September the 19th, 2016,
Amsterdam, The Netherlands

Floriana Gargiulo & Timoteo Carletti

The hidden relationship between Zipf's law and segregation

cnrs
dépasser les frontières

UNIVERSITÉ DE NAMUR

naxys
Namur Center for Complex Systems

 timoteo.carletti@unamur.be

Segregation

Our societies are faced to social inequality and segregation dictated by race, religion, social status or incomes differences.

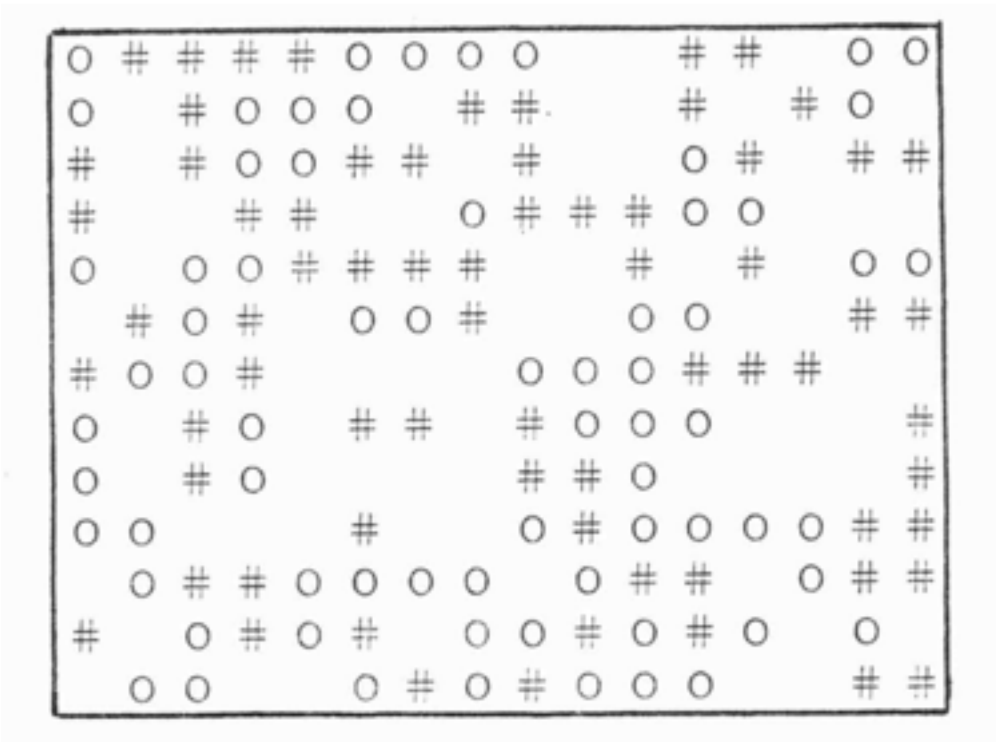
Vusimuzi settlement. Johnny Miller
www.unequalscenes.com

Palestinian loss of land, 1947-2005.

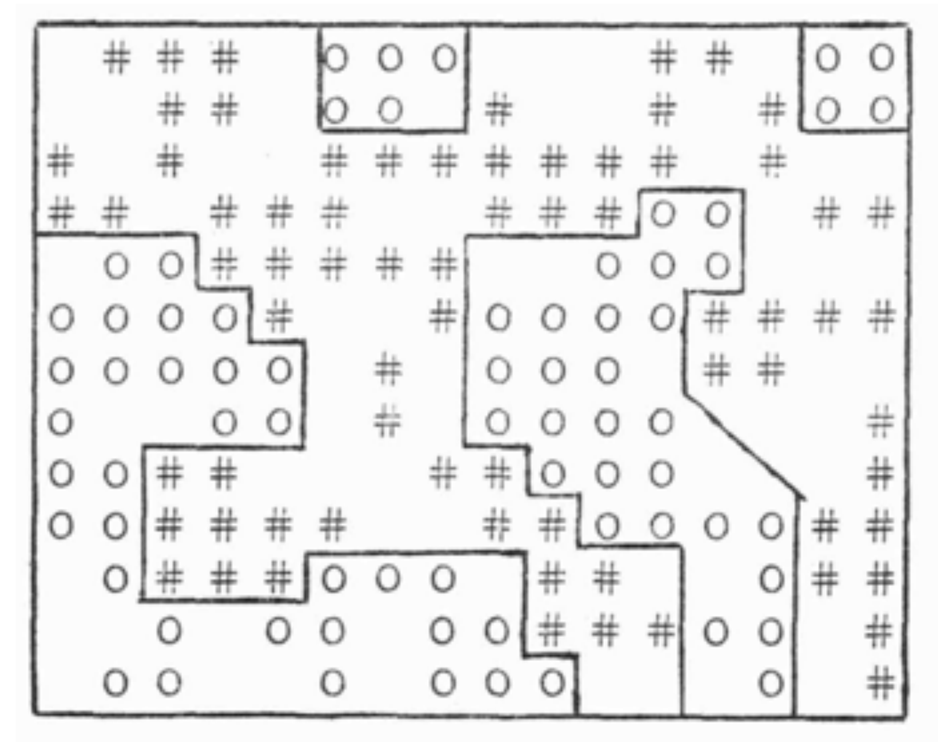
Schelling models

Schelling, T.C. American Econ. Rev. 59: 488-493. (1969)
Schelling, T.C. J. Math. Soc. 1: 143-186. (1971)

Initial spatially homogeneous population



Final clustered population



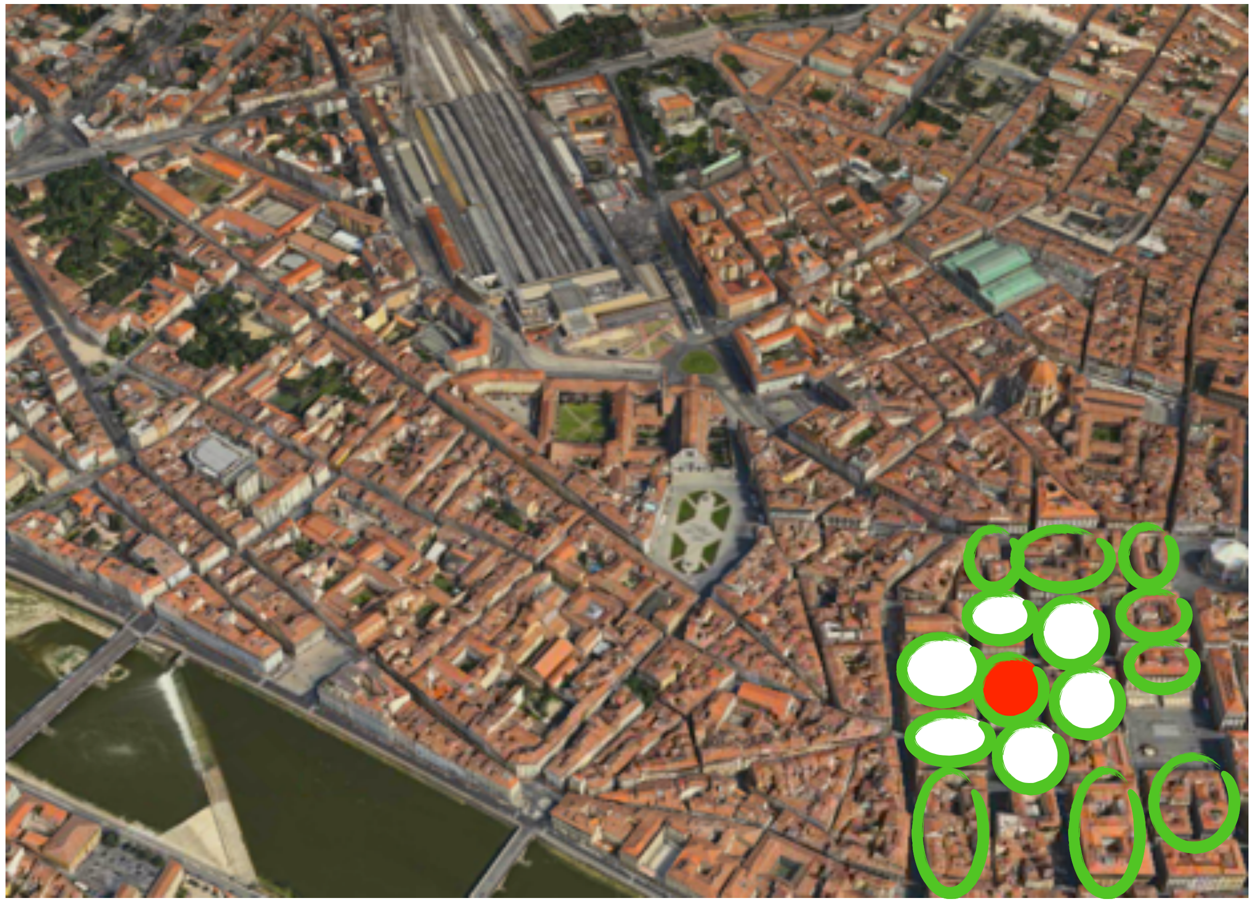
Global segregation emerges from local individual decisions, an agent decides to move if her fitness is too low.

Schelling models: main ideas

- Agents live in a chessboard, each cell can be empty or occupied by at most one agent (# or o). Agent density $\rho = N/D^2$
- Agents' fitness = fraction of similar agents in their neighbourhood;
- Agents are unhappy if their fitness is smaller than a given threshold;
- Several possibilities do exist to choose the new location (swap agents positions, fitness should strictly increase, geographical distance, network proximity, ...)

Metapopulation Schelling model

- People live in buildings forming districts, separated by well defined structures (streets, rails, rivers, hills, squares, ...);



- Well mixed behaviour inside each district.
  Interactions with “first”neighbouring districts.

Metapopulation Schelling model

- Fitness of agent type A in district i
  n_i^X # agents of type X in district i

$$\boxed{f_{1,i}^A = \frac{\sum_{j\in i} n_j^A}{\sum_{j\in i}(n_j^A + n_j^B)}}$$

- Agent of type A is unhappy in district i, if $f_{1,i}^A < 1-\epsilon$
  $\epsilon \in (0,1)$: the larger the more tolerant to diversity

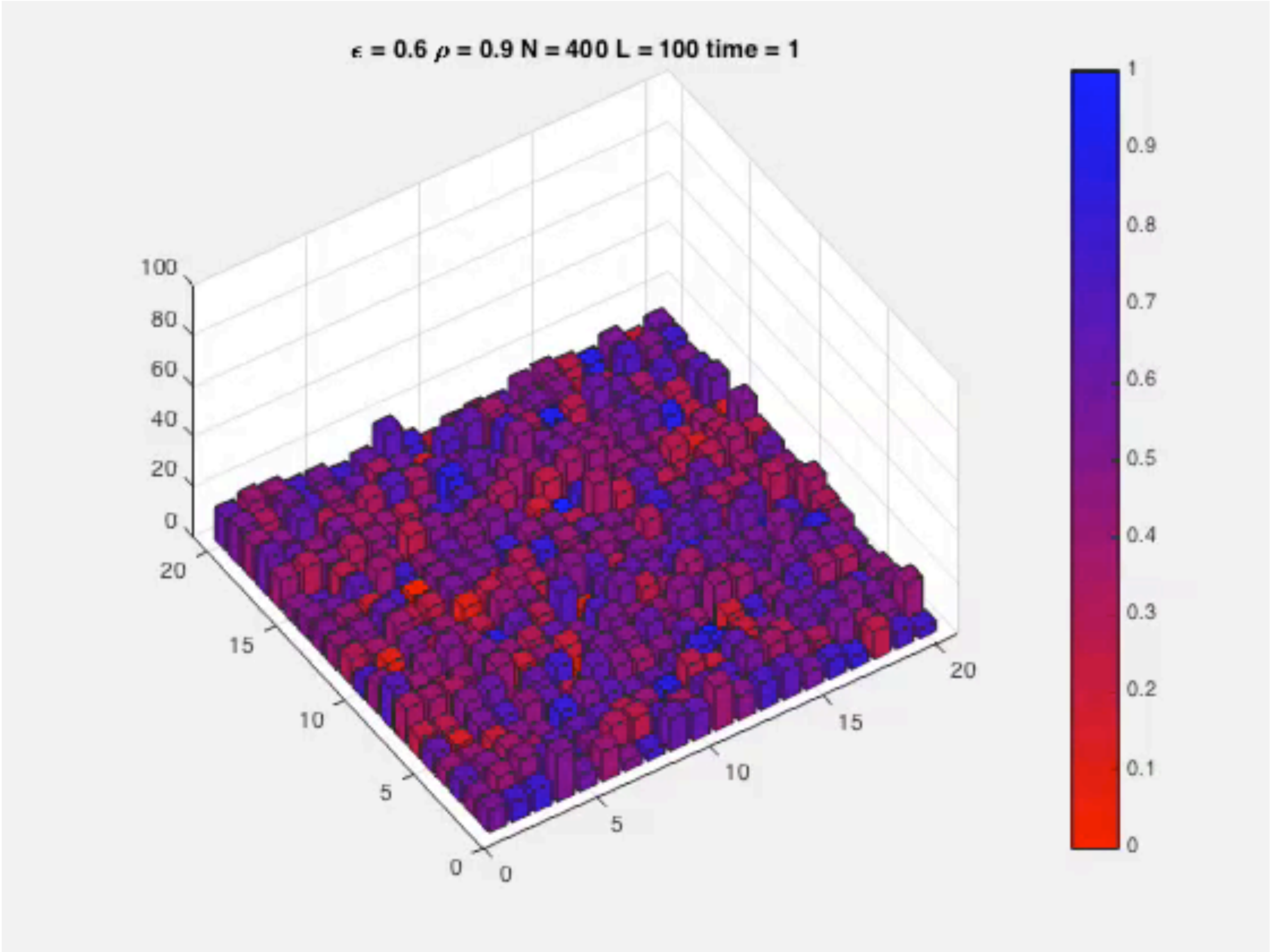
- Each district can allow for at most L agents (A or B or empty). Vacancies ρLN, A & B $(1-\rho)LN/2$, $\rho \in (0,1)$, N districts.

Metapopulation Schelling model: some remarks

- 2D square lattice but extension to networks is possible
[Y. Gandica et al., Chaos, Solitons and Fractals, 90, (2016) 46–54]

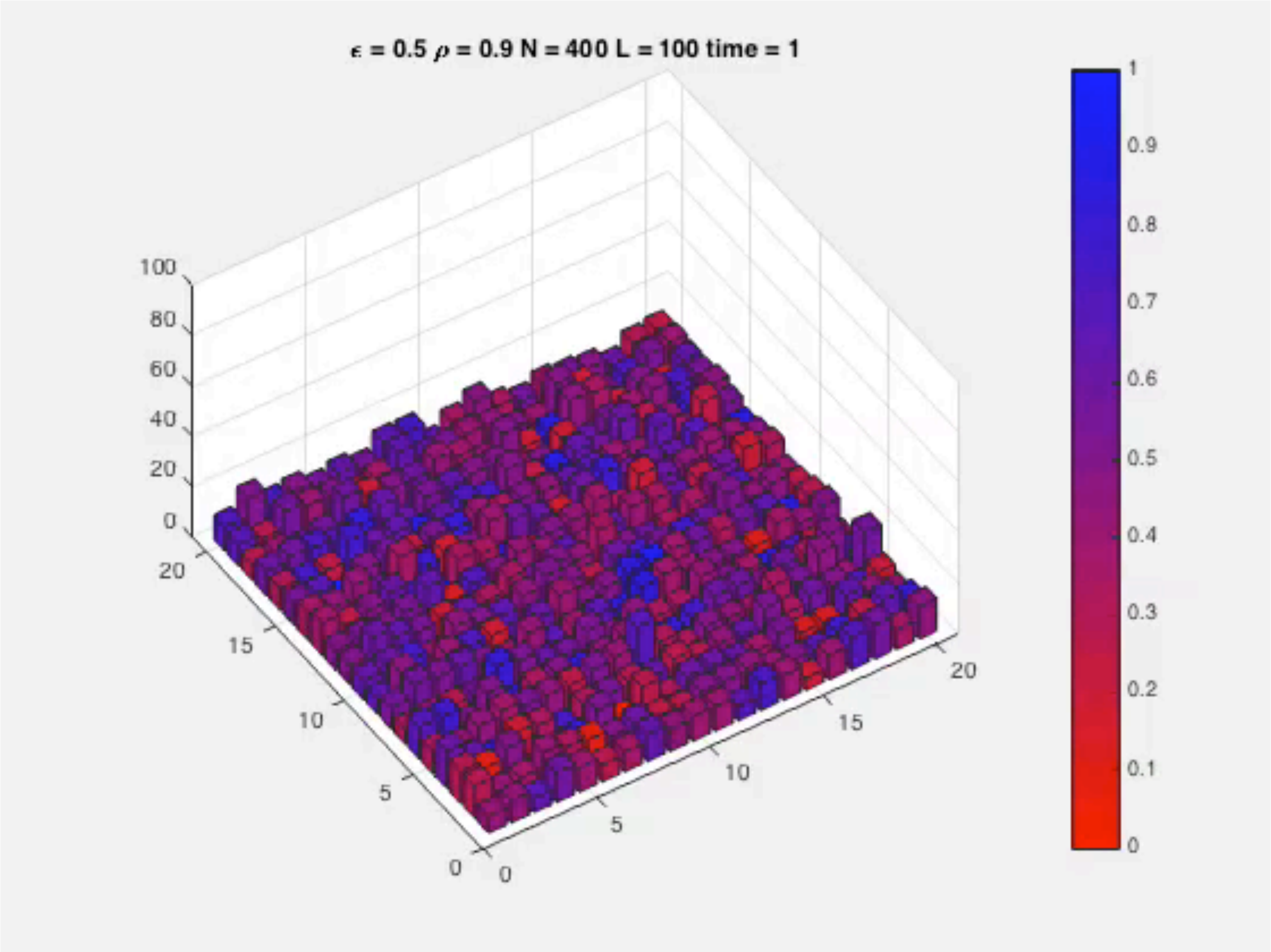
- Agents relocate themselves aiming at increasing their fitness. However partial information + hazard can induce a random choice of the final destination (if space is available there).

Some numerical results: $\epsilon = 0.6\,,\ \rho = 0.9\,,\ N = 400\,,\ L = 100$



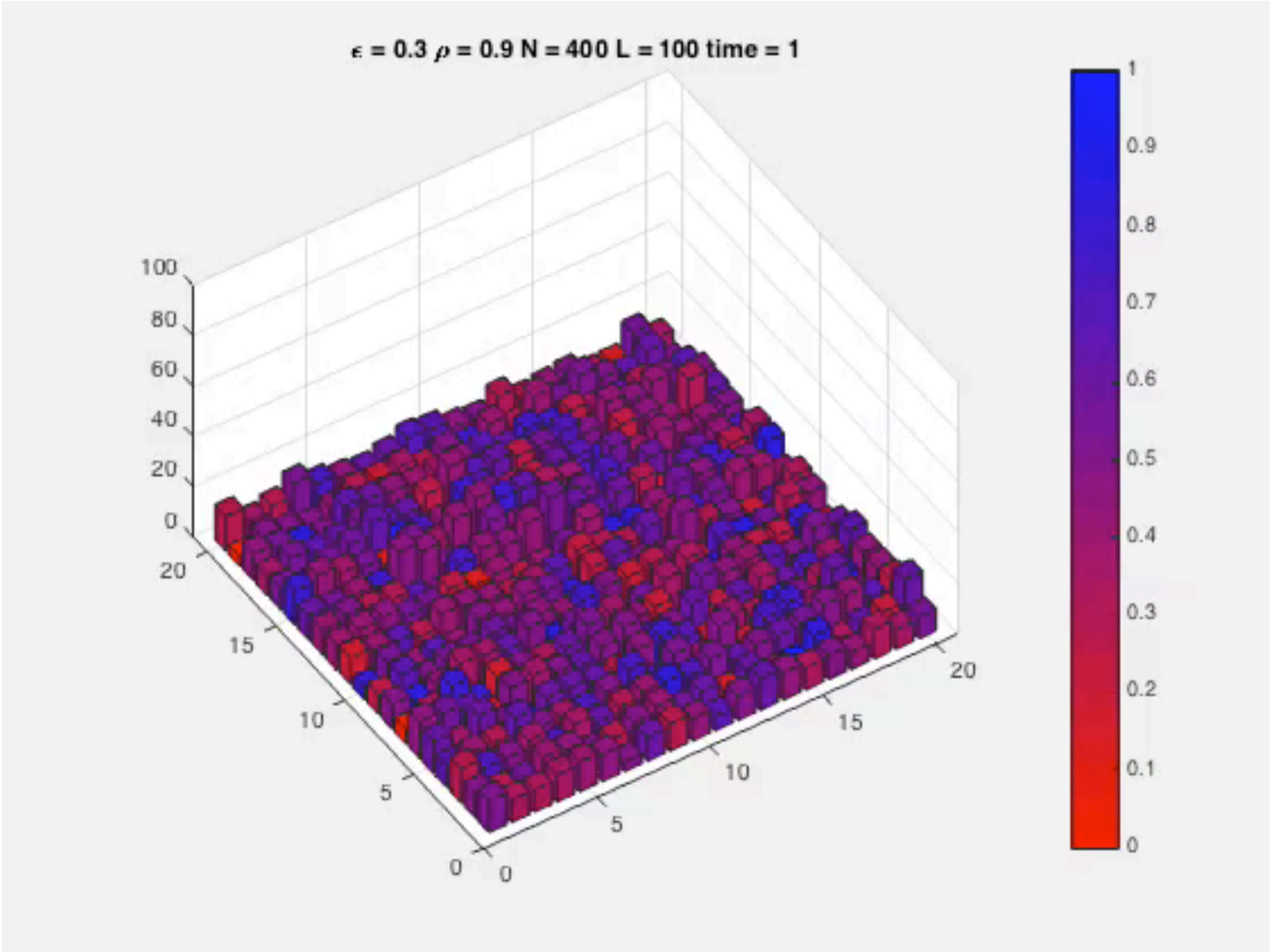
Agent of type A is unhappy if $f^A_{1,i} < 1 - \epsilon = 0.4$
hence she tolerates up to 60% of agents of type B

Some numerical results: $\epsilon = 0.5\,,\ \rho = 0.9\,,\ N = 400\,,\ L = 100$



Agent of type A is unhappy if $f_{1,i}^{A} < 1 - \epsilon = 0.5$
hence she tolerates up to 50% of agents of type B

Some numerical results: $\epsilon = 0.3\,,\ \rho = 0.9\,,\ N = 400\,,\ L = 100$



Agent of type A is unhappy if $f^A_{1,i} < 1 - \epsilon = 0.7$
hence she tolerates up to 30% of agents of type B

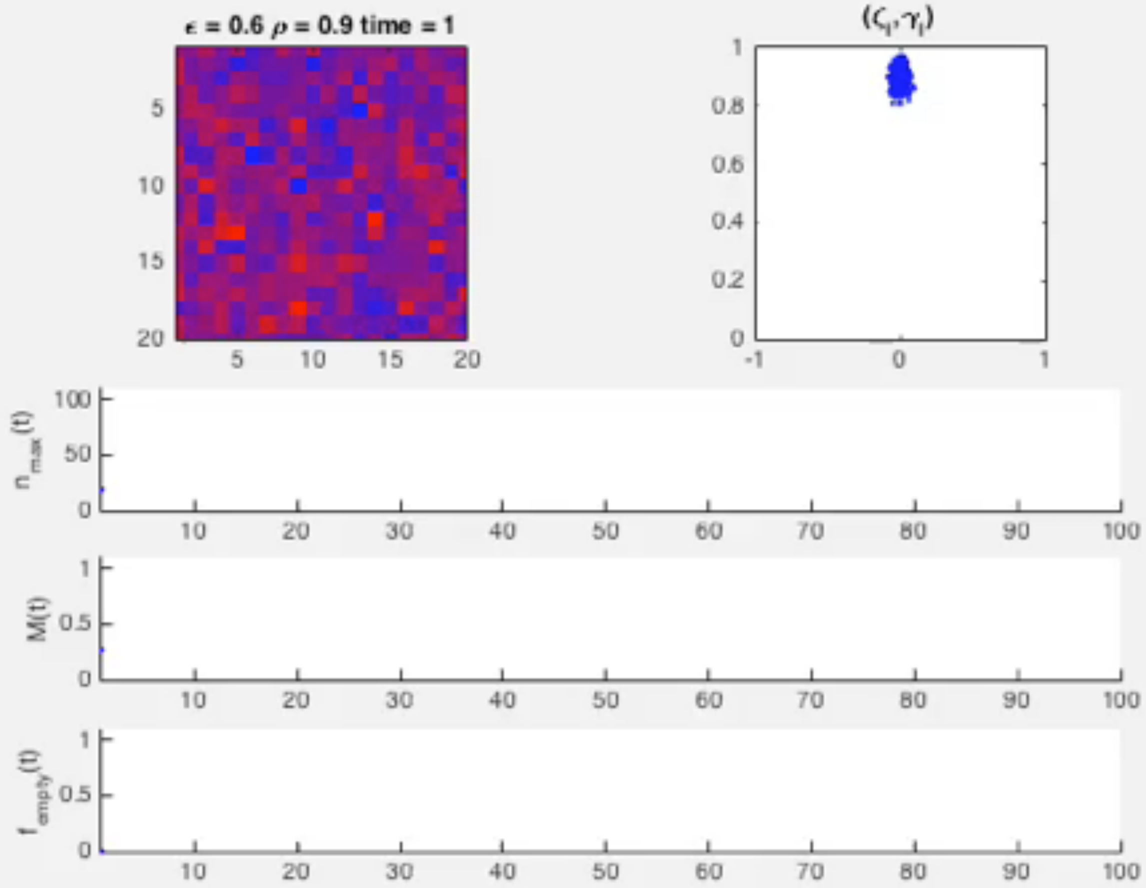
- Interface: (white) empty cells separating coloured cells.

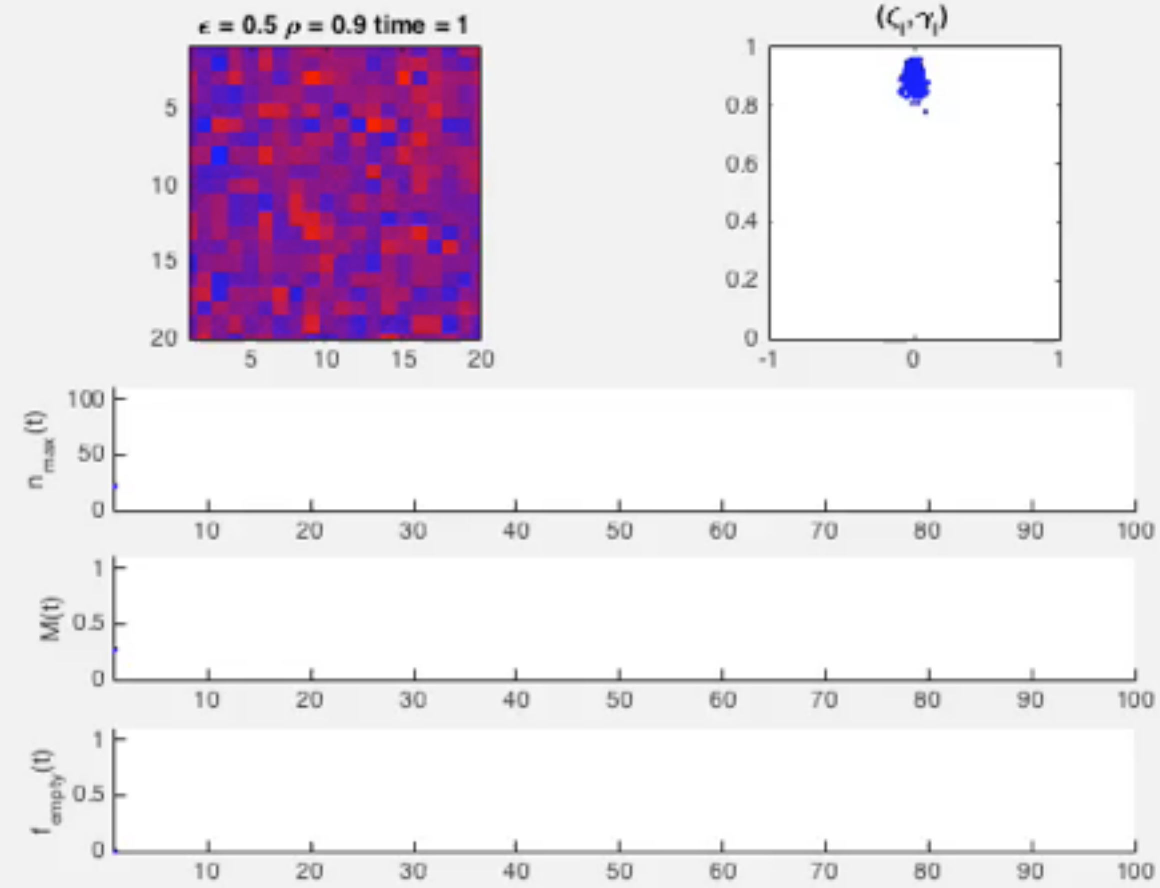
A measure of segregation

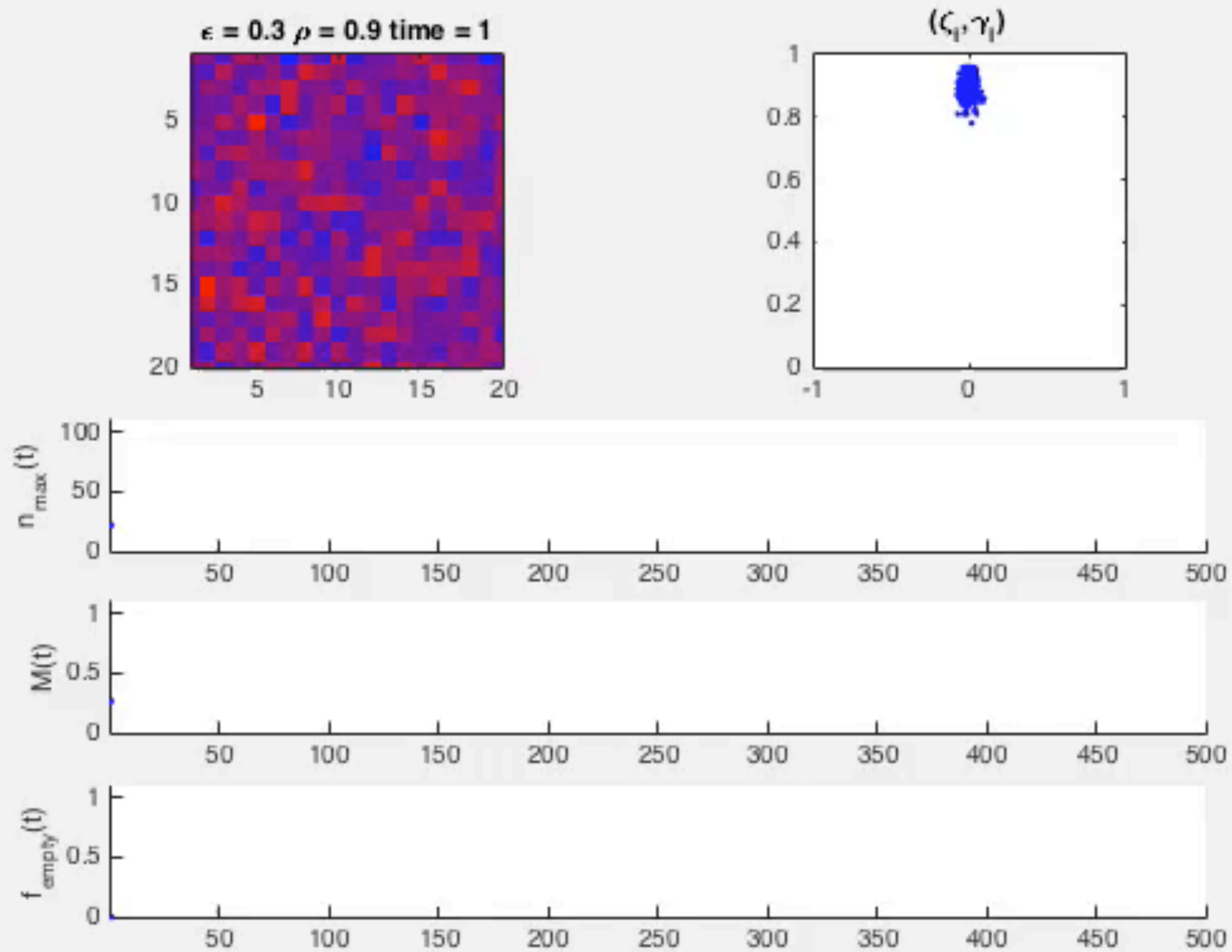
- "Magnetization" $$\langle \mu \rangle = \frac{1}{N} \sum_i \frac{|n_i^B - n_i^A|}{n_i^B + n_i^A}$$

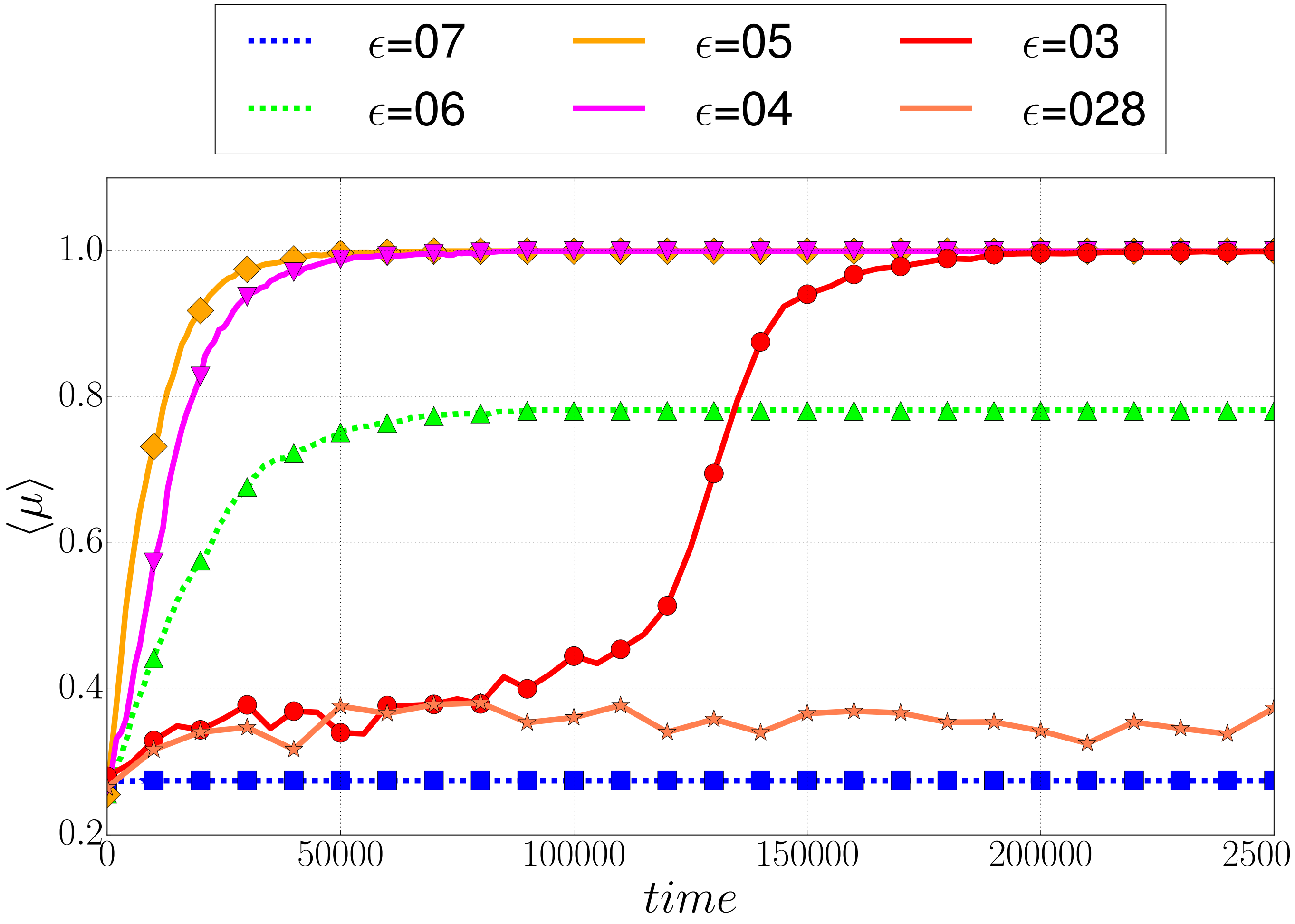
$\langle \mu \rangle \sim 0$ on average, nodes have the same number of A and B agents (well mixed population)

$\langle \mu \rangle \sim 1$ on average nodes a filled with a single type of agents (segregated population)

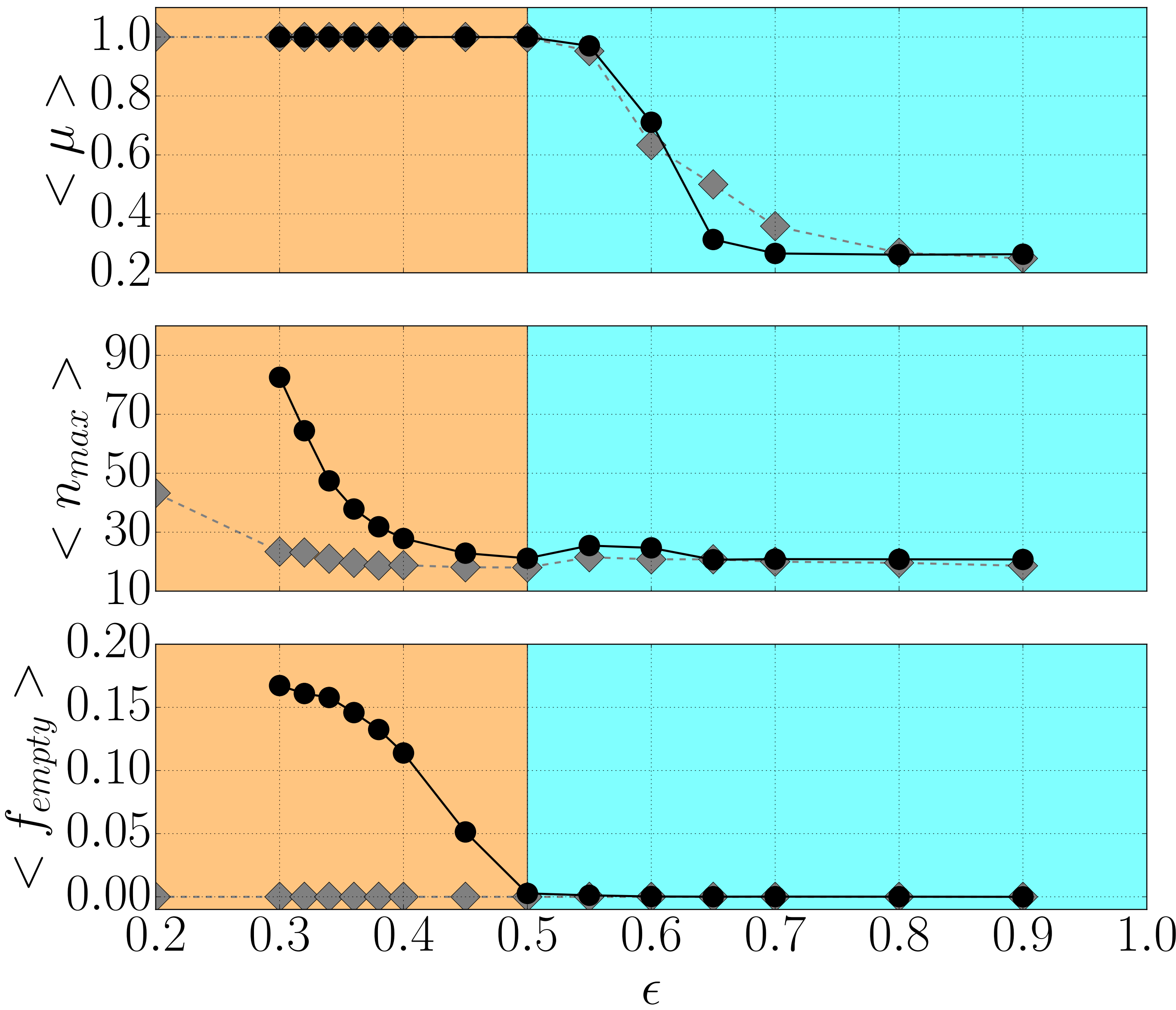

ε = 0.6 ρ = 0.9 time = 1
(ζ_l, γ_l)
n_max(t)
M(t)
f_empty(t)


ε = 0.5 ρ = 0.9 time = 1
(ζ_l, γ_l)
n_max(t)
M(t)
f_empty(t)


ε = 0.3 ρ = 0.9 time = 1
(ζ_l, γ_l)
n_max(t)
M(t)
f_empty(t)

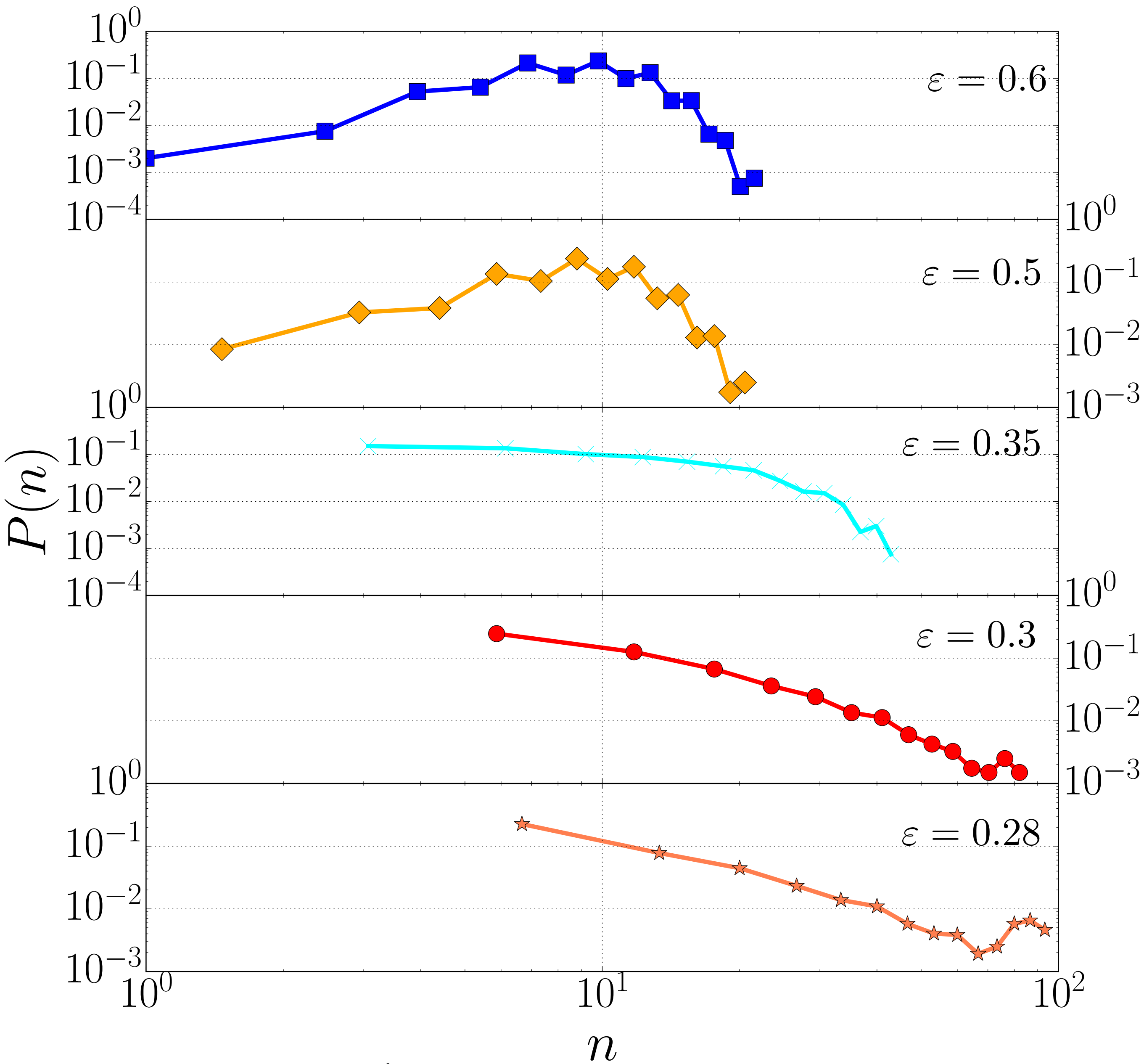


ϵ=07
ϵ=06
ϵ=05
ϵ=04
ϵ=03
ϵ=028
⟨μ⟩
time



$< \mu >$
$< n_{max} >$
$< f_{empty} >$
ϵ

Heterogeneity: Pareto law



$$n_i = n_i^A + n_i^B$$

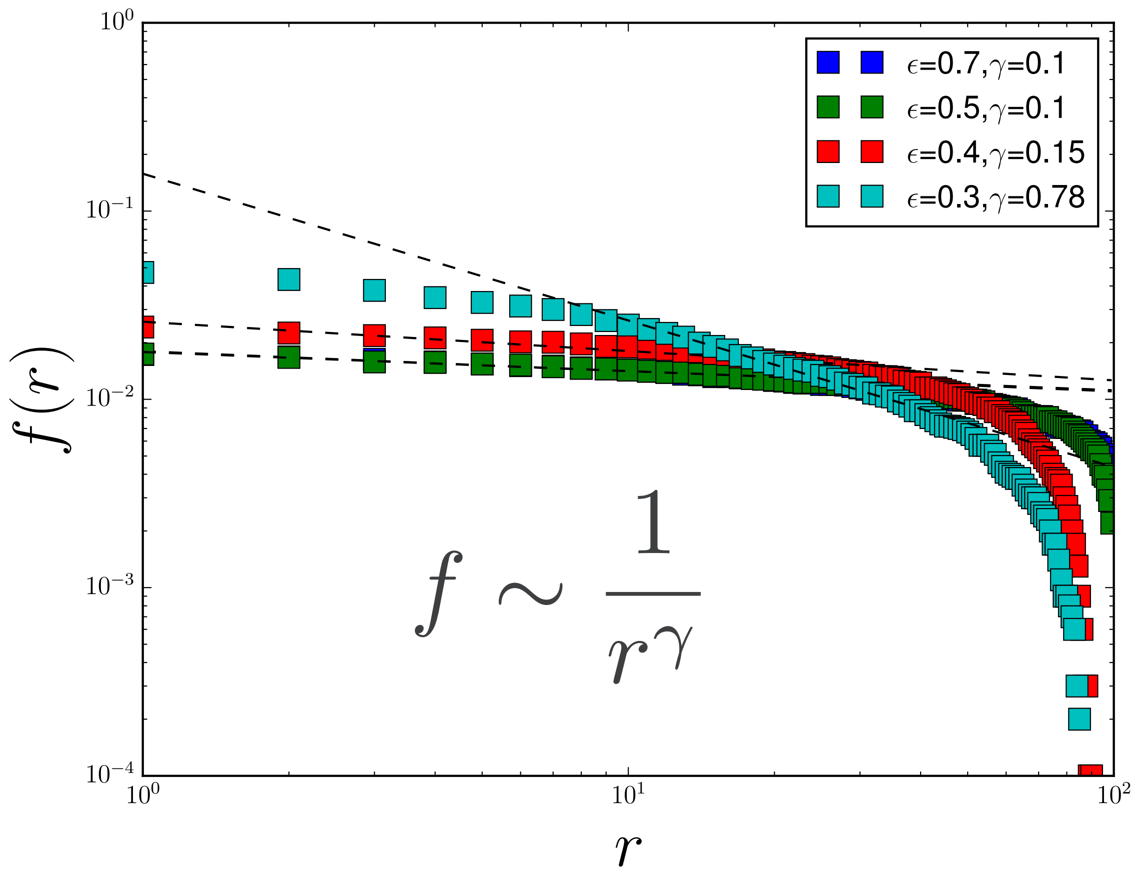
node total population

$$f_i = n_i / N$$

fraction of population in node i

$$r_i$$

population ranking of nodes



Pareto distribution in real cities

- Zipf's / Pareto laws are striking empirical phenomena observed in geographical studies.

- It is "universal": being valid across several countries and also several scales: regional, metropolitan.

> Maybe this emerging phenomenon could be a consequence of the segregation dynamics.

Segregation versus density?

- The model is too simplistic to describe the real evolution of a complex urban system (several endogenous and exogenous factors are missing)

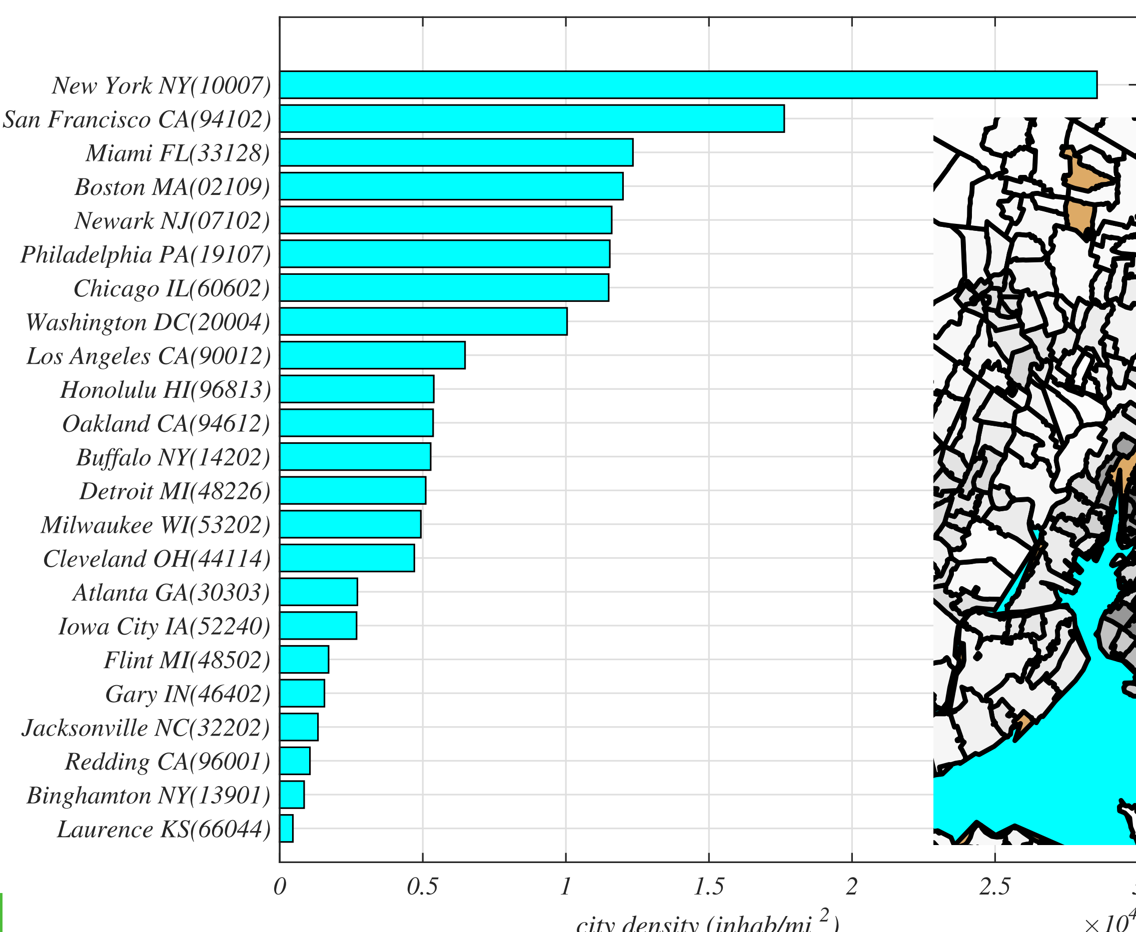
Working Questions

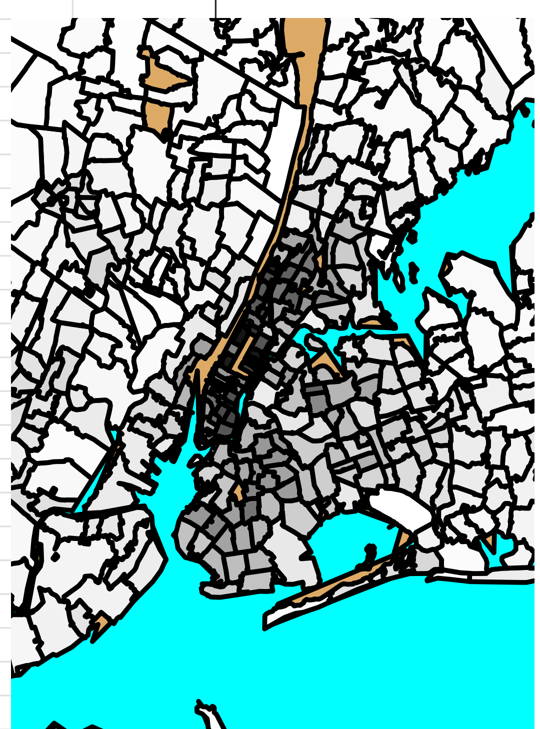
Does there exist in real urban systems, a relationship between the segregation level and population heterogeneity?

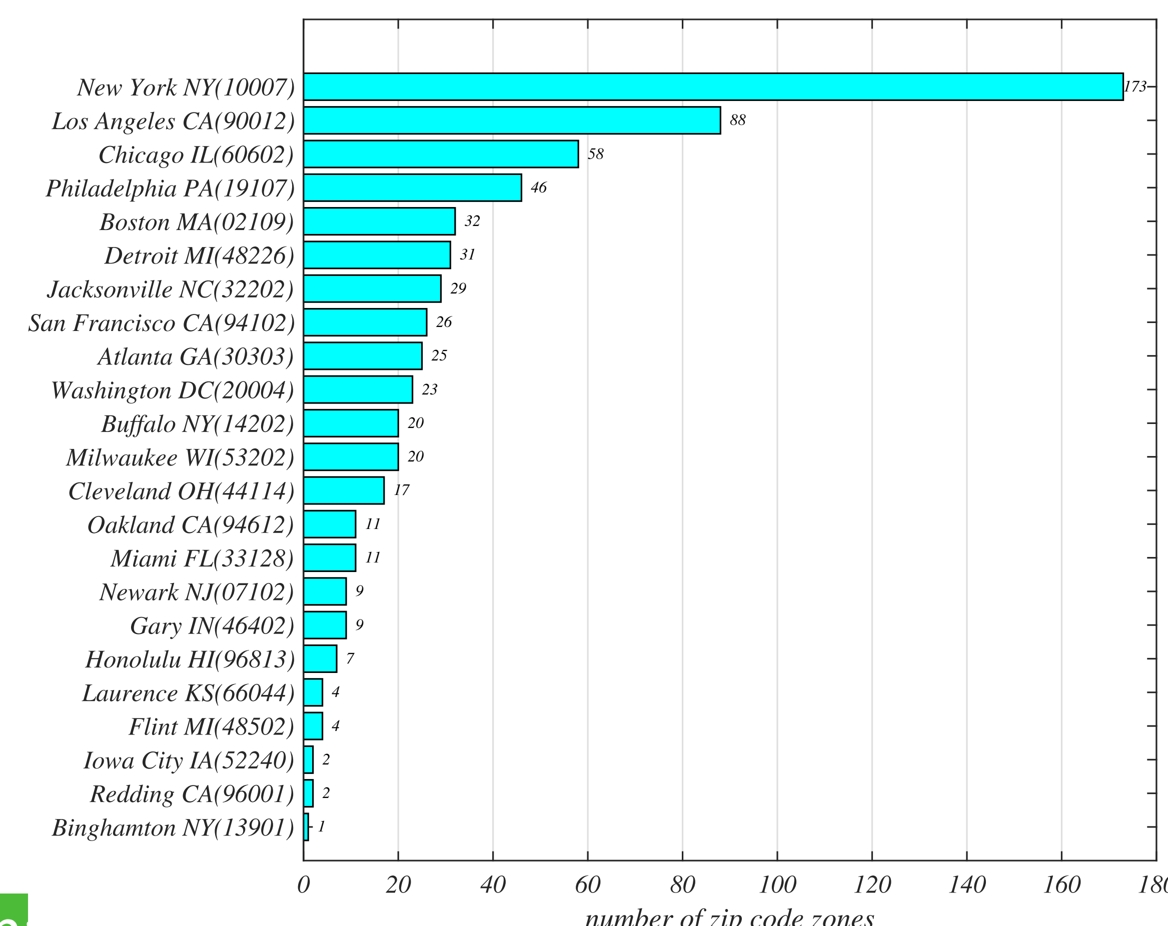
Is the segregation process, leading to a preferential racial organization inside the cities (i.e."density" preferences)?

The data

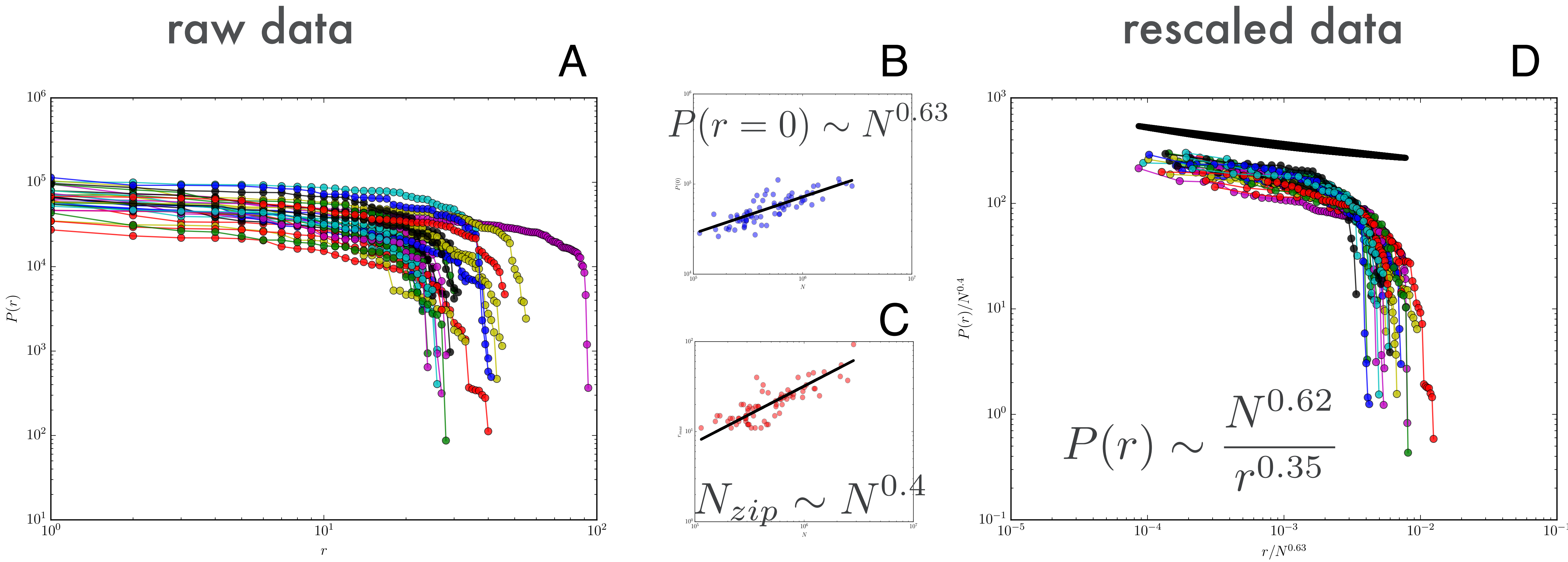
- Census data US 2010 at zip level (postal code zone)
- For each zip zone:
  - surface area (squared miles);
  - total population;
  - density (inhabitants per squared miles)
  - population per race: White, Black, Hispanic, Asiatic, native hawaiian and other Pacific islander, Native american and Alaska native, other.







Population heterogeneity across metropolitan areas



r : population ranking of zip zones

N : total population of the urban area

N_{zip} : number of internal zip codes inside the urban area

P(r=0) : maximal zip population in the urban area

Racial density preferences inside a metropolitan area

- **z-score:** $$z_{i,\alpha}^{MA} = \frac{n_{i,\alpha}^{MA} - E[X_{i,\alpha}^{MA}]}{V[X_{i,\alpha}^{MA}]}$$

MA : metropolitan area
i : zip zone
α: race (W,B,H,As,Pa,N)

$n_{i,\alpha}^{MA}$: population of race α in the i^th zip zone of MA

Null Model: binomial distribution (with constraints)

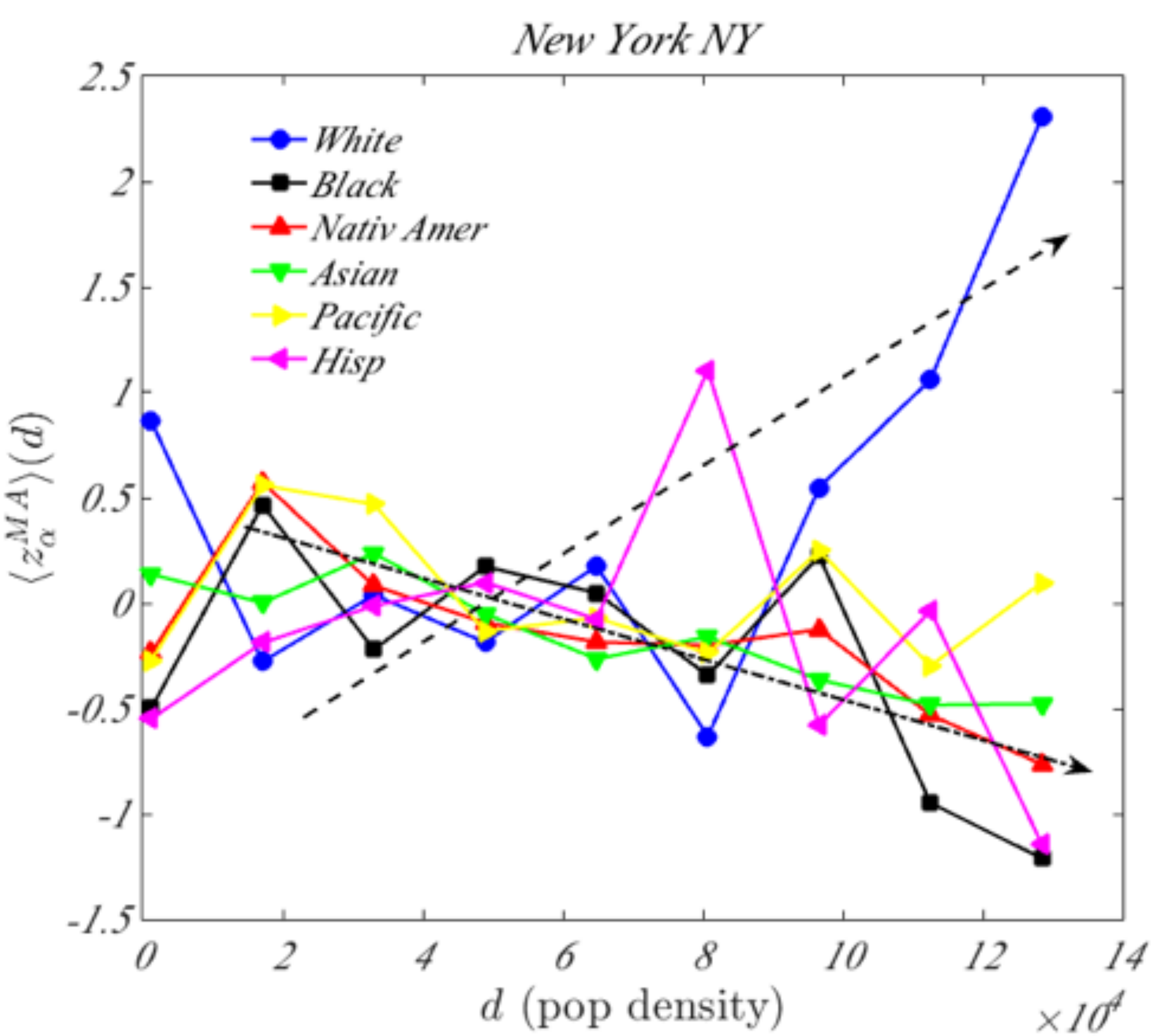
$$X_{i,\alpha}^{MA} \sim Binom(N_i^{MA}, p_\alpha)$$

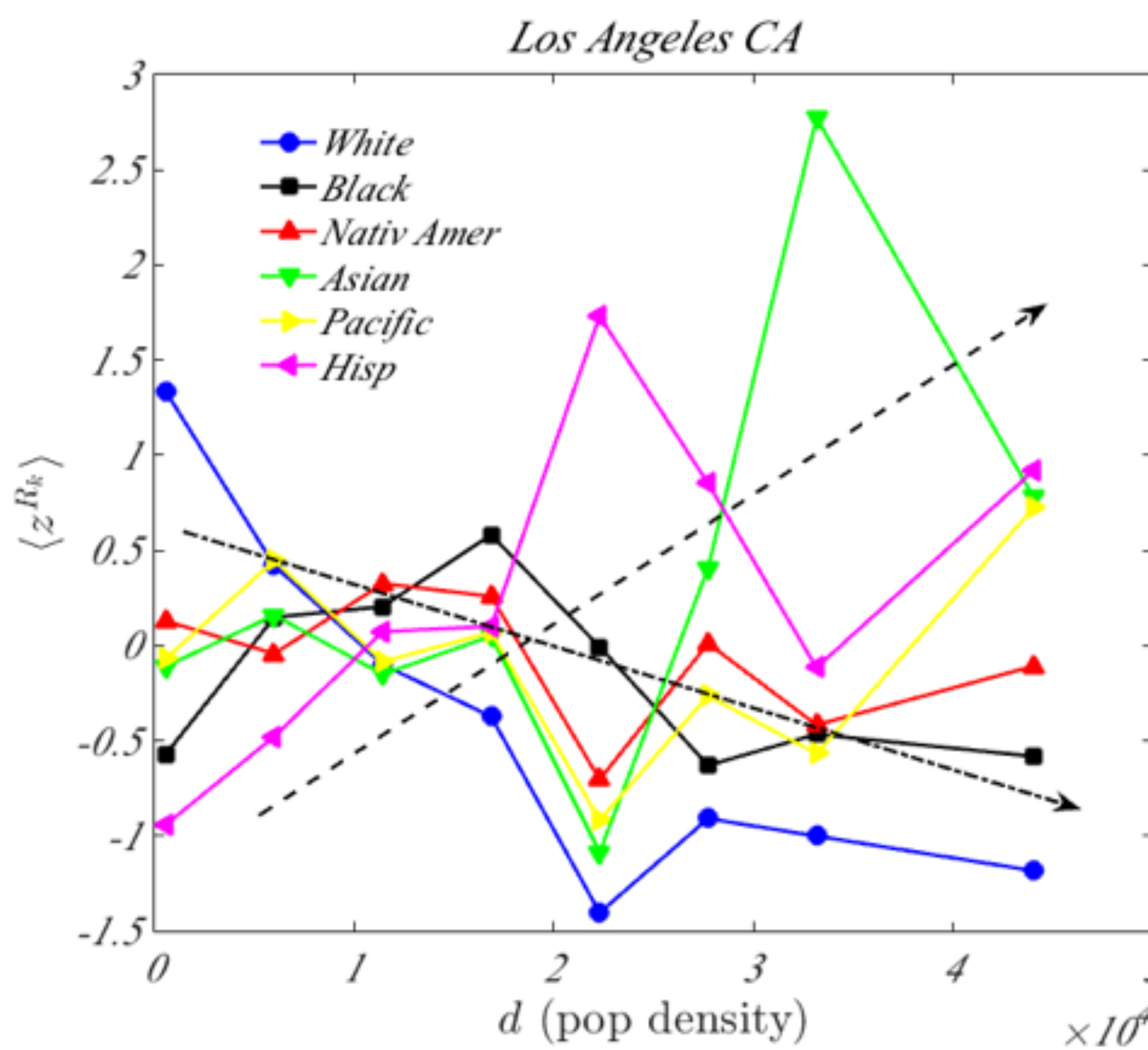
$$N_i^{MA} = \sum_\alpha n_{i,\alpha}^{MA} \qquad p_\alpha = \frac{\sum_i n_{i,\alpha}^{MA}}{\sum_{i\,\alpha} n_{i,\alpha}^{MA}}$$

Racial density preferences inside a metropolitan area

- spatially averaged z-score (on zip zones with same density d) $\langle z_\alpha^{MA} \rangle (d)$

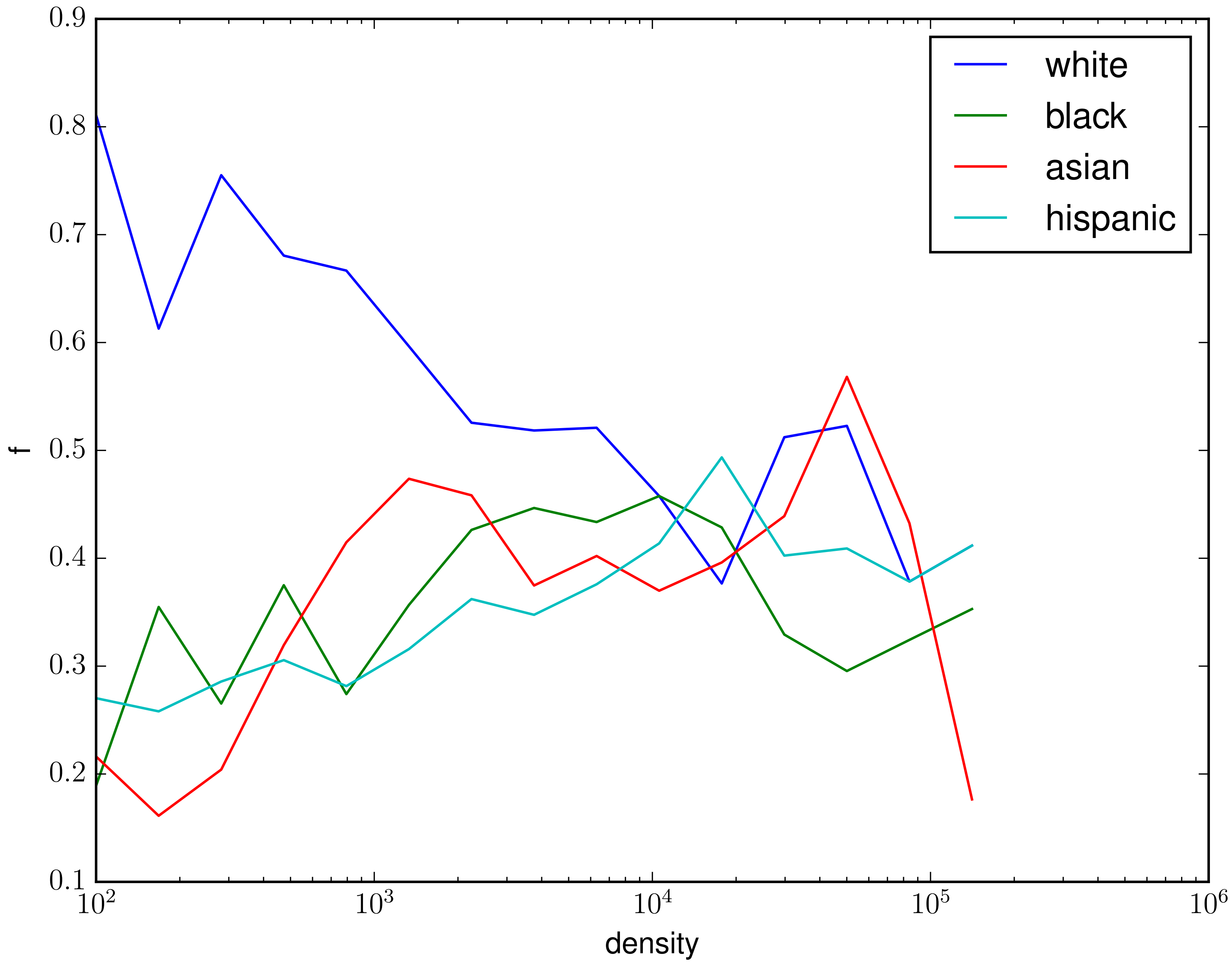
$\langle z_\alpha^{MA} \rangle (d) > 0 \; (< 0)$ race α is overrepresented (underrepresented)





Racial density preferences inside a metropolitan area

- fraction of zip zones where each race is overestimated

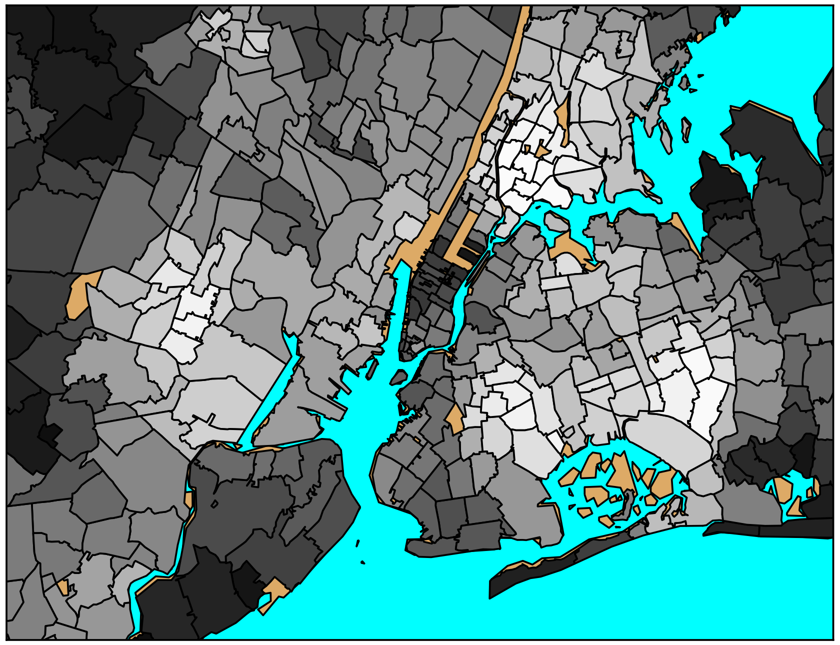


Racial density preferences inside a metropolitan area

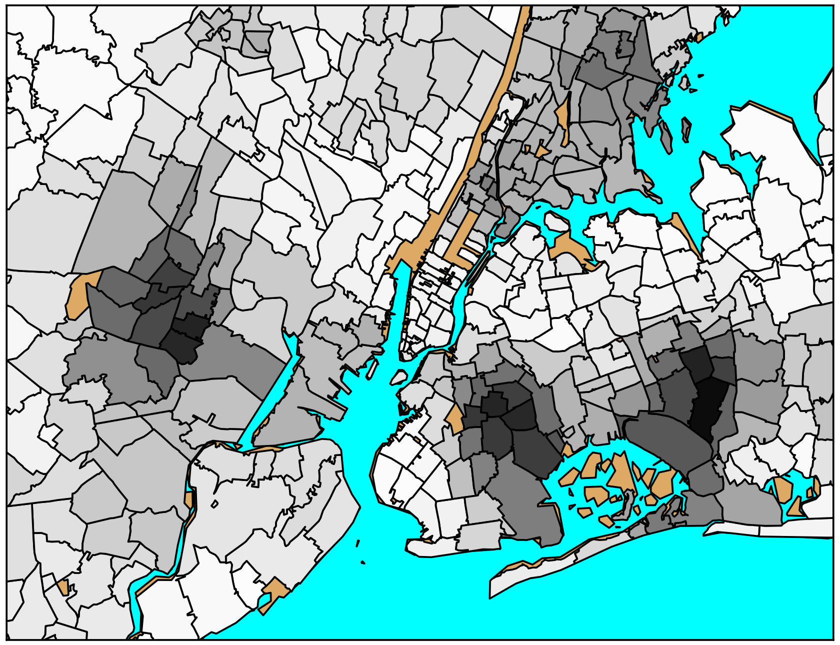
fitness metapopulation Schelling model

$$\boxed{f_{1,i}^{A} = \frac{\sum_{j\in i} n_j^A}{\sum_{j\in i}(n_j^A + n_j^B)}} \Rightarrow \boxed{f_{1,i}^{\alpha} = \frac{\sum_{j\in i} n_j^\alpha}{\sum_\alpha \sum_{j\in i} n_j^\alpha}}$$

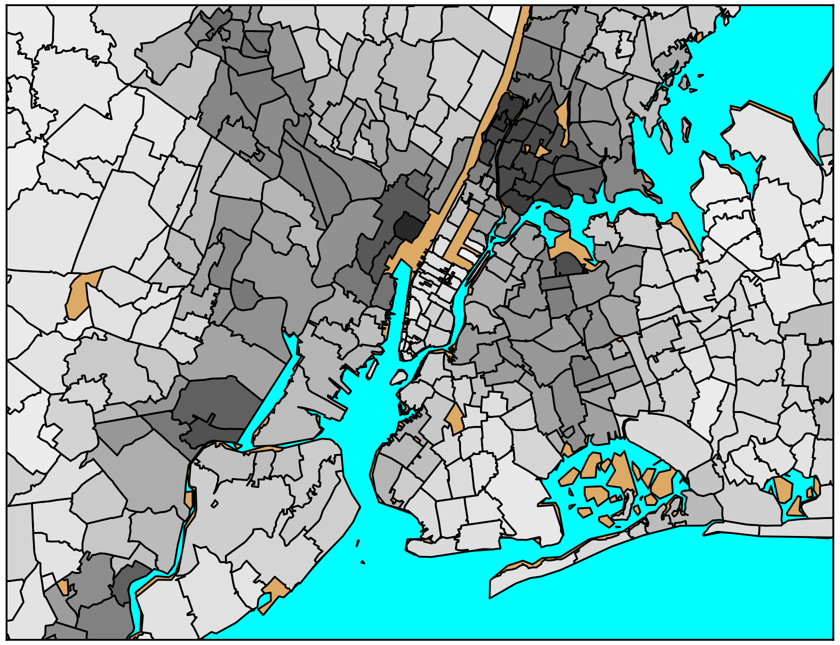
New York, NY



White



Black

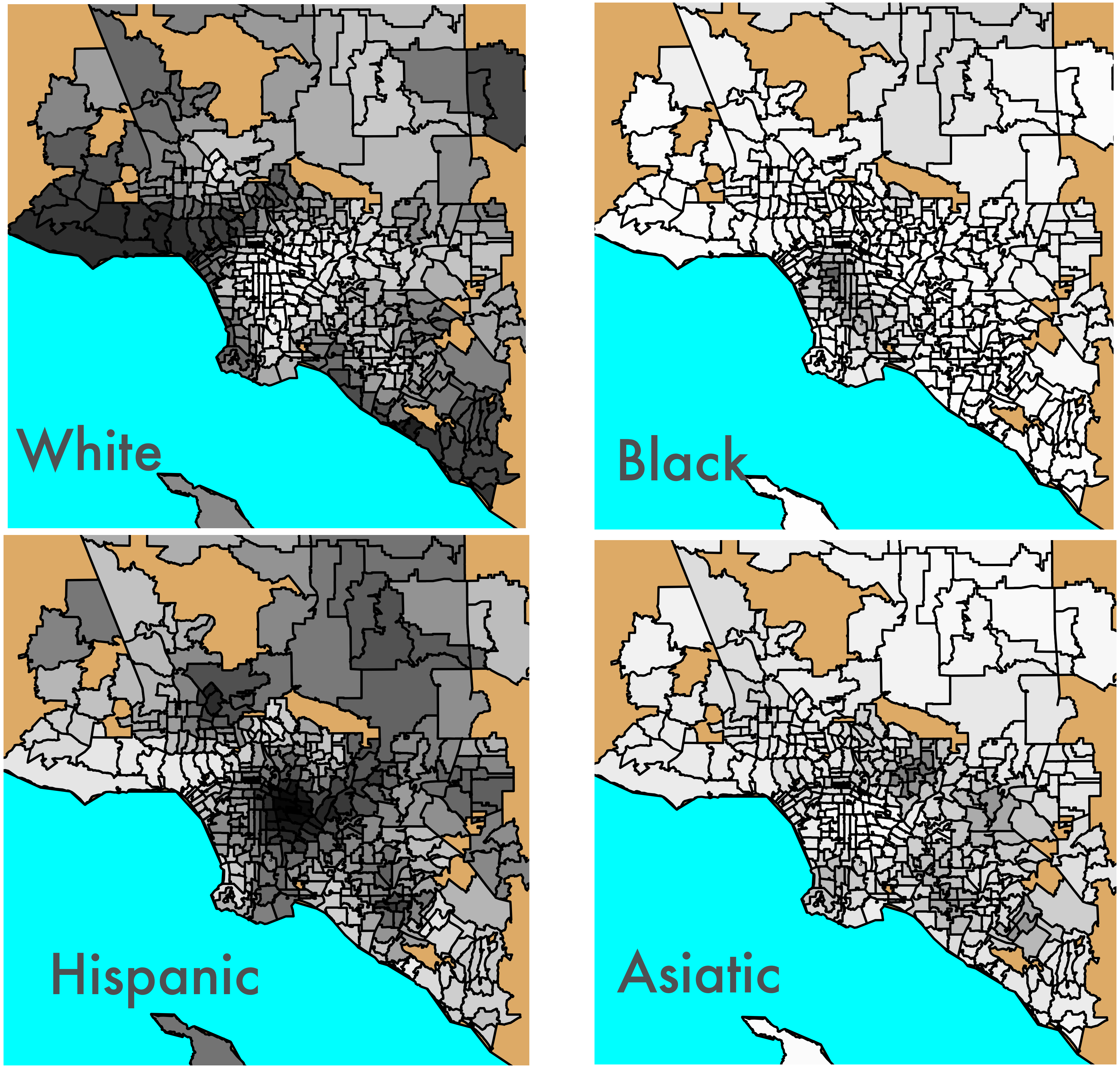


Hispanic

dark / light = high/low value of fitness

Racial density preferences inside a metropolitan area

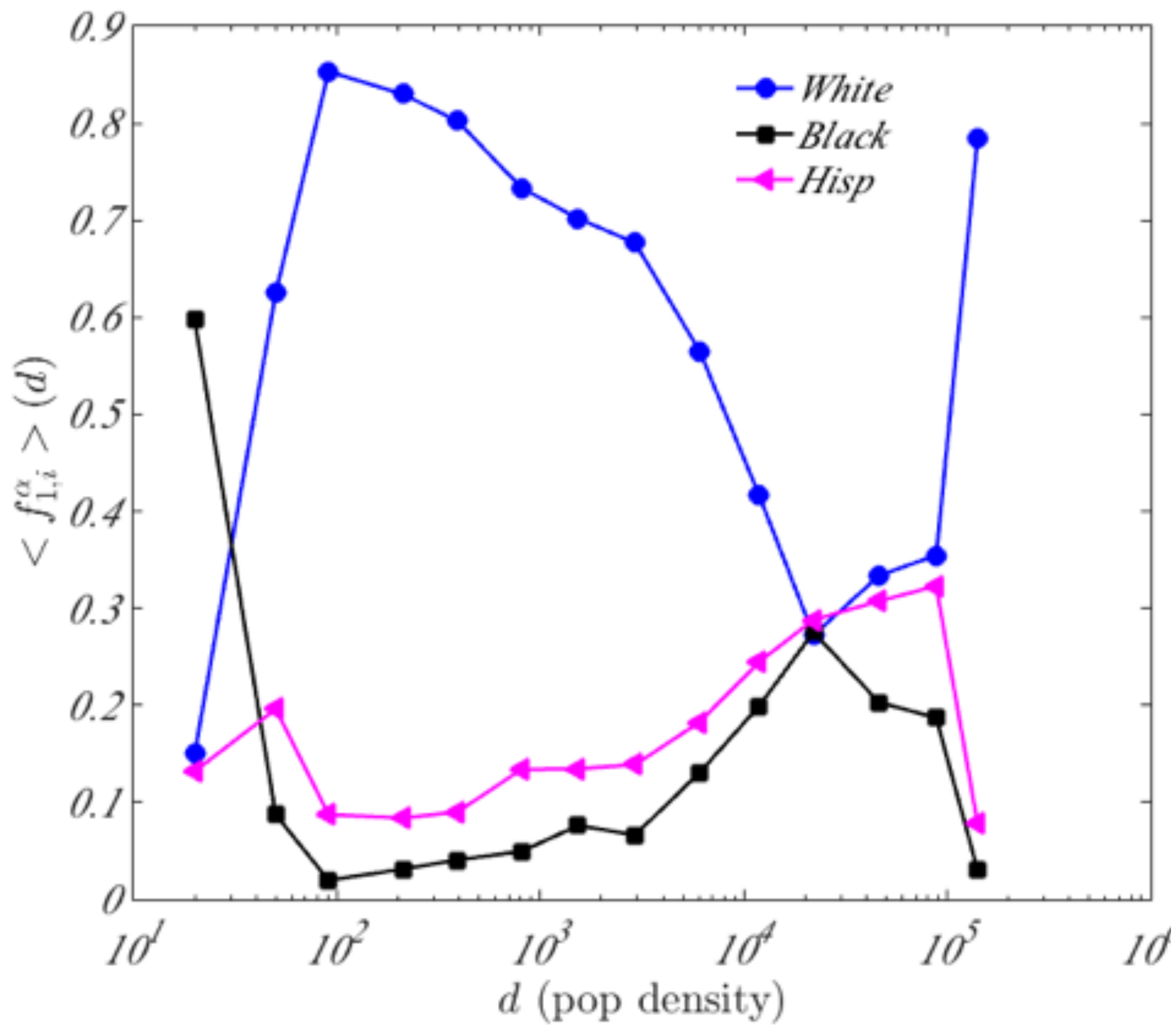
Los Angeles, CA



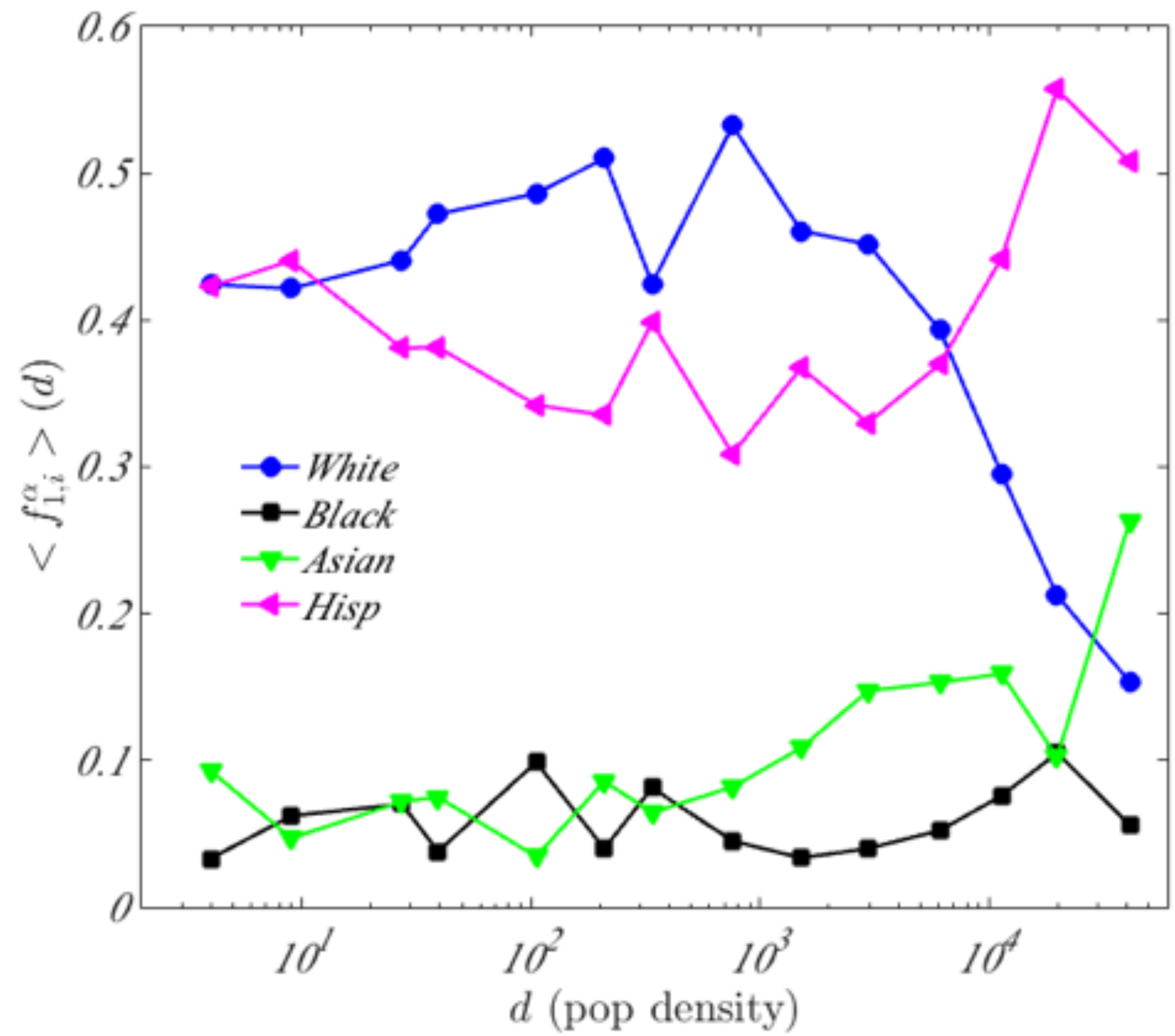
Racial density preferences inside the metropolitan area

spatially averaged fitness $\langle f^{\alpha}_{1,i} \rangle (d)$
(on zip zones with same density d)

New York, NY



Los Angeles, CA



Conclusions

- Improved Schelling model (metapopulation and topology). Analytical results available;
- Emergent Pareto laws: heterogeneity vs density;
- Study of several US metropolitan areas;
- Heterogeneity and racial density preferences.

September the 19th, 2016,
Amsterdam, The Netherlands

Floriana Gargiulo & Timoteo Carletti

The hidden relationship between Zipf's law and segregation

cnrs dépasser les frontières

UNIVERSITÉ DE NAMUR

naxys Namur Center for Complex Systems