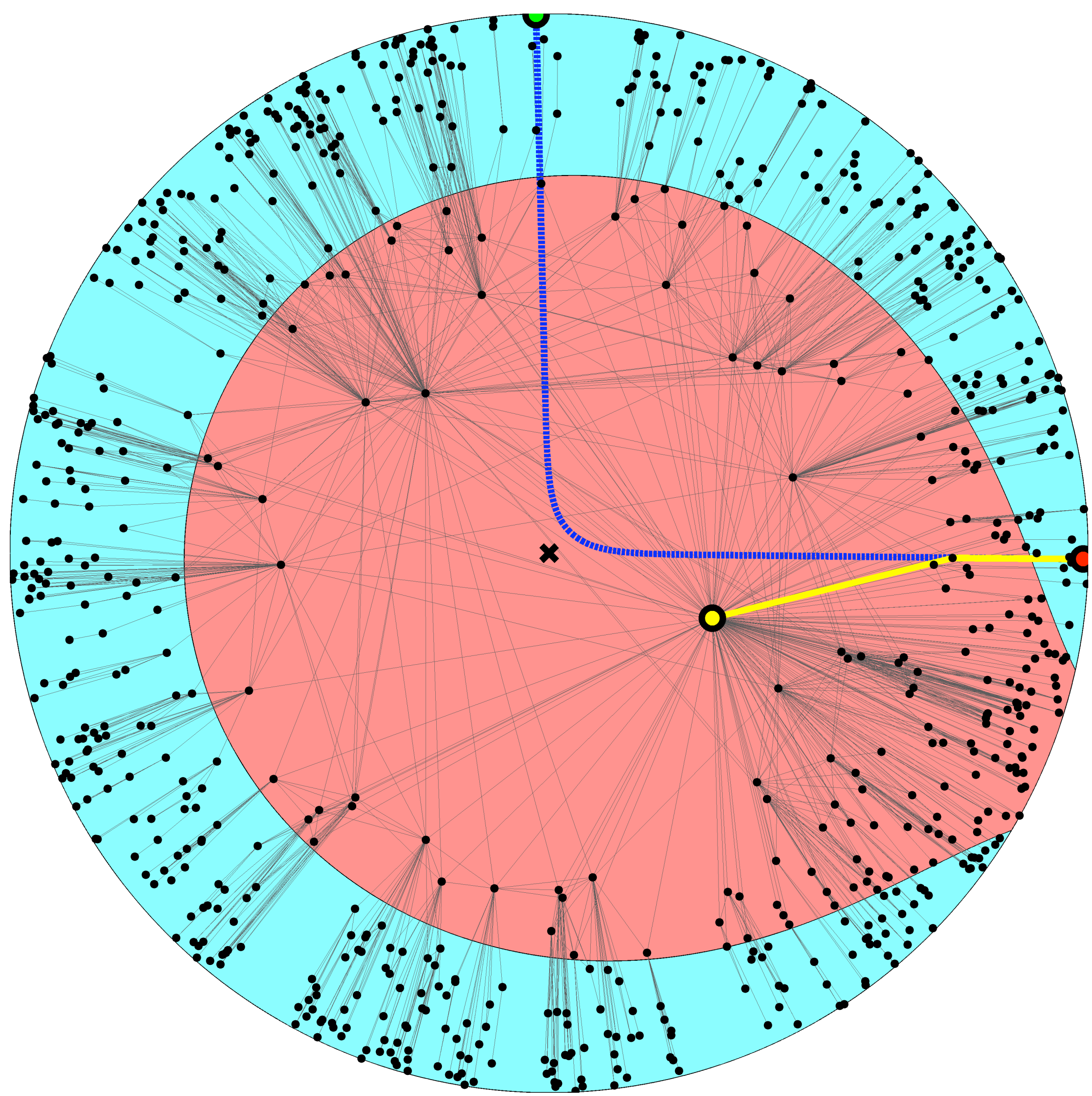


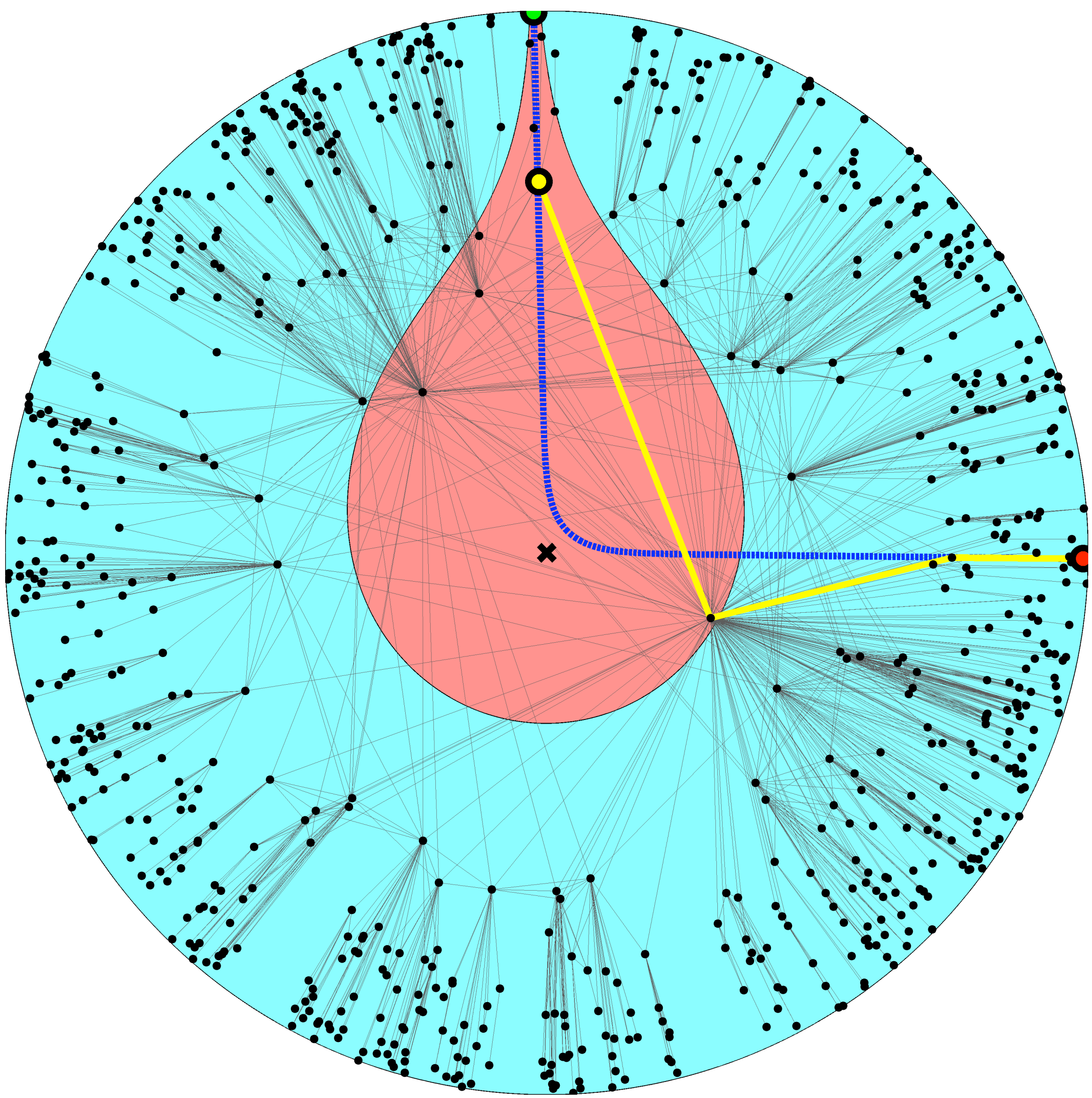


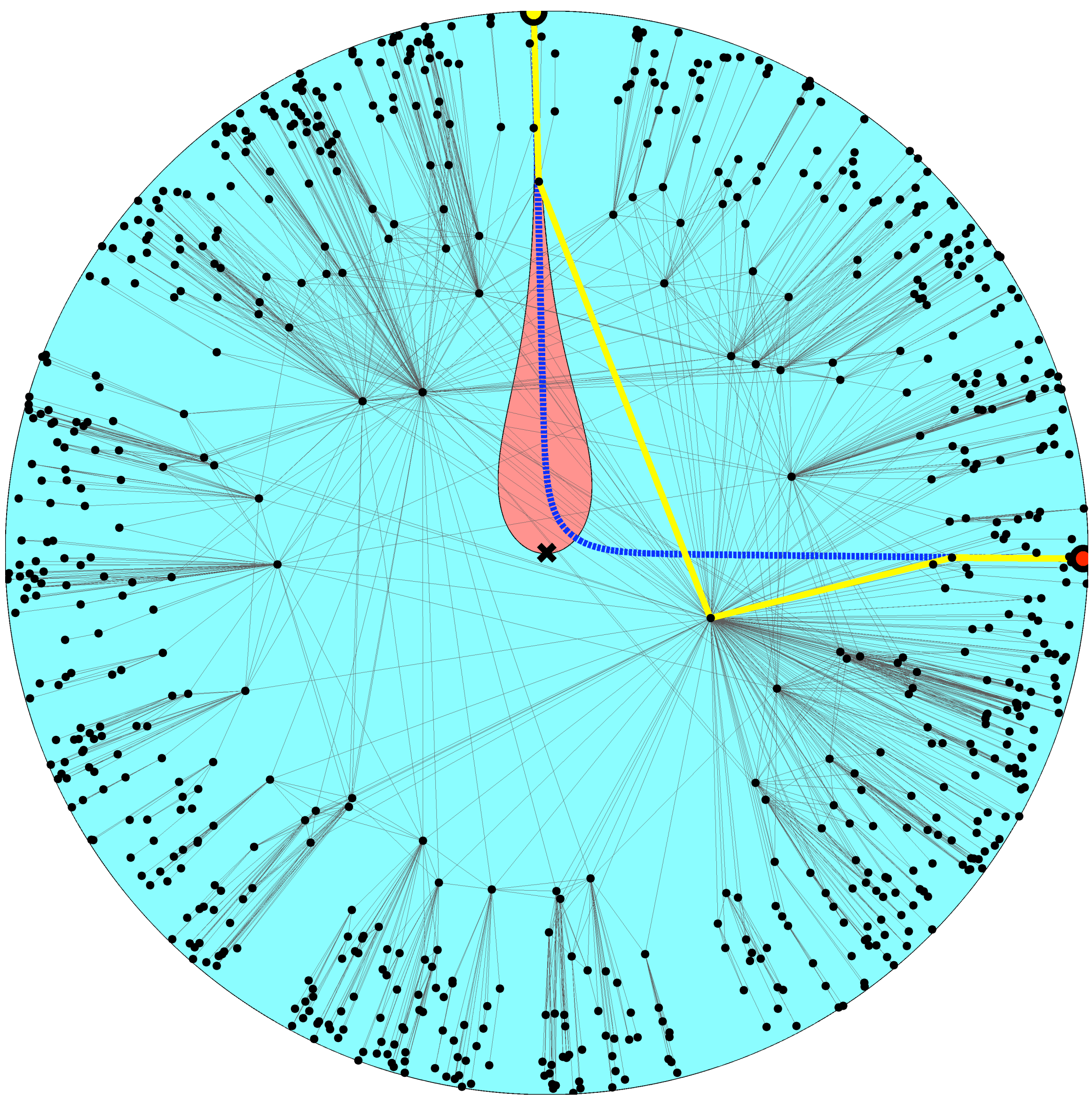


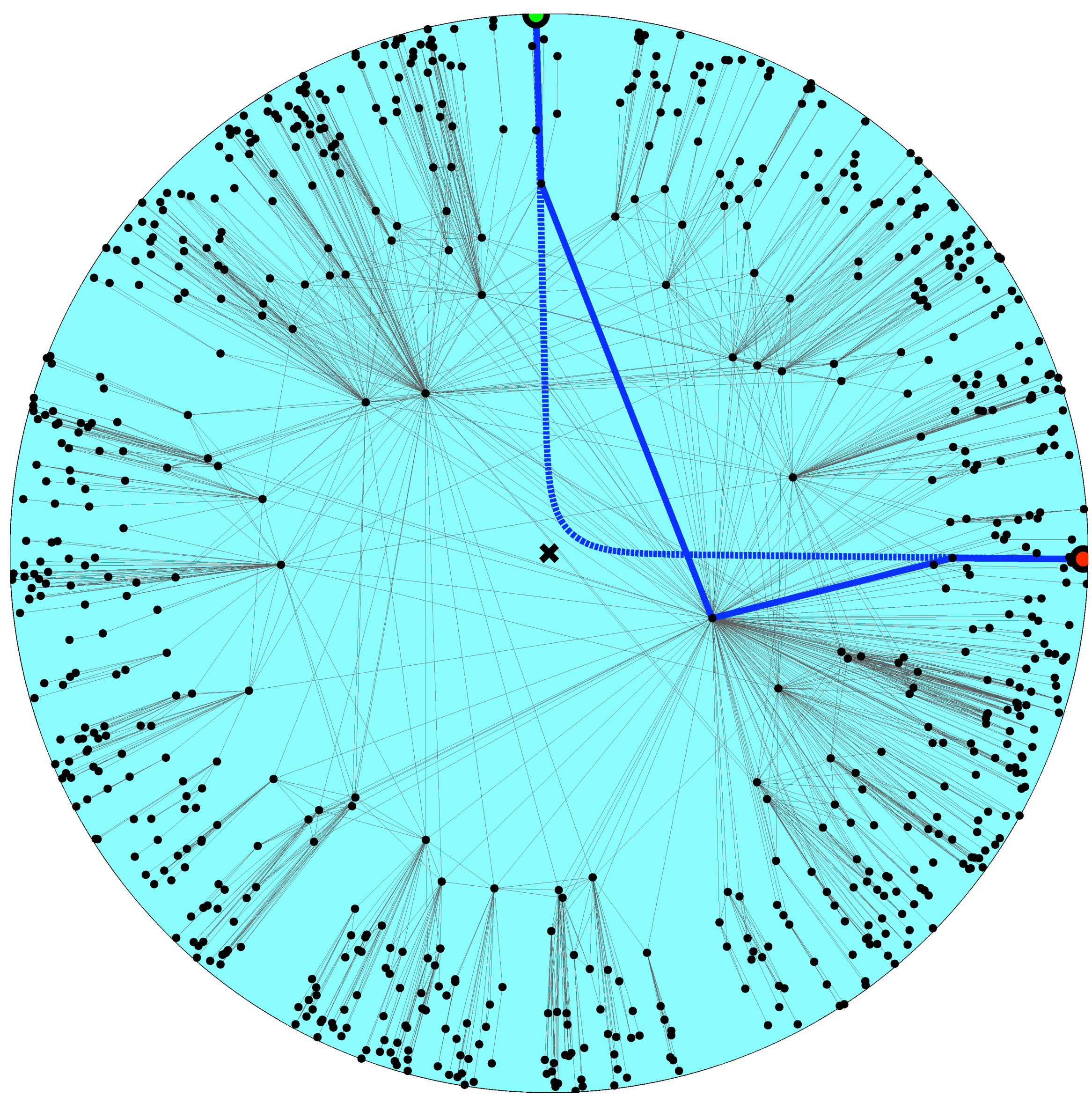












Navigation efficiency
(γ=*2.1*; *T*=*0*)

- Percentage of successful greedy paths

  99.99%

- Percentage of shortest greedy paths

  100%

- Percentage of successful greedy paths after removal of *x%* of links or nodes
  - *x*=10% → 99%
  - *x*=30% → 95%

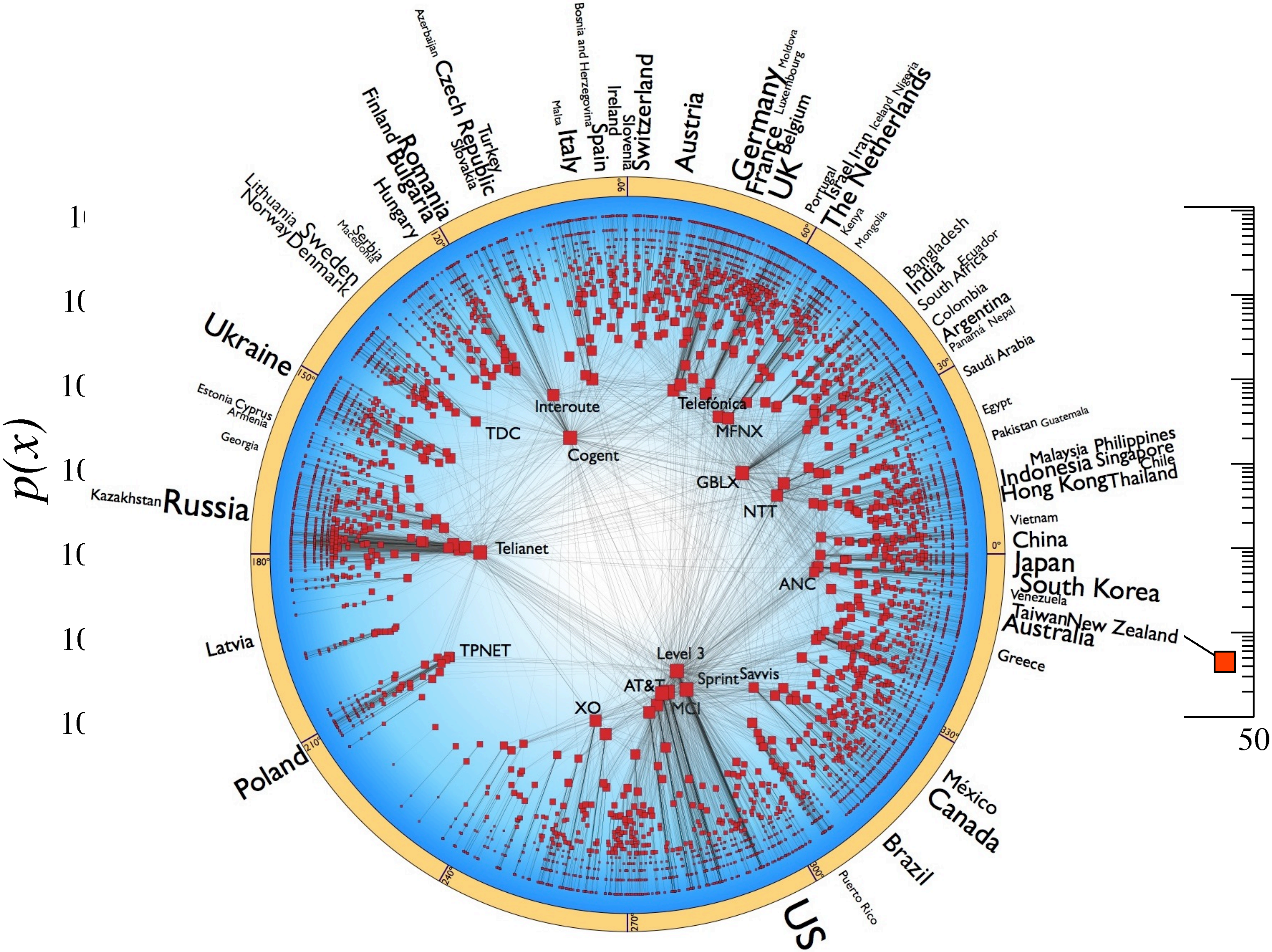
Mapping the real Internet using statistical inference methods

- Measure the Internet topology properties
  - N, $\langle k \rangle$, γ, c
- Map them to model parameters
  - R, ν, ζ, T
- Place nodes at hyperbolic coordinates (r,θ)
  - $k \sim e^{\zeta(R-r)/2}$
  - θ's are uniformly distributed on $[0,2\pi]$
- Apply the Metropolis-Hastings algorithm to find θ's maximizing the likelihood that Internet is produced by the model

Metropolis-Hastings

$$L = \prod_{i<j} p(x_{ij})^{a_{ij}} [1 - p(x_{ij})]^{1-a_{ij}}$$

- Compute current likelihood L_c
- Select a random node
- Move it to a new random angular coordinate
- Compute new likelihood L_n
- If $L_n > L_c$, accept the move
- If not, accept it with probability L_n / L_c
- Repeat


p(x)
50
Russia
Kazakhstan
Ukraine
Estonia
Cyprus
Armenia
Georgia
Lithuania
Norway
Sweden
Denmark
Macedonia
Serbia
Hungary
Finland
Romania
Bulgaria
Azerbaijan
Czech Republic
Turkey
Slovakia
Malta
Italy
Bosnia and Herzegovina
Spain
Ireland
Slovenia
Switzerland
Austria
Germany
France
UK
Moldova
Luxembourg
Belgium
Portugal
Israel
Iran
Iceland
Nigeria
The Netherlands
Kenya
Mongolia
Bangladesh
India
Ecuador
South Africa
Colombia
Argentina
Panamá
Nepal
Saudi Arabia
Egypt
Pakistan
Guatemala
Malaysia
Philippines
Indonesia
Singapore
Chile
Hong Kong
Thailand
Vietnam
China
Japan
South Korea
Venezuela
Taiwan
New Zealand
Australia
Greece
México
Canada
Brazil
Puerto Rico
US
Poland
Latvia
Interoute
TDC
Cogent
Telefónica
MFNX
GBLX
NTT
Telianet
ANC
TPNET
Level 3
AT&T
Sprint
Savvis
MCI
XO
0°
30°
60°
90°
120°
150°
180°
210°
240°
270°
300°
330°

Navigation efficiency (γ=*2.1*; *T*=*0.69*)

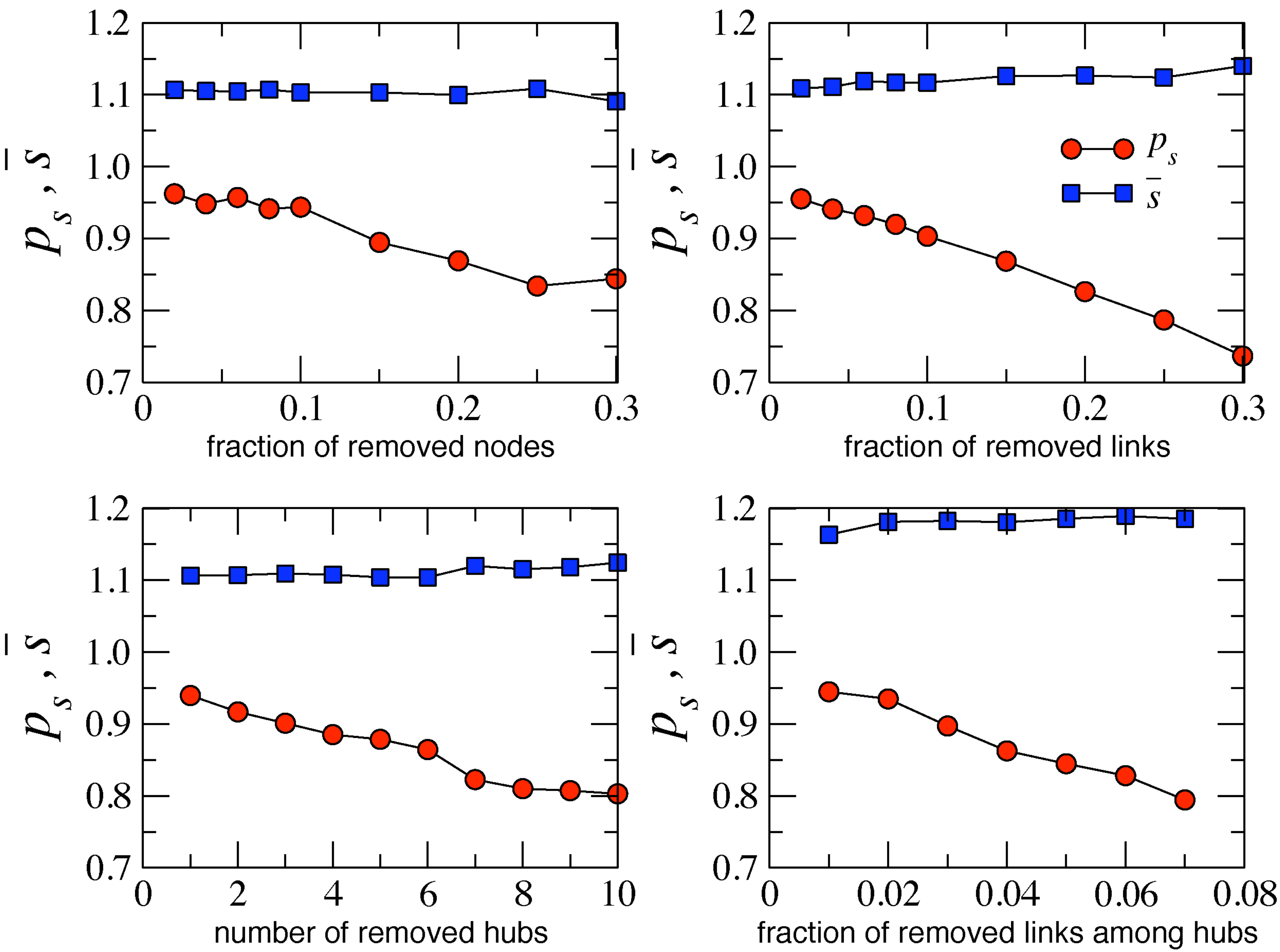
- Percentage of successful greedy paths

  97%

- Average stretch

  1.1

- Percentage of successful greedy paths after removal of *x%* of links or nodes:



1.2
1.1
1.0
0.9
0.8
0.7
0
0.1
0.2
0.3
p_s , $\bar{s}$
fraction of removed nodes
fraction of removed links
p_s
$\bar{s}$
number of removed hubs
2
4
6
8
10
0.02
0.04
0.06
0.08
fraction of removed links among hubs

Conclusions

- We have discovered a new *geometric* framework to study the structure and function of complex networks
- The framework provides a natural (no “enforcement” on our side) explanation of two basic structural properties of complex networks (*scale-free degree distribution and strong clustering*) as consequences of underlying hyperbolic geometry
- The framework subsumes the *configuration model and classical random graphs* as two limiting cases with degenerate geometric structures
- The framework explains how the hierarchical *structure* of complex networks ensures the maximum efficiency of their *function* – efficient navigation without global information
- The framework provides a basis for mapping real complex networks to their hidden/underlying hyperbolic spaces

Applications of network mapping

- Internet
  - optimal (maximally efficient/most scalable) routing
  - routing table sizes, stretch, and communication overhead approach their theoretical lower bounds
- Other networks
  - discover hidden distances between nodes (e.g., similarity distances between people in social networks)
    - “soft” communities become areas in the underlying space with higher node densities
  - tell what drives signaling in networks, and what network perturbations drive it to failures (e.g., in the brain, regulatory, or metabolic networks)

- D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá,
  Curvature and Temperature of Complex Networks,
  Physical Review E, v.80, 035101(R), 2009

- F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat,
  Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces,
  INFOCOM 2010

- M. Boguñá, F. Papadopoulos, and D. Krioukov,
  Sustaining the Internet with Hyperbolic Mapping,
  Nature Communications (to appear), 2010

- M. Boguñá, D. Krioukov, and kc claffy,
  Navigability of Complex Networks,
  Nature Physics, v.5, p.74-80, 2009

- M. Boguñá and D. Krioukov,
  Navigating Ultrasmall Worlds in Ultrashort Time,
  Physical Review Letters, v.102, 058701, 2009

- M. Ángeles Serrano, D. Krioukov, and M. Boguñá,
  Self-Similarity of Complex Networks and Hidden Metric Spaces,
  Physical Review Letters, v.100, 078701, 2008