# Voter model, opinion diffusion and mobility networks


*MAXI SAN MIGUEL*

JUAN FERNANDEZ-GRACIA

KRZYSZTOF SUCHECKI

JOSE J. RAMASCO

VICTOR M. EGUILUZ

IFISC

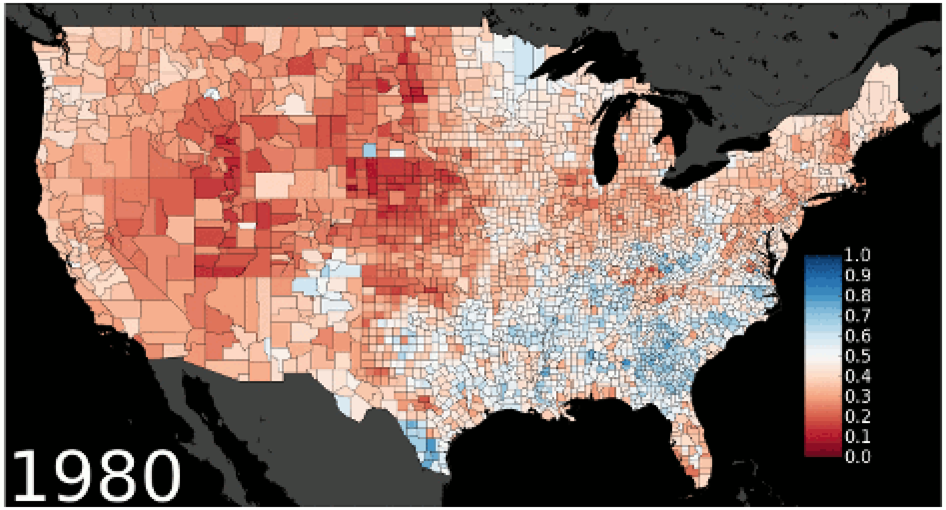
CSIC

Universitat de les Illes Balears

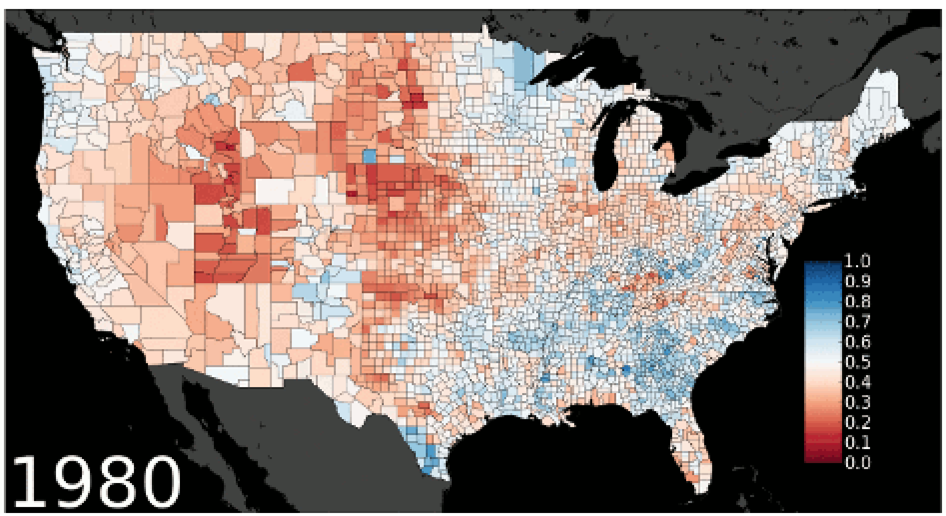
## Raw data

## US presidential elections 1980-2012

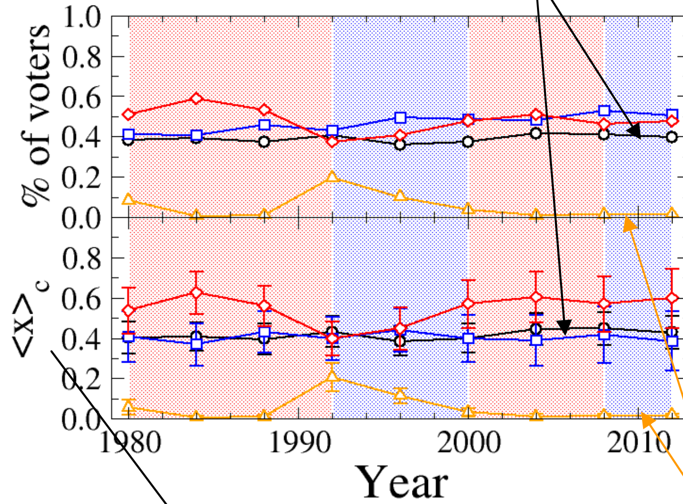
Evolution of democrat shares




Evolution of republican shares




Global percentage of votes for each party and turnout




Average per county percentage of votes for each party and turnout

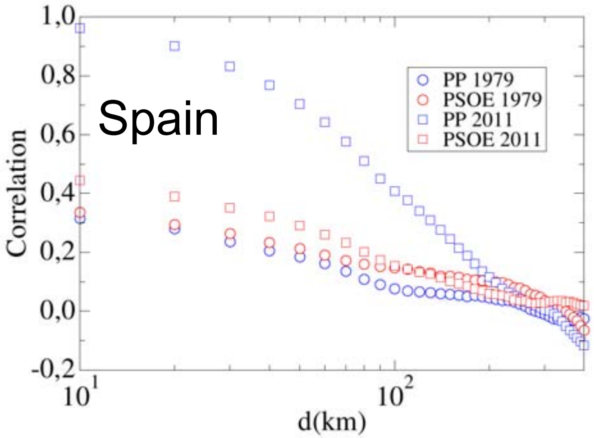
Other parties

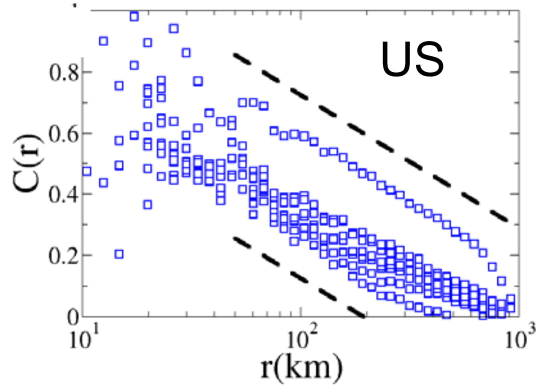
Basically two option system

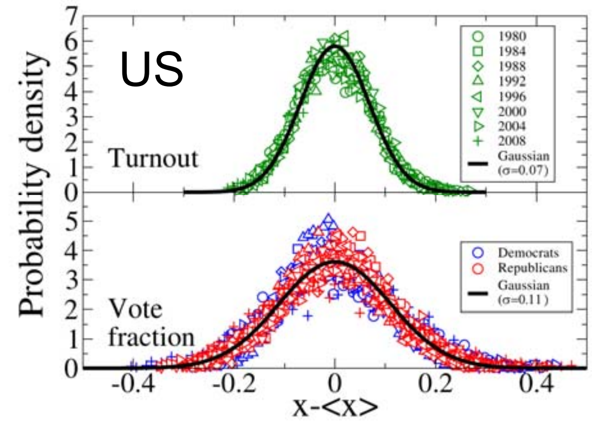
Blue: Democrat

Red: Republican

# Irrespective of the winner!



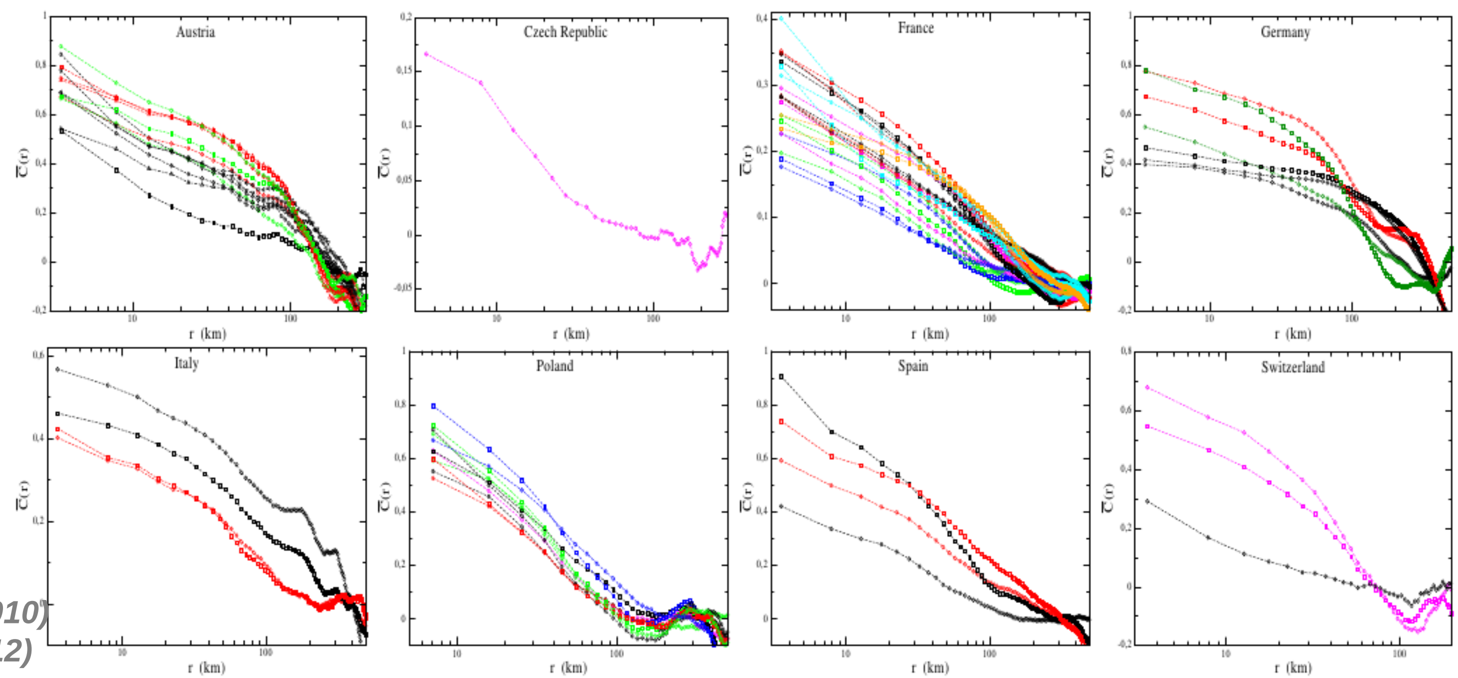




**IBM**

**Diffusive model?**

**SPATIAL CORRELATIONS**

~log decay




*C. Borghesi et al.*
*Eur. Phys. J. B 75, 395-404 (2010)*
*PLoS ONE 7(5):e36289,05 (2012)*

Ingredients of a social influence model:

a) **Interaction mechanism:** Imitation as basic manifestation of social influence.

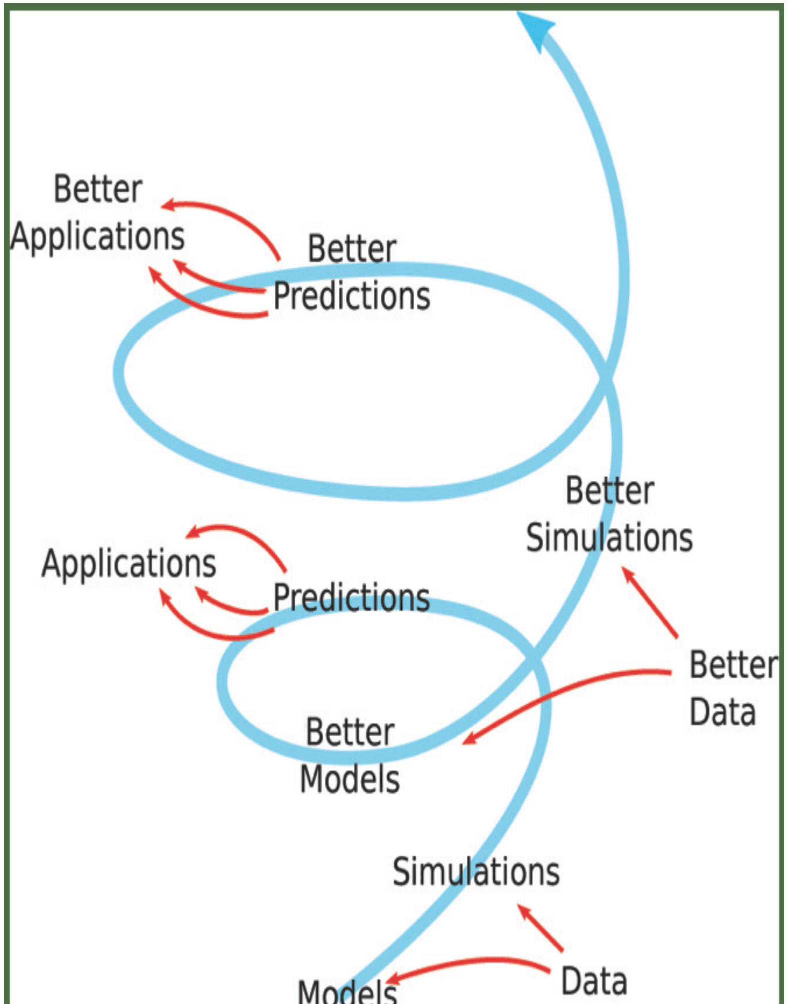
THEORY

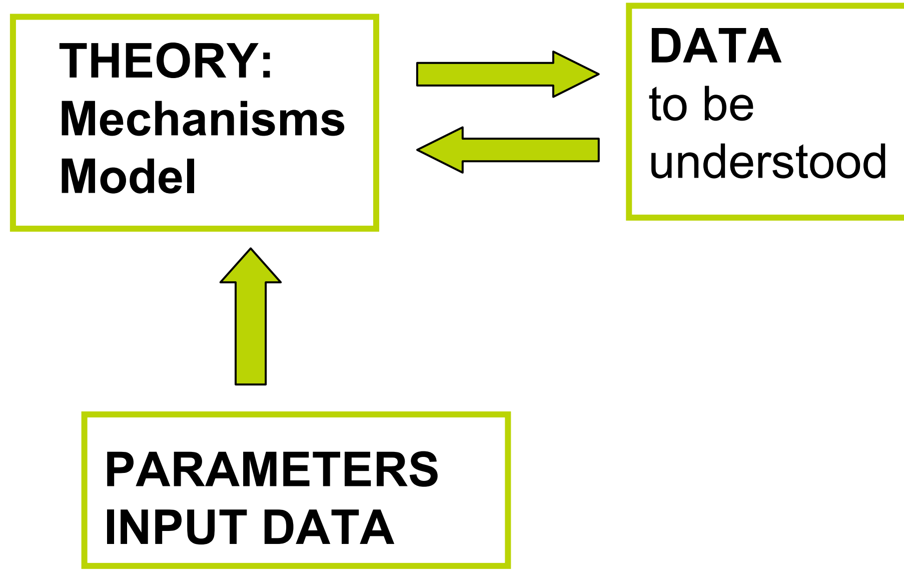
MODELLING

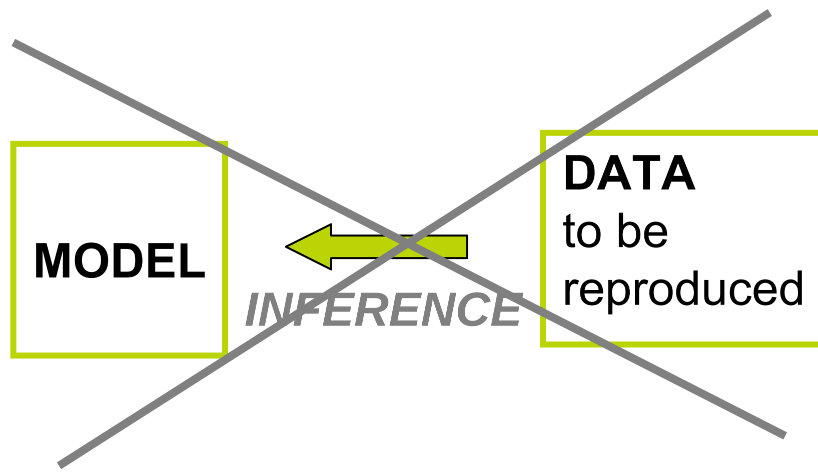
b) **Social context:** Set of all possible interactions of an individual with any other peer. We model it as a network of interactions from census data for population and mobility.

INPUT DATA

# MODELS and DATA



Ingredients of a social influence model:

a) Interaction mechanism: Imitation as basic manifestation of social influence.

**THEORY**

**MODELLING**

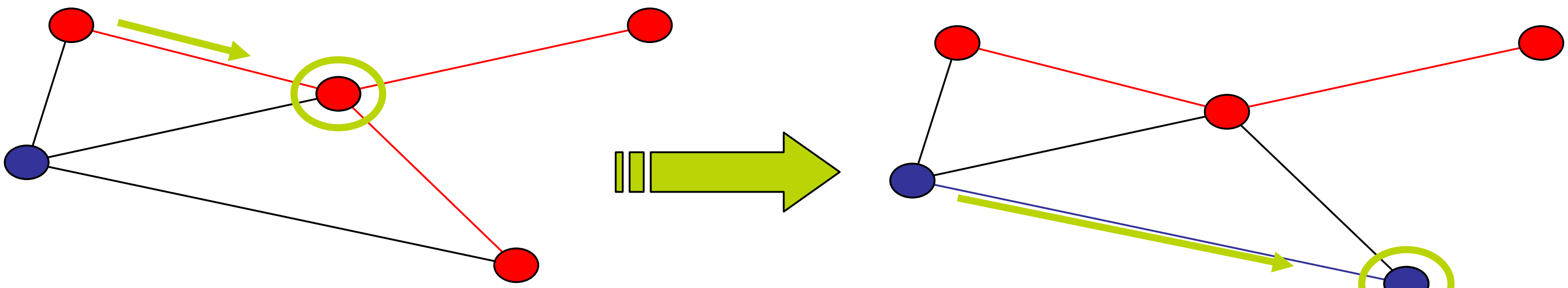
***What can we learn from models of simple social behavior?***

***Reaching agreement by imitation?***


Herding Behavior

**Voter model** (Ann Prob. 1975)

Two options:



Interaction: copy the state of one of your neighbors at random

## Voter Model as a diffusion model

$\sigma_i = \pm 1$ → Value of the state at site i.

$k_i$ degree

$a_{ij}$ adjacency

$$\langle \dot{\sigma}_i \rangle = \frac{1}{k_i} \sum_j L_{ij} \langle \sigma_j \rangle$$

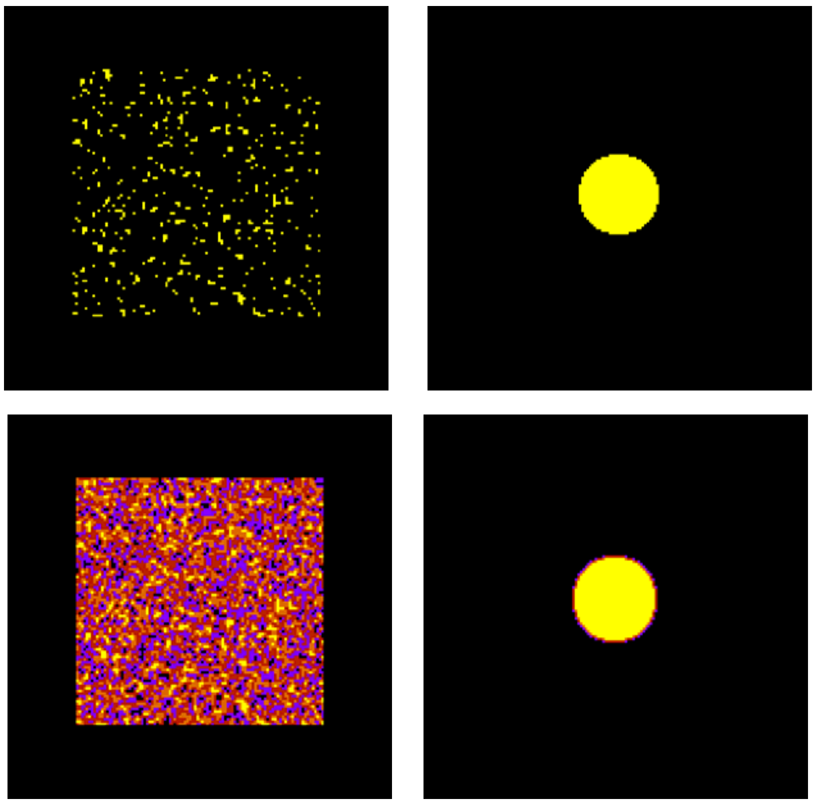
$\sum_j L_{ij} = 0$ Laplacian matrix

$$L_{ij} = a_{ij} - k_i \delta_{ij}$$

$$\frac{d}{dt} \frac{\sum_i k_i \langle \sigma_i \rangle}{\sum_i k_i} = 0$$

Conserved quantity:
Ensemble average weighted magnetization

*Klemm et al, Sci. Rep. 2, 292 (2012)*



For a noisy diffussive model, spatial correlations decay logarithmically in 2d.

$$\dot{\psi} = \nabla^2 \psi + \eta$$

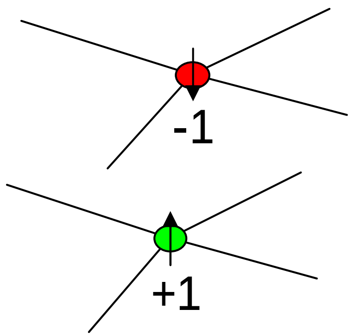
$$\Rightarrow \langle \psi(\vec{r}) \psi(\vec{r'}) \rangle \xrightarrow{d=2} a - b \ln(|\vec{r} - \vec{r'}|)$$

$\eta$ → Uncorrelated gaussian white noise.

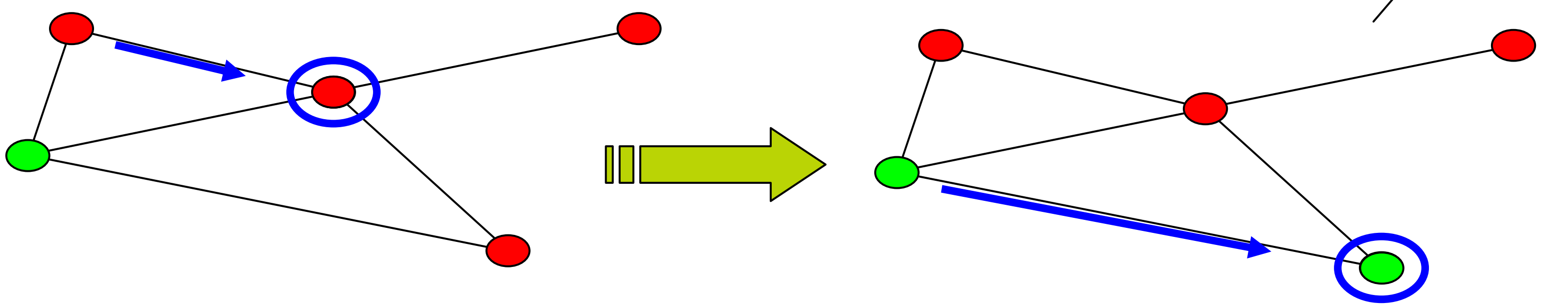
Spatial correlation and average over realizations do not commute

# Voter Model: IMITATION DYNAMICS

Ann. Probability (1975)

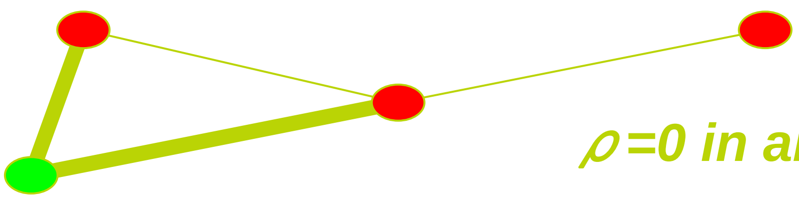



- "Voters" located in the nodes of a network have "opinions" $\sigma_i$=1 or $\sigma_i$=-1.
- ***A randomly chosen voter takes the opinion of one of its neighbors chosen at random***



***Question: When and how one of the two absorbing states (agreement or consensus) is reached by imitation dynamics?***

Order Parameter: Average interface density



$$\rho = \frac{1}{2N\langle k\rangle}\left(\sum_{i=1}^{N}\sum(1-\sigma_i\sigma_j)\right)$$

***$\rho$=0 in absorbing state***

Interface: a link connecting nodes with different states.

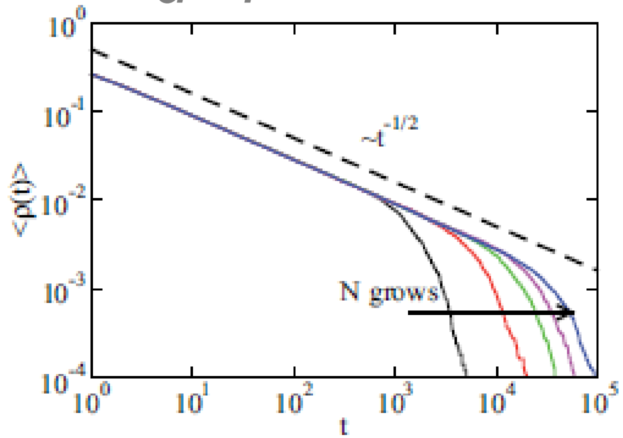
# Voter Model in regular networks

$$<\rho>\sim\begin{cases} t^{-1/2}, & d=1 \\ (\ln t)^{-1}, & d=2 \\ \xi - b t^{-d/2}, & d>2 \end{cases} \qquad \tau\sim\begin{cases} N^2, & d=1, \text{ time to reach absorbing state} \\ N\ln N, & d=2, \text{ time to reach absorbing state} \\ N, & d>2, \text{ survival time of metastable state} \end{cases}$$

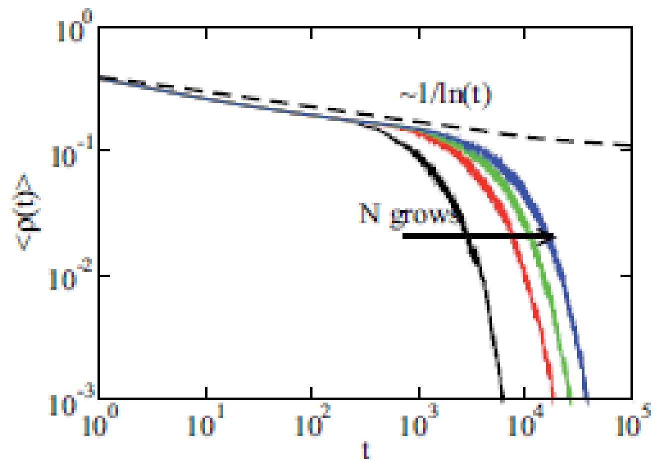
## *d=1,2: Coarsening/Ordering*

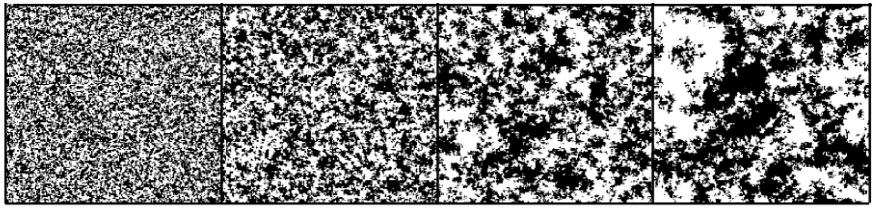
Unbounded growth of domains of absorbing states

*d=1*




*d=2*



*Coarsening driven by interfacial noise*

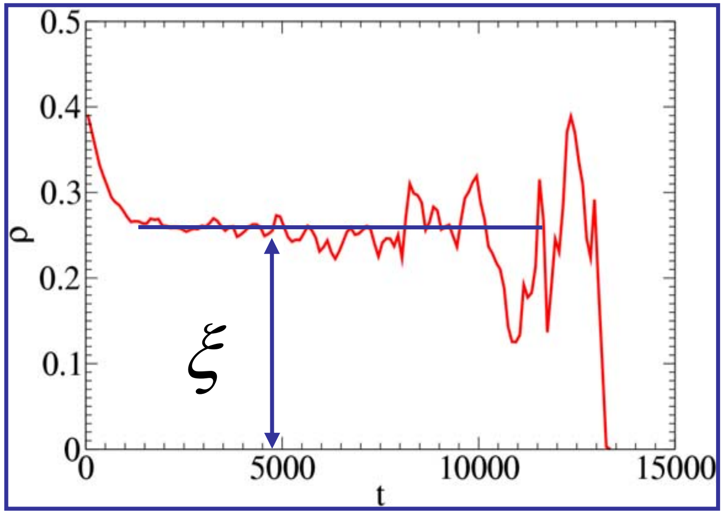
## d>2 regular and complex networks

$<\rho>\sim\xi$

$\tau(N)\approx N$, survival time of metastable state

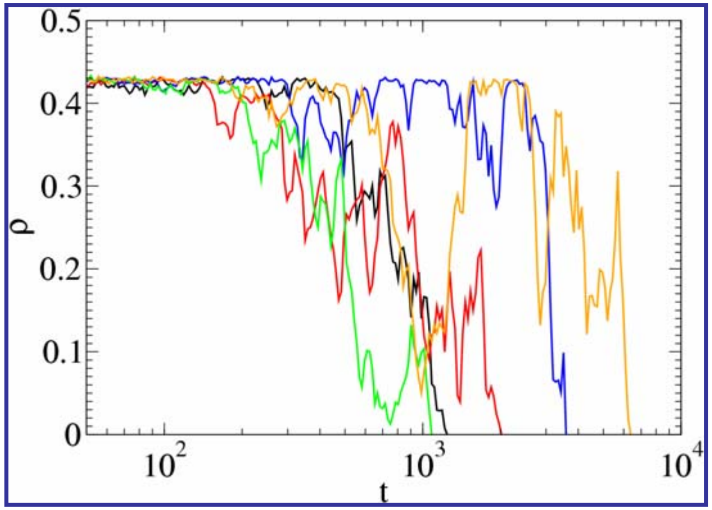
### *d>2: No Coarsening : Dynamical Metastability*

**Disordered states.**




$l=\xi^{-1}$ *Characteristic size of ordered domain*

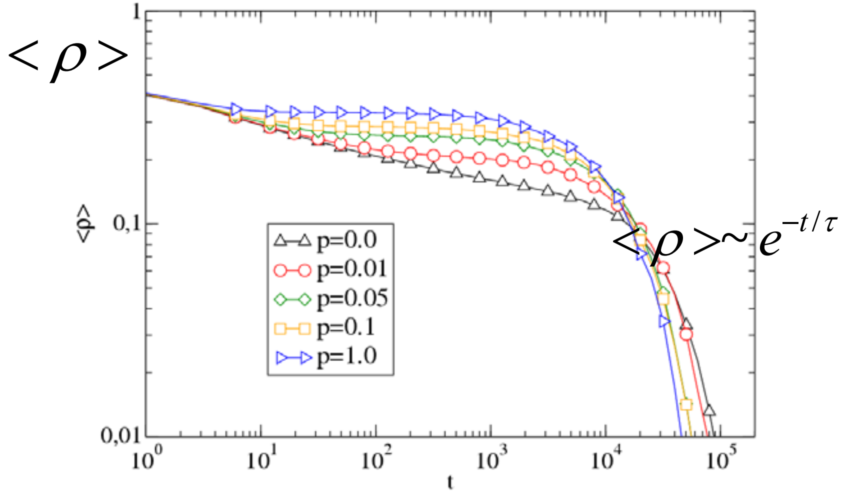
**Finite size fluctuations take the system to an absorbing state**

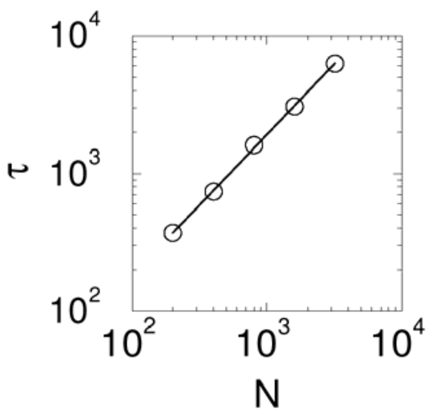



$<\rho>\sim e^{-t/\tau}$

# Voter Model in Complex Networks

## Small World Networks



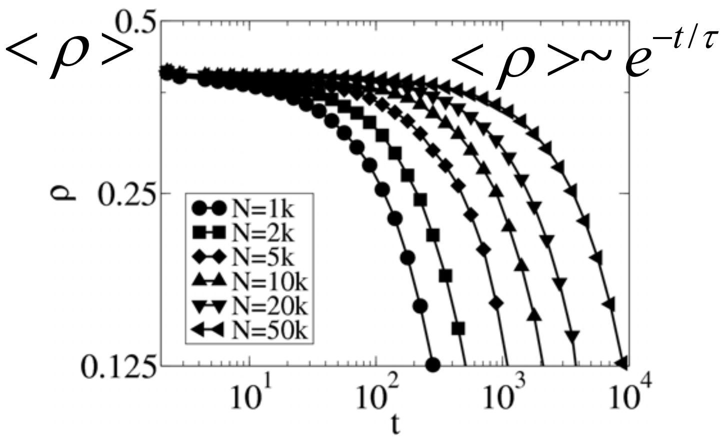



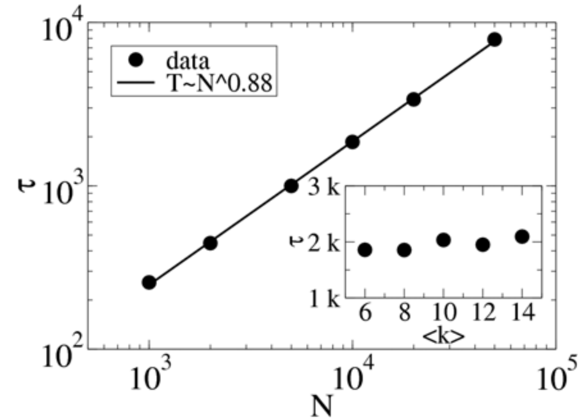

Survival time scales as in regular d>2:

$\tau \sim N$

Castellano et al, Eur. Phys. Lett. 63,153(2003)

## Scale Free Barabasi-Albert Nets




$\tau \sim N \ln N$

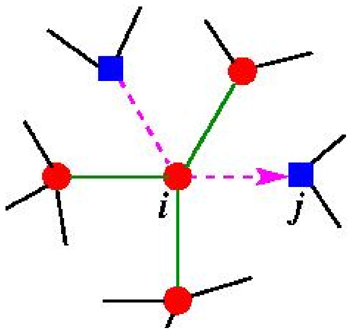
Suchecki et al, Eur. Phys. Lett. **69**,228(2005)

F. Vazquez, V.M. Eguiluz, New J. Phys. 10 (2008) 063011

## UNCORRELATED NETWORKS

$n_+$ active links

$k-n_+$ inert links



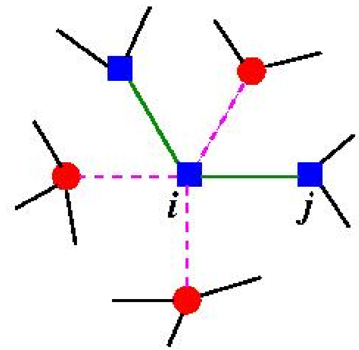

$$P(-\rightarrow+)=\frac{1}{2}\frac{n_+}{k}$$

Node of degree *k*

$s_i=-1$, $n_+$ + neighbors

$k-n_+$ active links

$n_+$ inert links




$$\Delta\rho=\frac{2(k-2n_+)}{<k>N}$$

$$\Delta m=2/N$$

$s_i=+1$

Coupled eqs. for <m> and <ρ>

Mean field link approx. for *Prob* $(k, -, n_+)$:

Neglect 2nd nearest-neighbor correlations

# Voter Model in Uncorrelated Networks

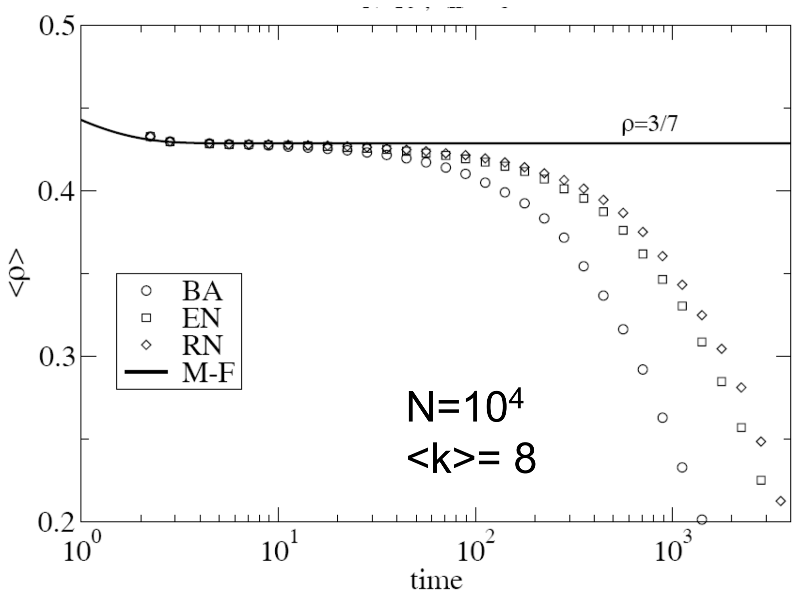
*F. Vázquez et al. Phys. Rev. Lett. 100, 108702 (2008)*

## Mean Field Link Dynamics:

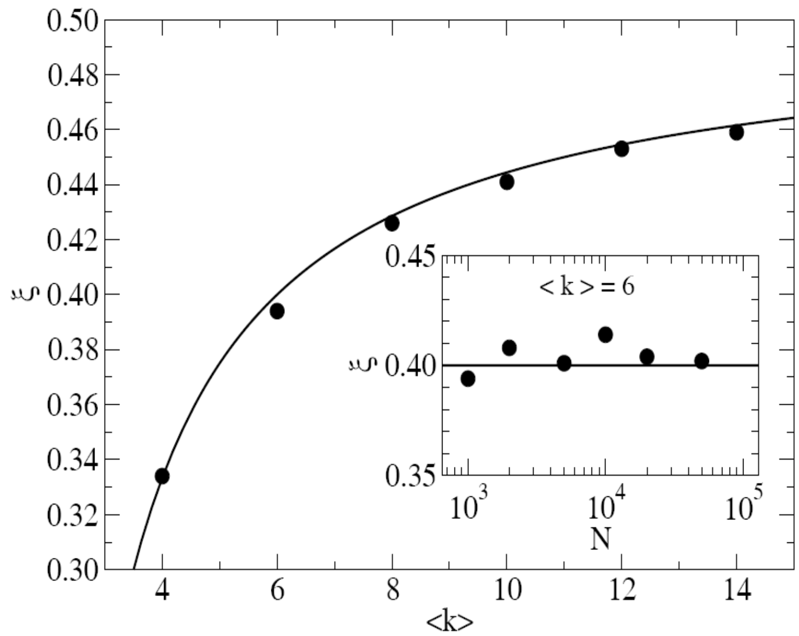
$$\rho^{s} = \xi = \frac{<k> - 2}{2(<k> - 1)}$$

*Single parameter theory*

## Network topology independence




## Barabasi-Albert Scale Free Networks



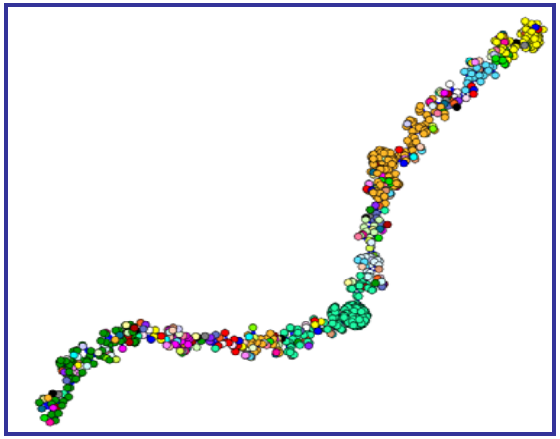
K. Suchecki, et al Phys. Rev. E 72, 036132(2005)

# Role of dimensionality

*1D Scale free net?*

*Structured SF:* ***SSF***

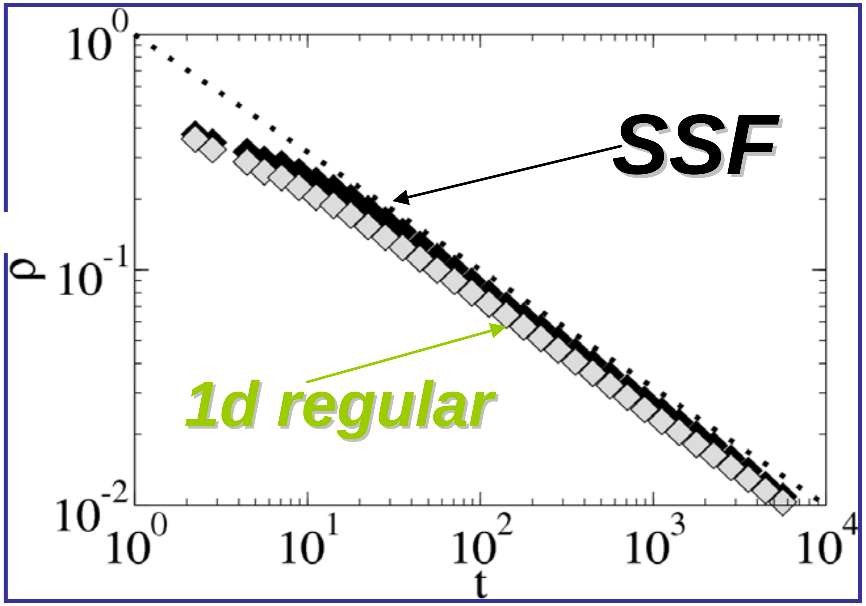
*Klemm and Eguíluz,*
*Phys. Rev. E* ***65****,036123 (2002)*



*Scale free but*
*high clustering and 1d*

***P(k)*** $\sim$ ***k*** $^{-3}$

***L*** $\sim$ ***N*** ***C*** $\sim$ ***N***$^0$




**SSF** $\rho \sim t^{-1/2}$

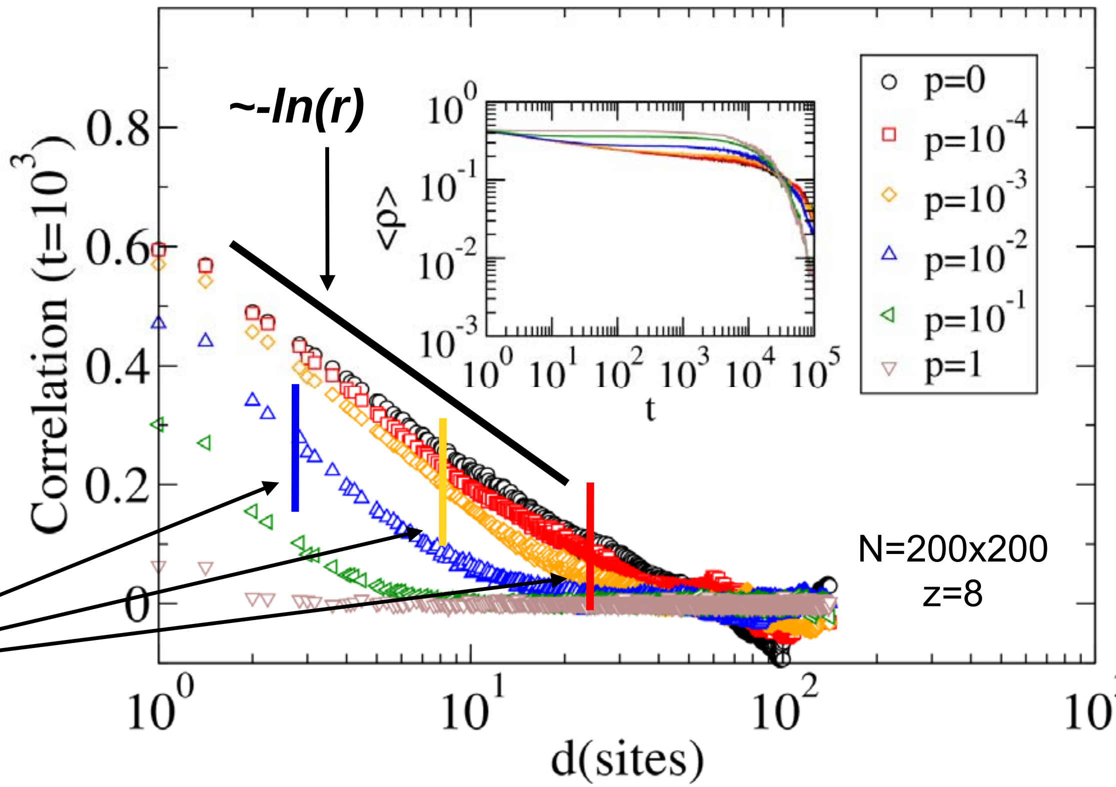
$\tau_1 \approx N^2$

***Dimensionality*** *determines when voter dynamics orders the system*

***Degree*** *distribution or network* ***disorder*** *are not relevant*

## 2d Small World Network

A small world network in 2d is locally a 2d network.




$$r_{max}(p) = \frac{1}{\sqrt{\frac{z}{2}\pi p}}$$

**Small-world topology destroys correlations.**

**Logarithmic scaling observed only locally .**

# REACHING AGREEMENT BY IMITATION?

- Role of topology of interactions (tie heterogeneity)

- Coevolution (non persistent ties):

  Imitating neighbors vs Choosing neighbors

  *F. Vazquez, et. al, Phys. Rev. Lett.* ***100****, 108702 (2008)*

- Heterogeneity in the timing of interactions

  *J. Fernández-Gracia, et al. Phys. Rev. E* ***84****, 015103 (2011)*

- Imitation vs Rational Behavior

  *D. Vilone, et al. Sci. Rep.* ***2****, 686 (2012)*

Ingredients of a social influence model:

a) Interaction mechanism: Imitation as basic manifestation of social influence.

THEORY

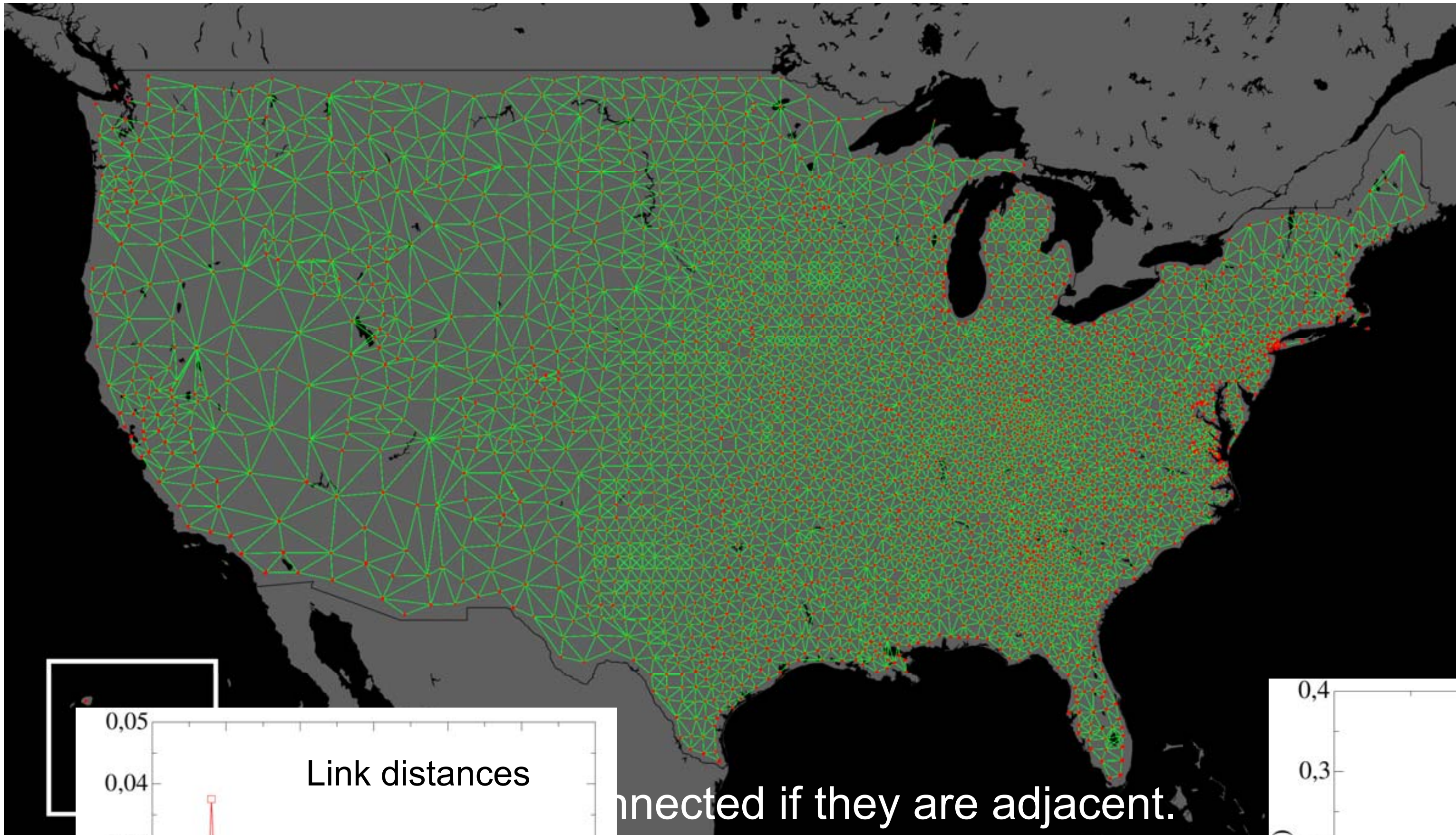
MODELLING

b) Social context: Set of all possible interactions of an individual with any other peer. We model it as a network of interactions from census data for population and mobility.

INPUT DATA

## US: geographical adjacency of populations

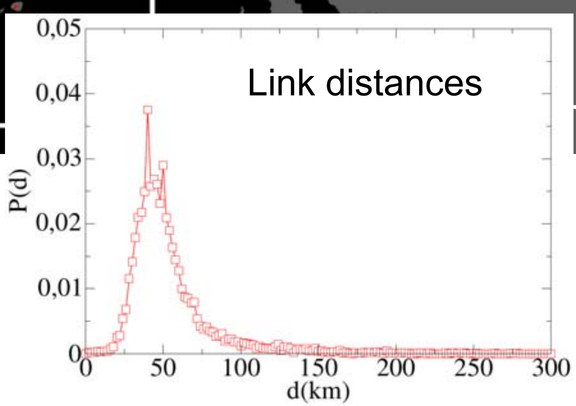
Undirected, unweighted network.

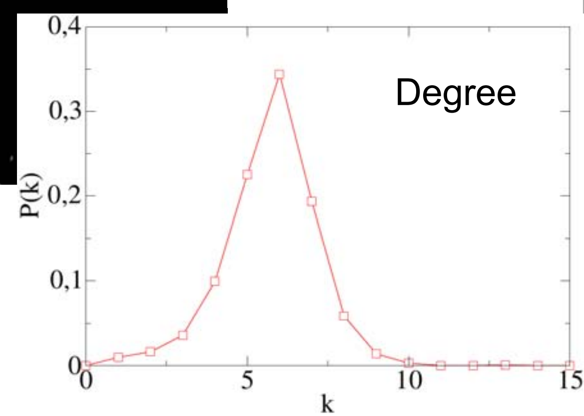



N=3114
<k>~6

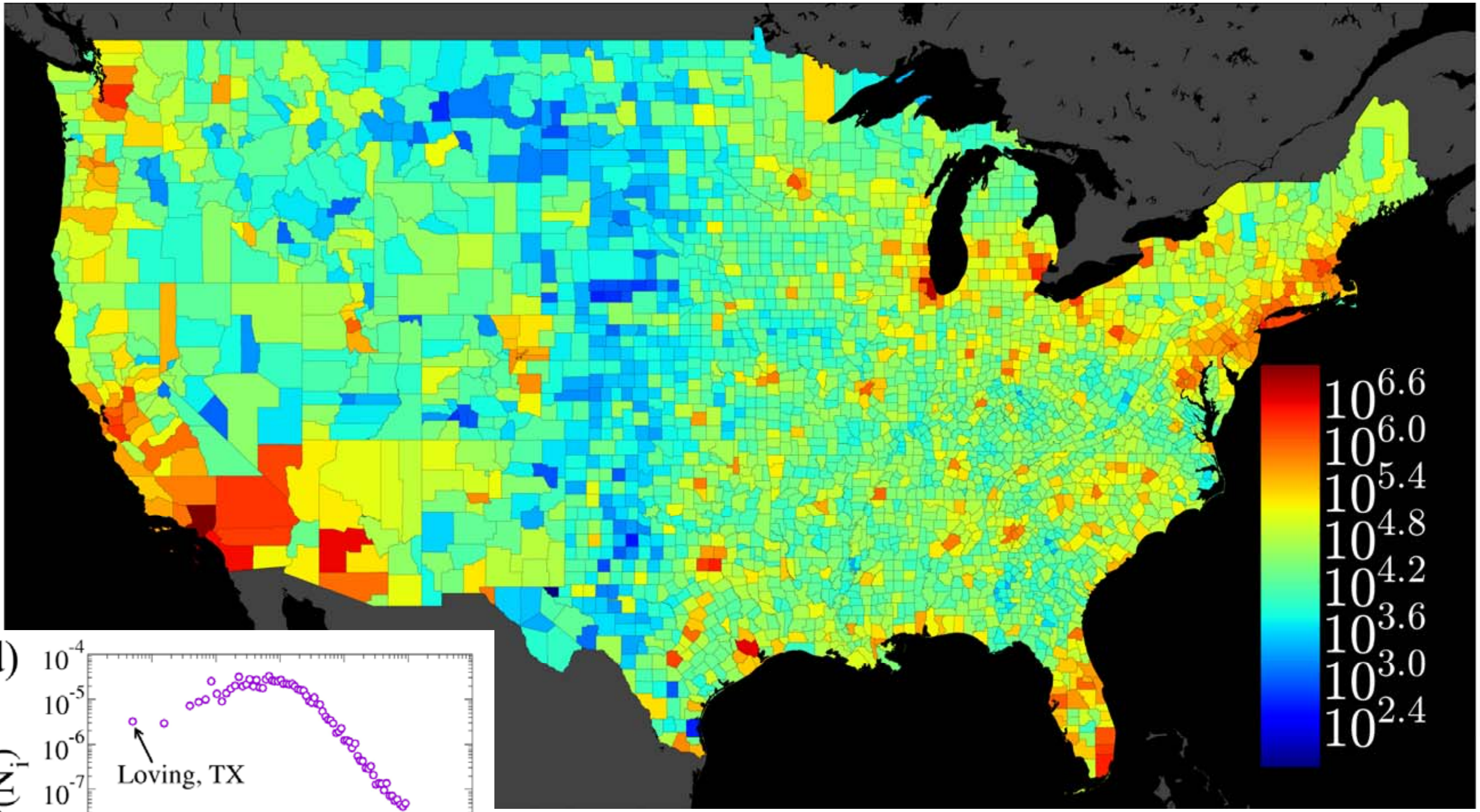
<C> = 0.431
<L> = 26.86

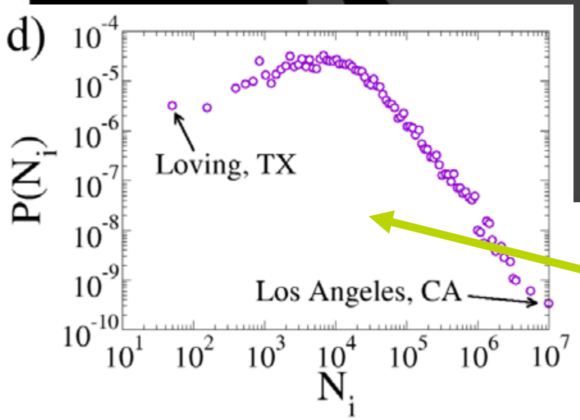
$\langle C \rangle_{Hex} = 0.4$
$\langle L \rangle_{Hex} = 22.15$



## US: Heterogeneous population distribution



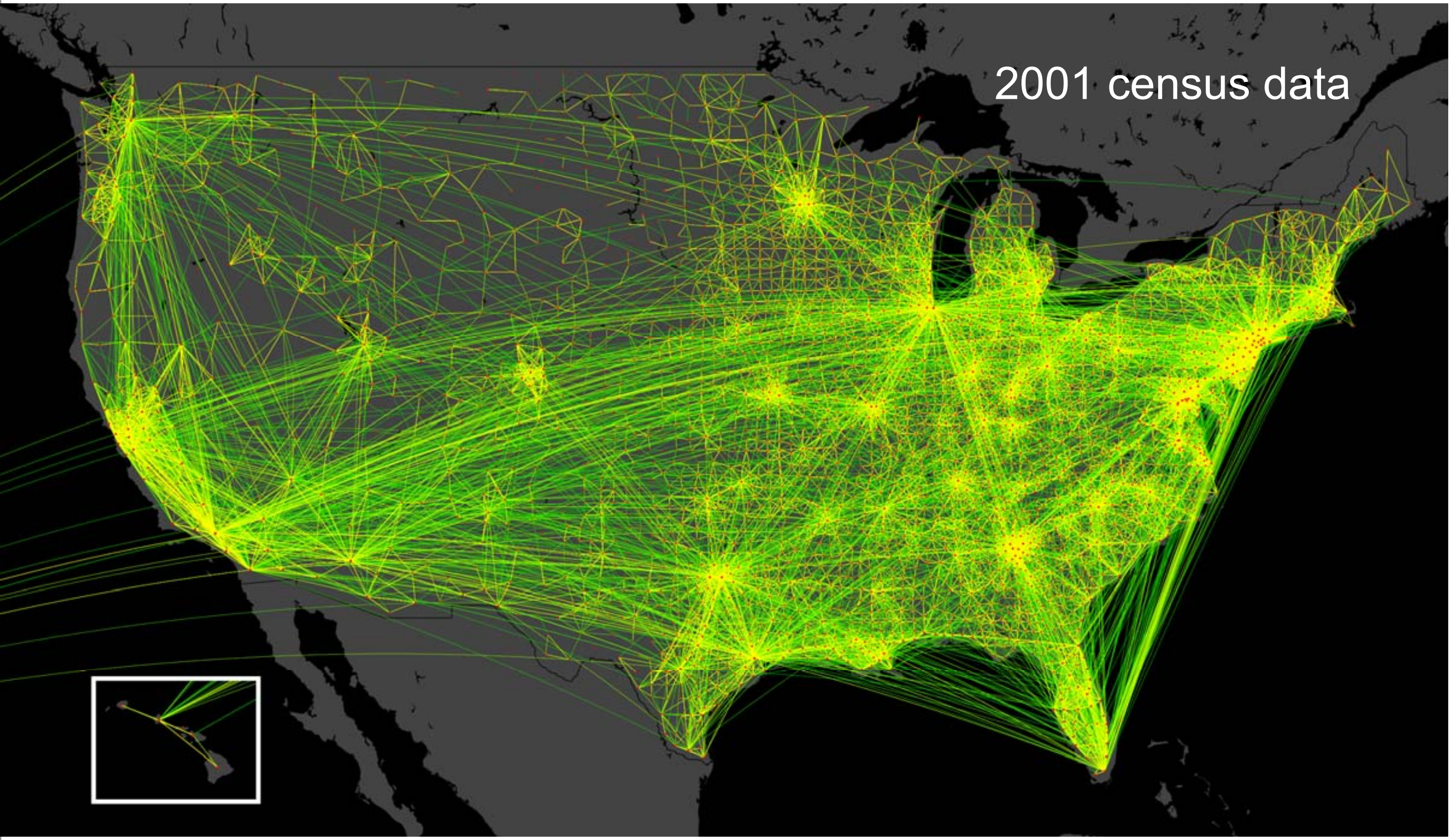




Populations

## Voting population and mobility networks

### US: commuter network for human mobility

Directed, weighted network.




N=3114
<k>~50

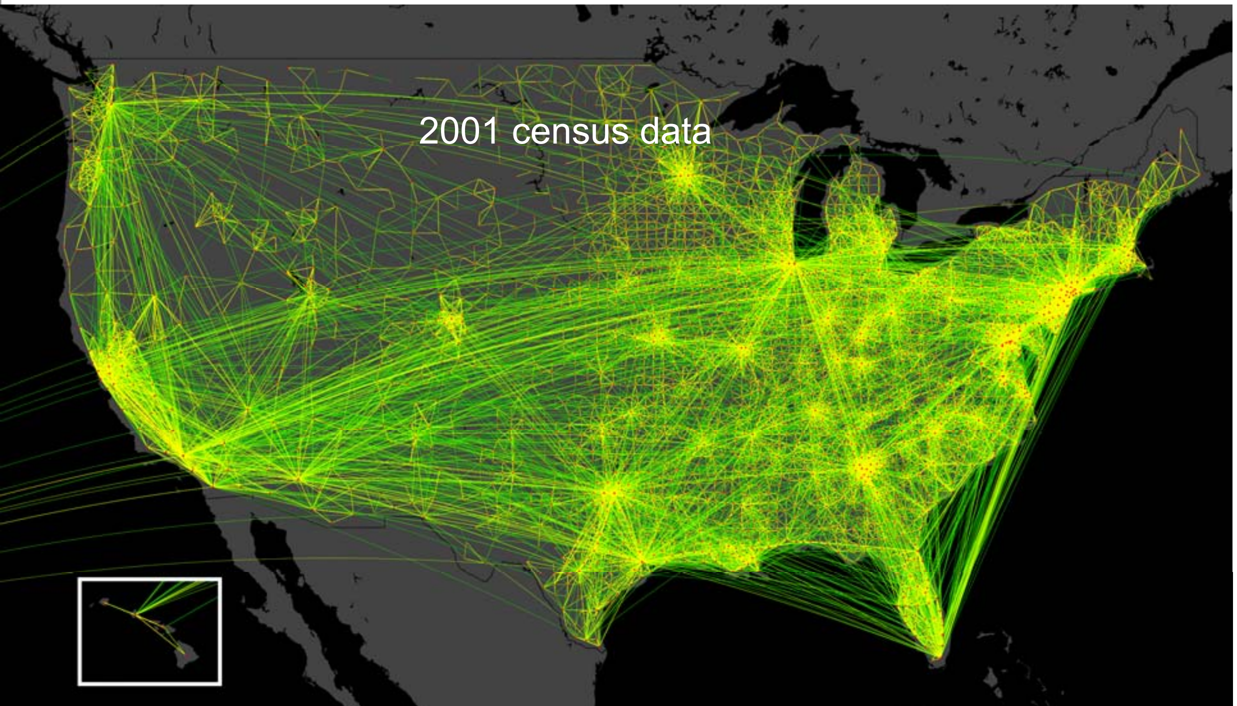
<C> = 0.36
<L> = 2.31

For 10 independent randomizations of degree sequence:

<C> = 0.088
<L> = 2.15

Counties linked by fluxes of commuters. Color indicates number of commuters. 20% of all connections shown.

# Voting population and mobility networks



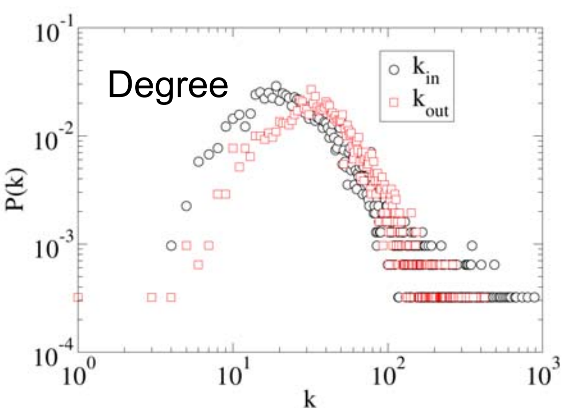

## US: commuter network for human mobility

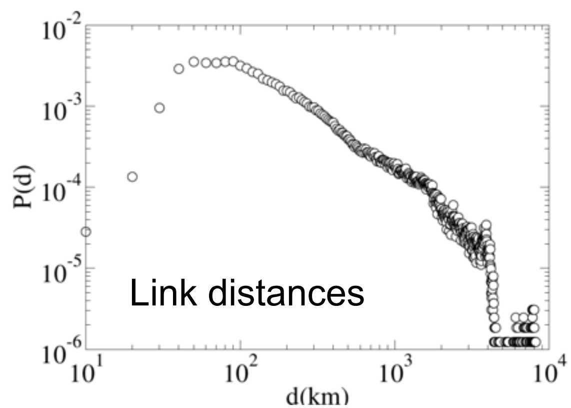
Directed, weighted network.

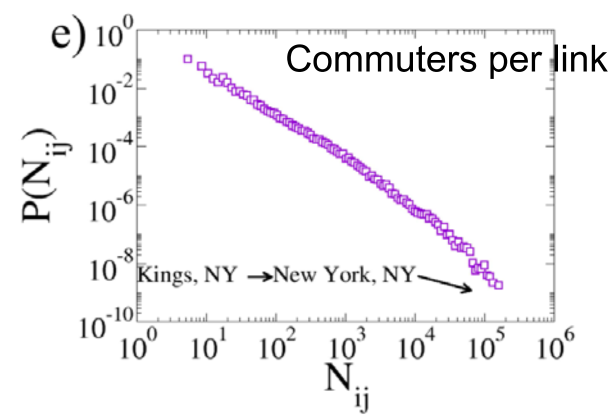
N=3114
<k>~50

Not a 2d network

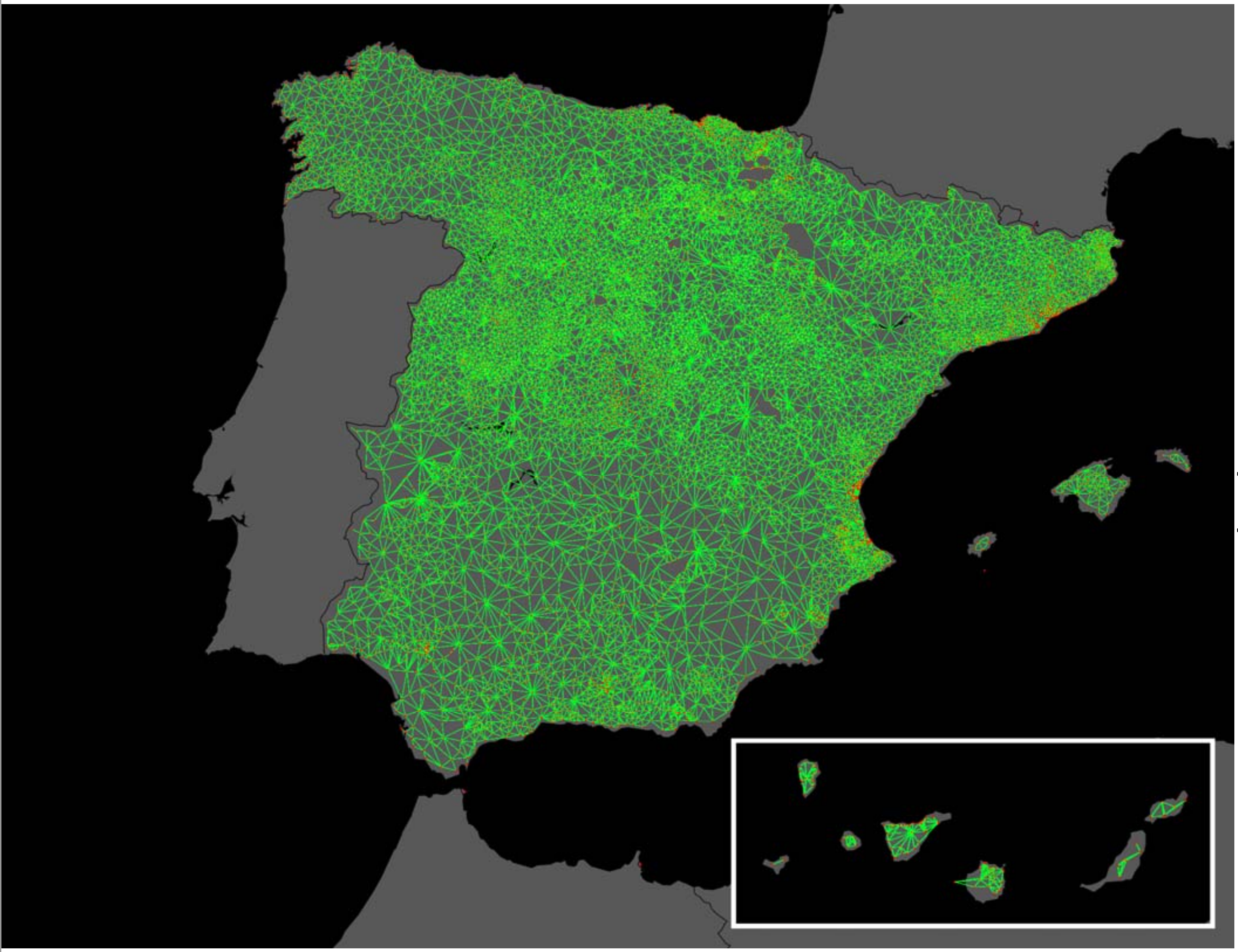
**Heterogeneous network in many characteristics**



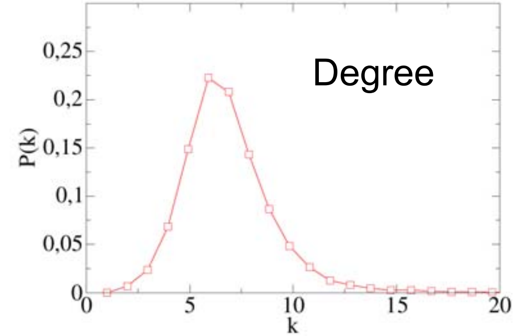
## Spain: geographical adjacency of populations



Municipalities in Spain connected if they are adjacent.

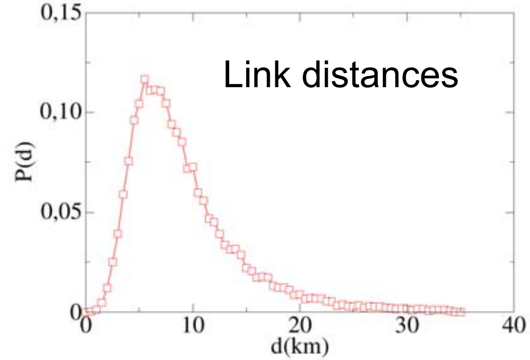
Undirected, unweighted network.

N=8104 <k>~6

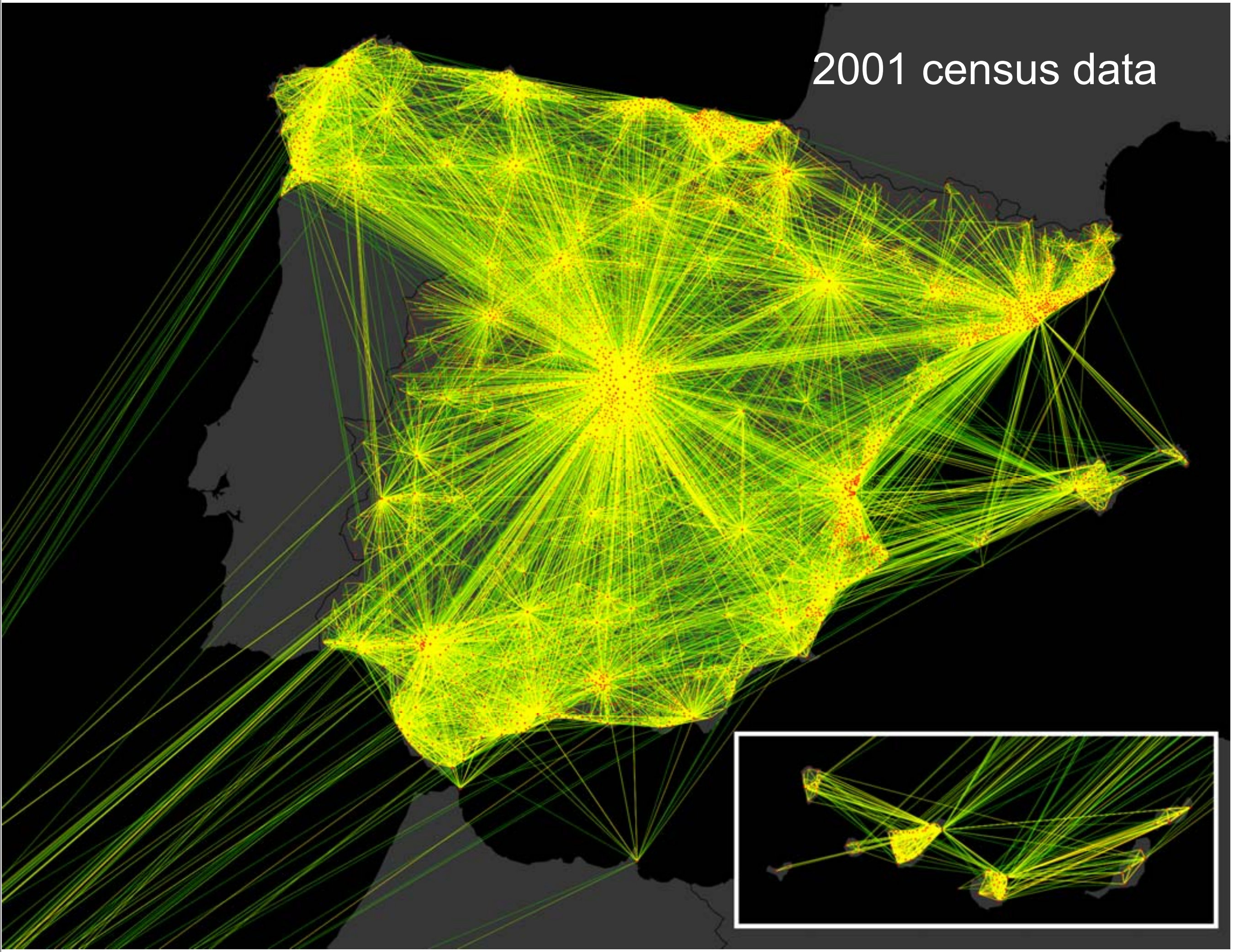



$\langle C \rangle = 0.49$; $\langle C \rangle_{Hex} = 0.4$

$\langle L \rangle = 25.79$; $\langle L \rangle_{Hex} = 35.32$



## Spain: commuter network for human mobility




Directed, weighted network.

N=8104
<k>~34

<C> = 0.54
<L> = 2.34

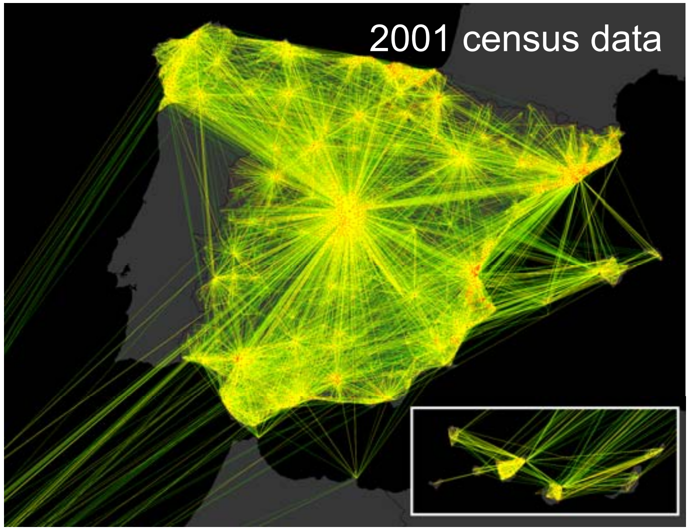
For 10 independent randomizations of degree sequence:

<C> = 0.096
<L> = 2.53

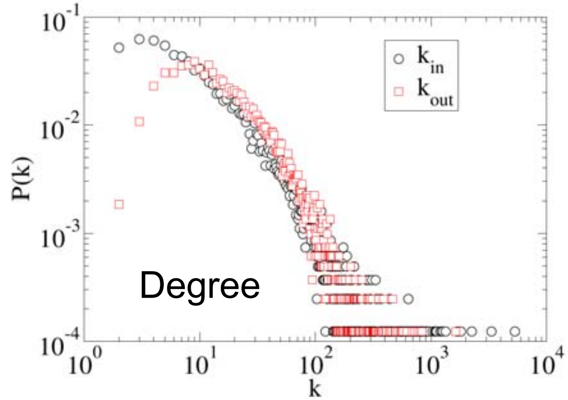
Municipalities linked by fluxes of commuters. Color indicates number of commuters. 20% of all connections shown.

# Voting population and mobility networks

## Spain: commuter network for human mobility



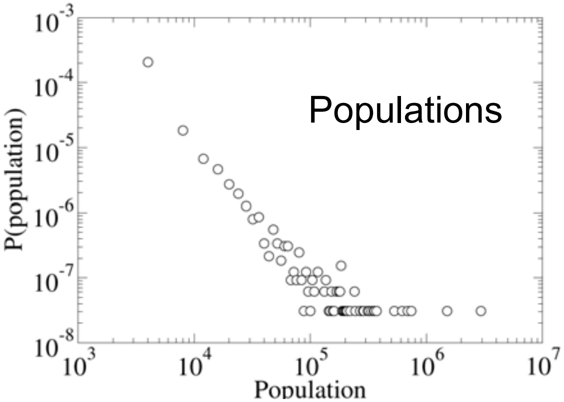

Directed, weighted network.

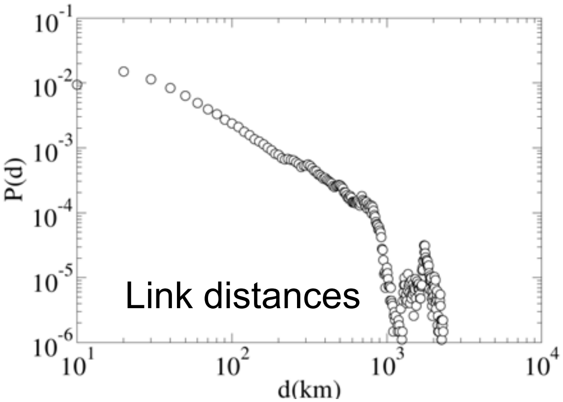


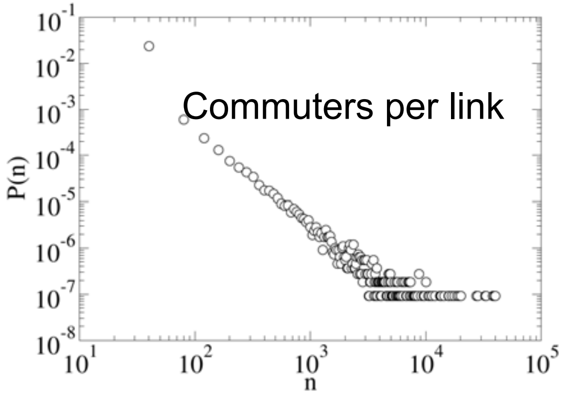

N=8104

<k>~34

**Not a 2d network**




Heterogeneous network in many characteristics

Ingredients of a social influence model:

a) **Interaction mechanism:** Imitation as basic manifestation of social influence.

THEORY

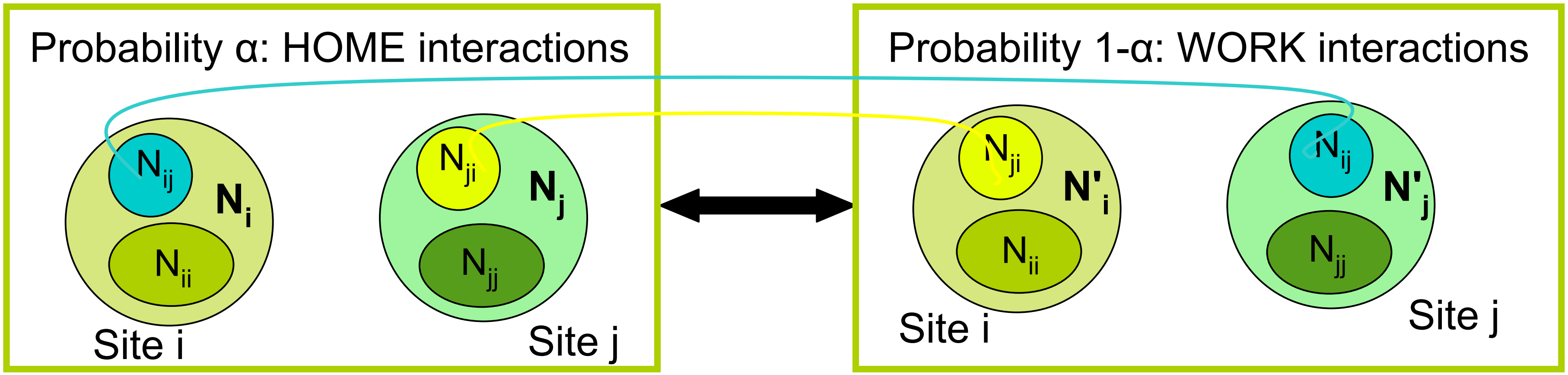
MODELLING

b) **Social context:** Set of all possible interactions of an individual with any other peer. We model it as a network of interactions from census data for population and mobility.

INPUT DATA

# Mobility: Metapopulation Voter Model

- N agents with a binary variable (state, opinion,...) with voter-like interaction
- There are $N_{sites}$ (counties).
- Each agent is considered in two sites: where she lives and where she works.



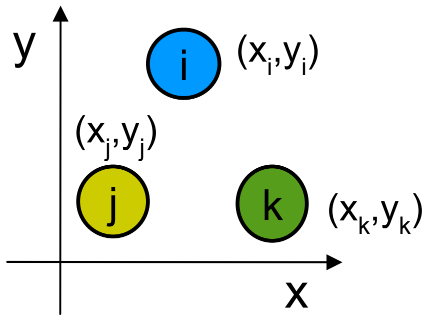

- $N_{ij}$ = # of agents living in i and working in j.
- $N_i$ = # number of agents living in i = $N_{ij} + N_{ii}$
- $N'_i$ = # number of agents working in i.

**An agent interacts with probability α with anyone in $N_i$: lives where she lives.**

**With probability 1-α interacts with anyone in $N'_i$ : works where she works.**

# Networks in Metapopulation Voter Model

## Geography: adjacency network

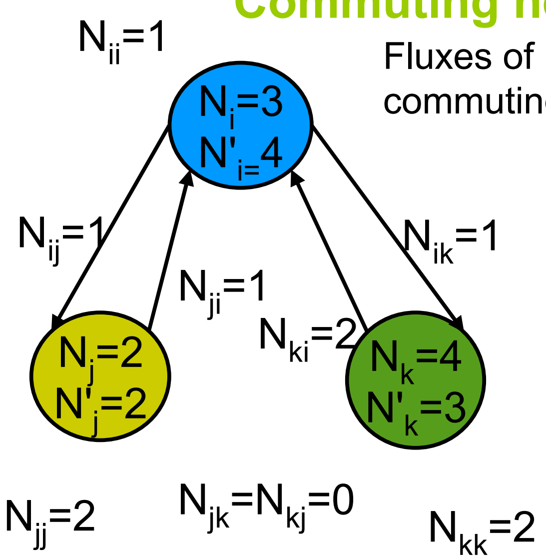



Geographic sites

Euclidean distances between pairs of sites

## Commuting network

Fluxes of commuting agents.

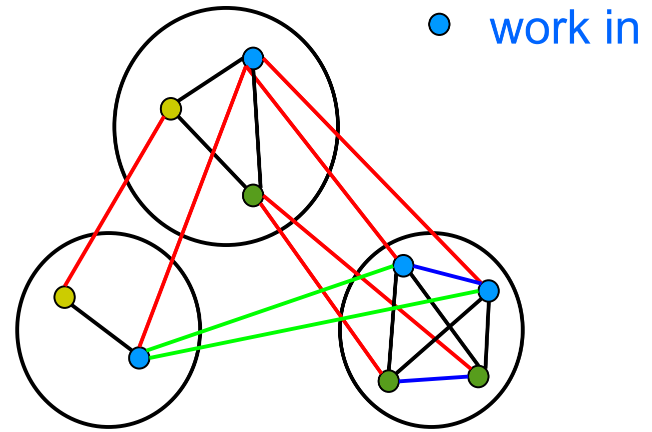



$N_i$: # living in i.

$N'_i$: # working in i.

$N_{ij}$: # living in i and working in j.

## Interaction network




| Link | P | Meaning |
|---|---|---|
| — (black) | α | Live in same place, work in different. |
| — (red) | 1-α | One works where the other works and lives. |
| — (green) | 1-α | Live in different places, work in same place. |
| — (blue) | 1 | Live in same place, work in same place. |

# Metapopulation Voter model: Analytical insights

## Parameters (census)

$N_{ij}$: number of agents living in i and working in j. $N_i = \sum_j N_{ij}$

$X_i, Y_i$: location of city i. $N'_i = \sum_j N_{ji}$

## Variables

$V_{ij}$: number of agents living in i and working in j holding opinion +1. $v_{ij} = \frac{V_{ij}}{N_{ij}}$

Correlations $\langle v_i v_j \rangle$ of densities $v_i = \frac{\sum_j V_{ij}}{N_{ij}}$

## Transition rates

$$r^+_{ij}(V_{ij} \to V_{ij}+1) = (N_{ij} - V_{ij})\left[\alpha \frac{V_i}{N_i} + (1-\alpha)\frac{V'_j}{N'_j}\right]$$

$$r^-_{ij}(V_{ij} \to V_{ij}-1) = V_{ij}\left[\alpha \frac{N_i - V_i}{N_i} + (1-\alpha)\frac{N'_j - V'_j}{N'_j}\right]$$

**Simulation step:** Update in random order each of the $V_{ij}$

## Master equation

$$\frac{\partial P(\{V_{ij}\};t)}{\partial t} = \sum_{i,j} \Big[ r^{+}_{ij}(\{\ldots,V_{ij}-1,\ldots\} \to \{V_{ij}\}) P(\{\ldots,V_{ij}-1,\ldots\};t)$$
$$+ r^{-}_{ij}(\{\ldots,V_{ij}+1,\ldots\} \to \{V_{ij}\}) P(\{\ldots,V_{ij}+1,\ldots\};t)$$
$$- \left( r^{+}_{ij}(\{V_{ij}\} \to \{\ldots,V_{ij}+1,\ldots\}) + r^{-}_{ij}(\{V_{ij}\} \to \{\ldots,V_{ij}-1,\ldots\}) \right) P(\{V_{ij}\};t) \Big]$$

## Langevin equation

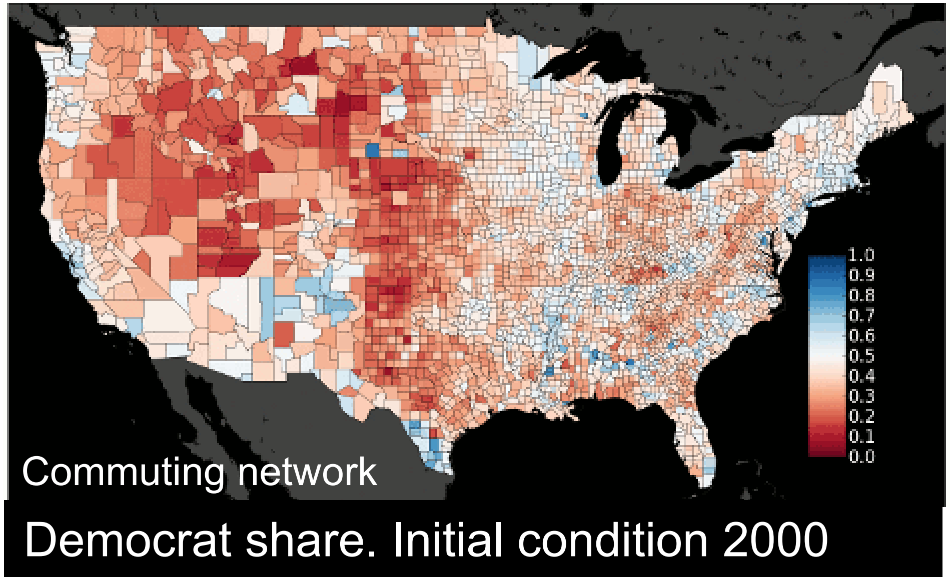
$$\frac{dv_{ij}}{dt} = \alpha \sum_{l} \left( \frac{N_{il}}{N_i} - \delta_{jl} \right) v_{il} + (1-\alpha) \sum_{l} \left( \frac{N_{lj}}{N'_j} - \delta_{li} \right) v_{lj}$$
$$+ \frac{1}{\sqrt{N_{ij}}} \sqrt{(1-2v_{ij}) \left( \alpha \frac{\sum_l N_{il} v_{il}}{N_i} + (1-\alpha) \frac{\sum_l N_{lj} v_{lj}}{N'_j} \right) + v_{ij}} \; \eta^{*}_{ij}(t).$$

$$\boxed{v_{ij} = \frac{V_{ij}}{N_{ij}}}$$

**Laplacian model**: Conserved quantity: $\sum_{ij} \langle V_{ij} \rangle$ <Number of + voters>

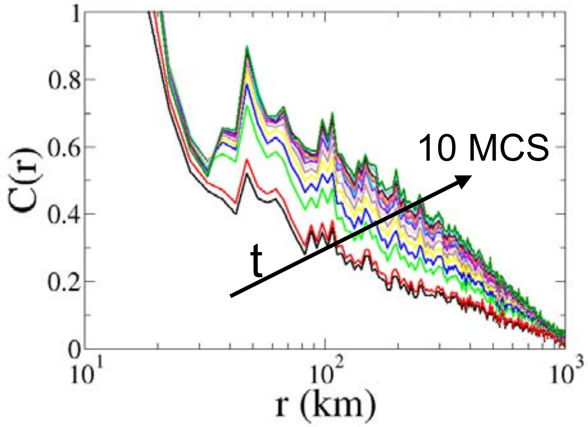
## US presidential elections: Metapopulation Voter Model simulations

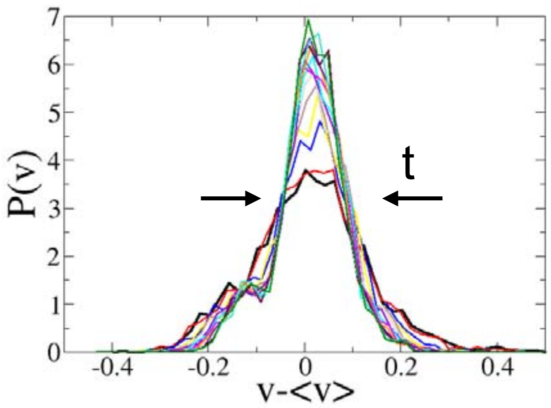
$\alpha$=1/2




Democrat share. Initial condition 2000

Diffusion process:
→ correlations grow, share distribution narrows.




**Extra ingredient needed for stationarity:**
**Imperfect imitation or External noise**

# Metapopulation Voter model: Analytical insights

## Parameters (census)

$N_{ij}$: number of agents living in i and working in j. $N_i = \sum_j N_{ij}$

$X_i, Y_i$: location of county i. $N'_i = \sum_j N_{ji}$

## Variables

$V_{ij}$: number of agents living in i and working in j holding opinion +1. $v_{ij} = \frac{V_{ij}}{N_{ij}}$

Correlations $\langle v_i v_j \rangle$ of densities $v_i = \frac{\sum_j V_{ij}}{N_{ij}}$

## Transition rates

Transition rates

$$r^+_{ij}(V_{ij} \to V_{ij}+1) = (N_{ij} - V_{ij})\left[\alpha \frac{V_i}{N_i} + (1-\alpha)\frac{V'_j}{N'_j}\right] + \boxed{N_{ij}\frac{D}{2}\eta^+_{ij}(t),}$$

$$r^-_{ij}(V_{ij} \to V_{ij}-1) = V_{ij}\left[\alpha \frac{N_i - V_i}{N_i} + (1-\alpha)\frac{N'_j - V'_j}{N'_j}\right] \boxed{+ N_{ij}\frac{D}{2}\eta^-_{ij}(t)}$$

Imperfect imitation

## Langevin equation

Langevin equation

$$\frac{dv_{ij}}{dt} = \alpha \sum_l \left(\frac{N_{il}}{N_i} - \delta_{jl}\right) v_{il} + (1-\alpha)\sum_l \left(\frac{N_{lj}}{N'_j} - \delta_{li}\right) v_{lj} \boxed{+ D\eta_{ij}(t)} \quad (7)$$

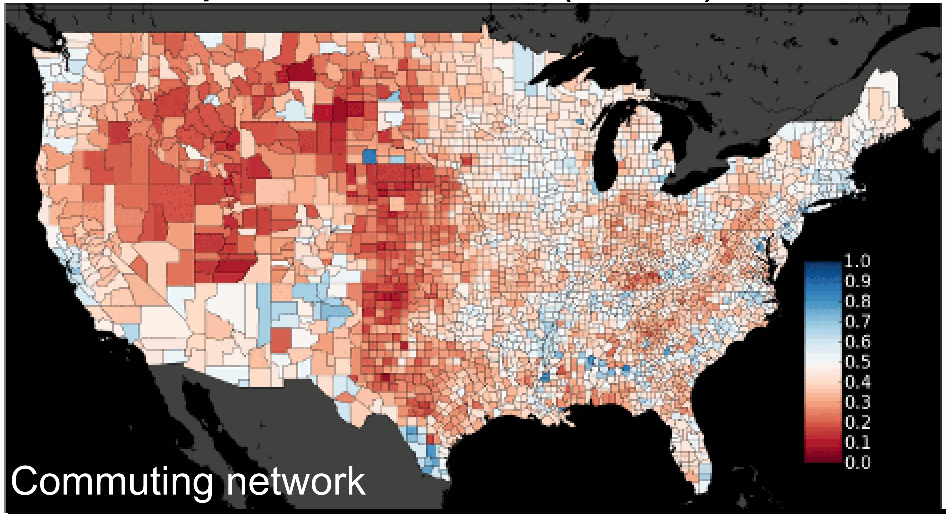
Laplacian model for $< v_{ij} >$

Conserved quantity

$$+ \frac{1}{\sqrt{N_{ij}}}\sqrt{(1-2v_{ij})\left(\alpha \frac{\sum_l N_{il} v_{il}}{N_i} + (1-\alpha)\frac{\sum_l N_{lj} v_{lj}}{N'_j}\right) + v_{ij}} \boxed{+ \frac{D}{2}\eta'_{ij}(t)}\eta^*_{ij}(t).$$

# US presidential elections: Metapopulation Voter Model simulations

*Imperfect* imitation (D=0.2)



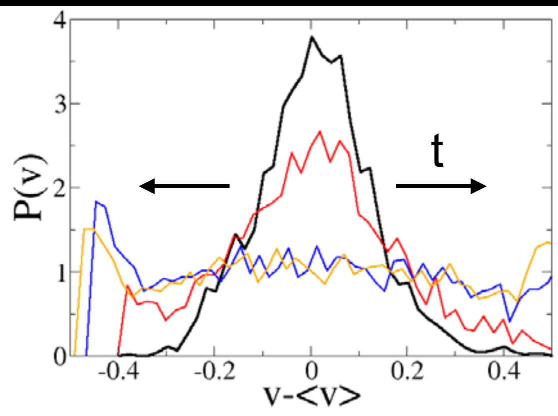

Democrat share. Initial condition 2000

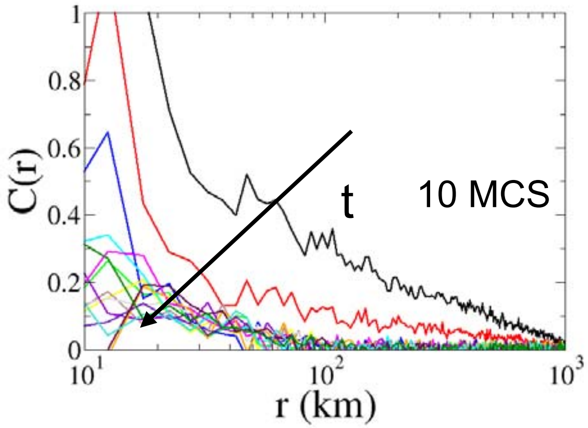
$$\frac{dv_{ij}}{dt} = \alpha \sum_l \left(\frac{N_{il}}{N_i} - \delta_{jl}\right) v_{il} + (1-\alpha) \sum_l \left(\frac{N_{lj}}{N'_j} - \delta_{li}\right) v_{lj} + D\eta_{ij}(t)$$

$$+ \frac{1}{\sqrt{N_{ij}}} \sqrt{(1-2v_{ij})\left(\alpha \frac{\sum_l N_{il} v_{il}}{N_i} + (1-\alpha)\frac{\sum_l N_{lj} v_{lj}}{N'_j}\right) + v_{ij} + \frac{D}{2}\eta'_{ij}(t)\eta^*_{ij}(t)}.$$

$\alpha=1/2$

Large noise intensity D produces a random field → correlations disappear, share distribution grows wider.



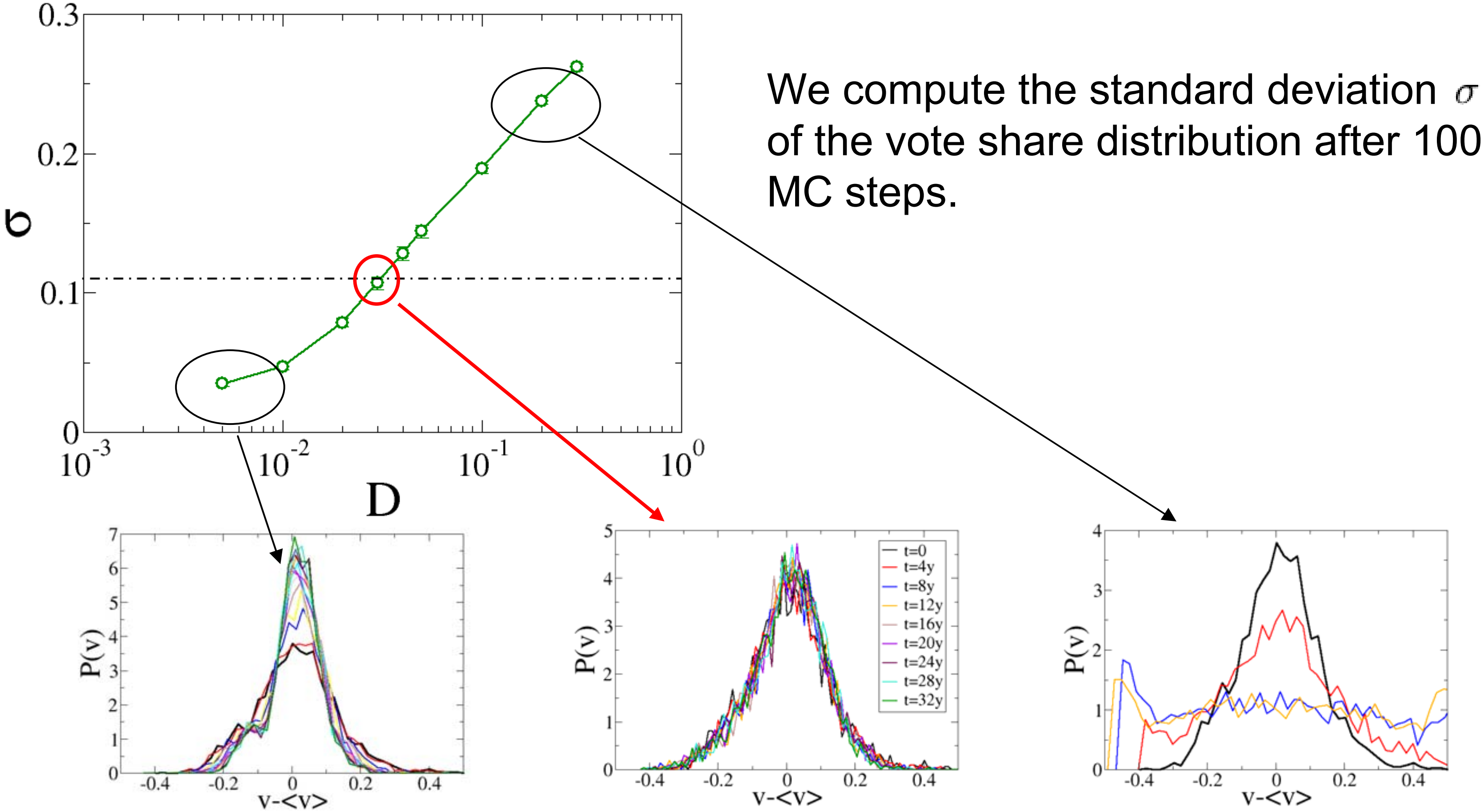

**Question:** ***Is there a value of D for which the process reproduces the noisy diffusive data between a diffusive variable a random field?***

# US presidential elections: Metapopulation Voter Model simulations

## Noise calibration

$\alpha$=1/2

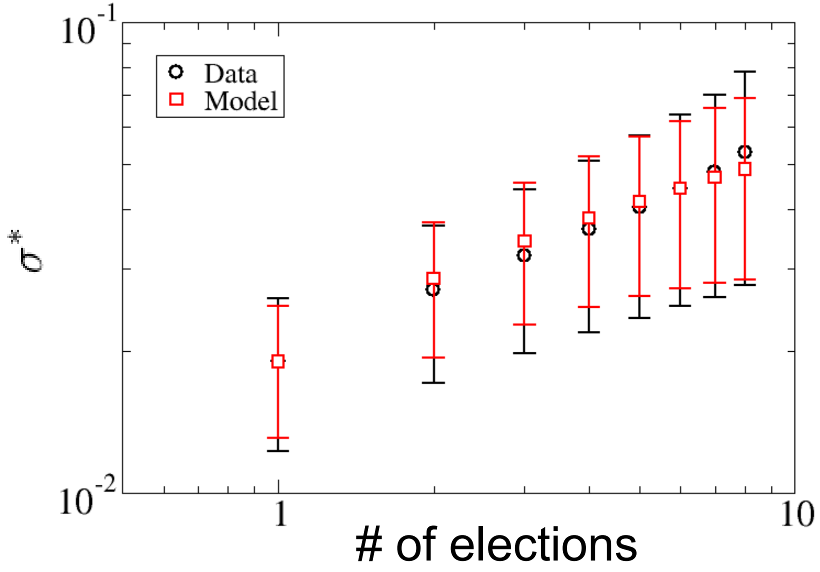
We compute the standard deviation $\sigma$ of the vote share distribution after 1000 MC steps.




**For D=0.03 the st. dev. of the vote share distribution remains stationary and fits the empirical value (data)**

## US presidential elections: Metapopulation Voter Model simulations

# Time calibration




For each time series of a county share we compute the standard deviation $\sigma^*$ as a function of the number of election points. Black symbols in the figure represent the average of this quantity over all counties.

Both data and model show a growth

$$\sigma^* \simeq t^{1/2}$$

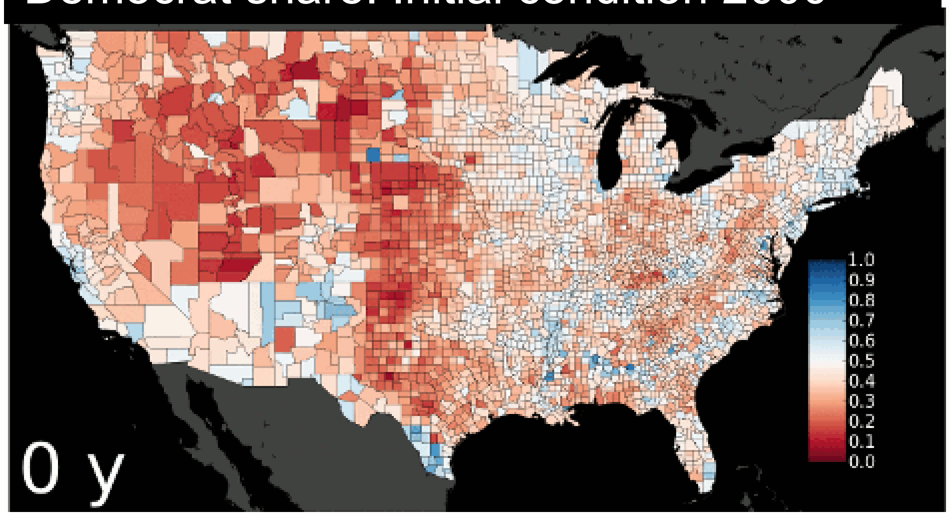
To make the curves overlapping we have to calibrate how many MC steps correspond to the time between two consecutive elections

**10 MC steps =4 years=1 election period**
**St. dev. of county share grows in time equally for model and data.**

# US presidential elections: Metapopulation Voter Model simulations

$\alpha=1/2$

Democrat share. Initial condition 2000



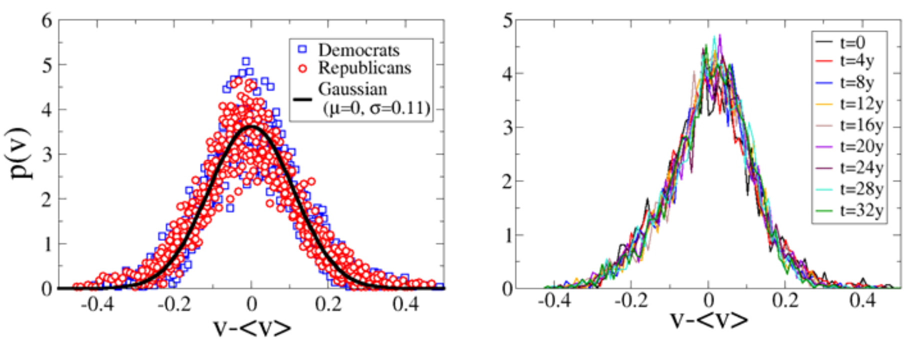

## Calibrated Model

Reproduces the statistical regularities found in election data

*(Single fitted parameter: D)*

### Vote share distribution

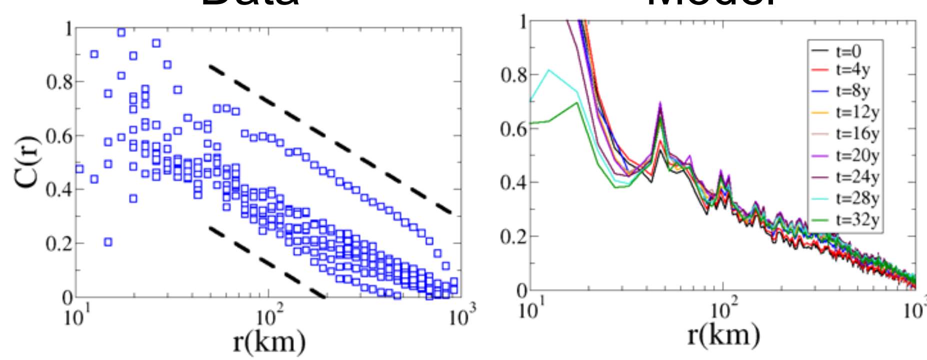
Data Model




St. dev. remains constant.

### Spatial correlations

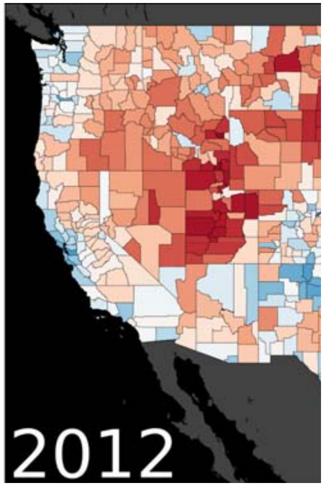
Data Model



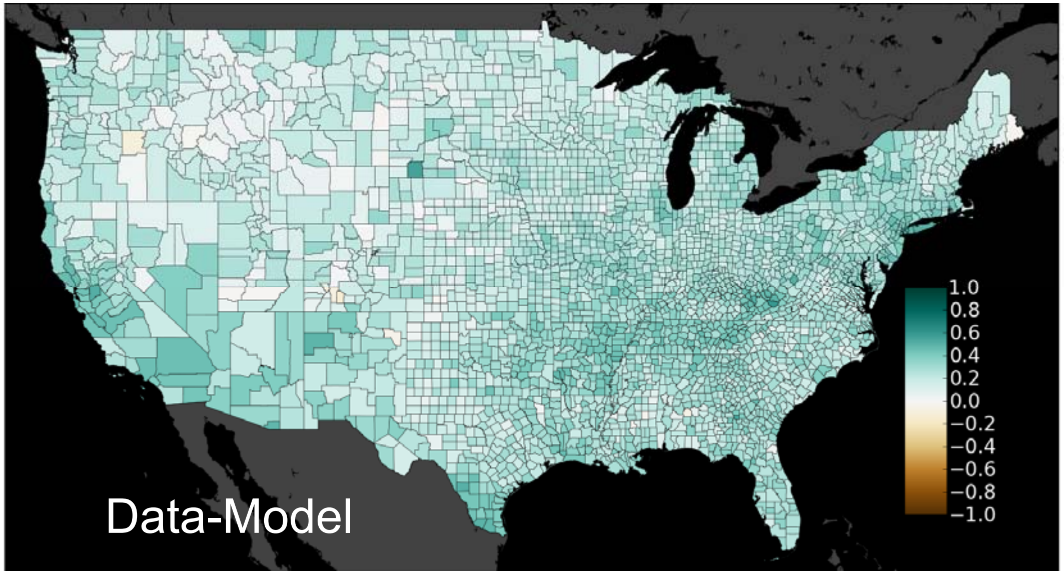

Logarithmic decay remains.

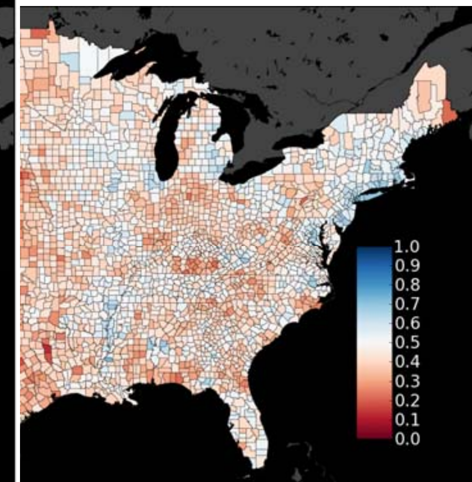
## Electoral predictions?

Results for democrat party

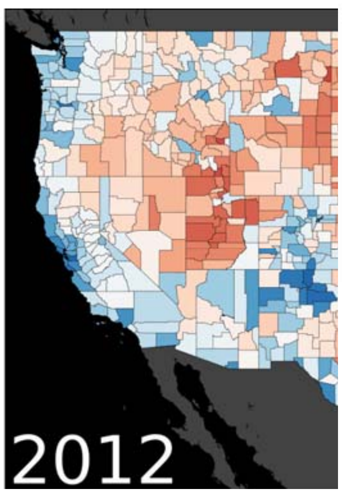



Data

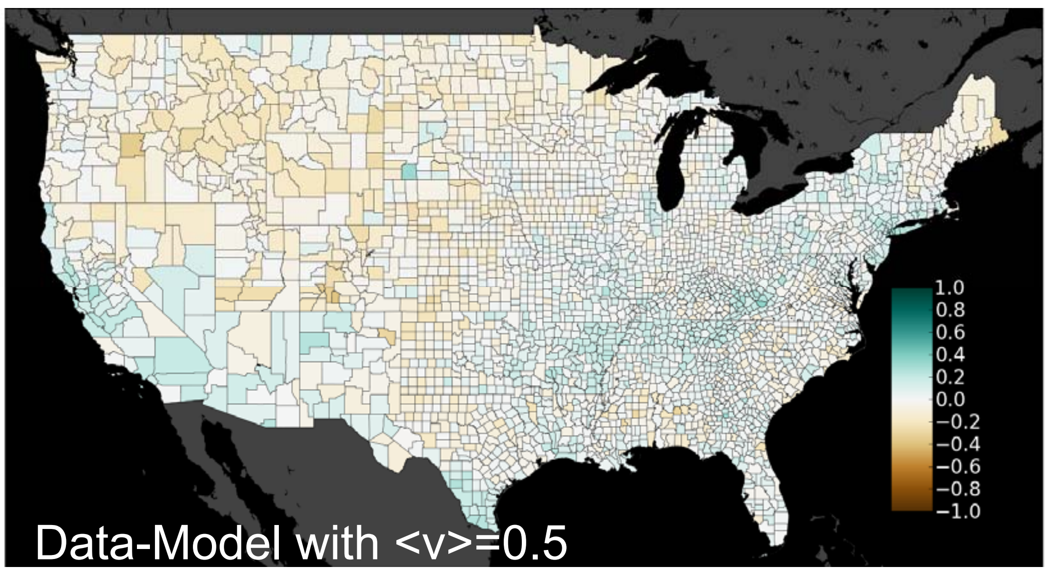


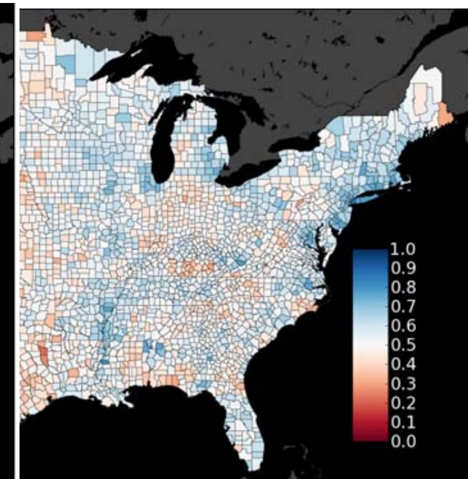




Model (i.c. 2000)



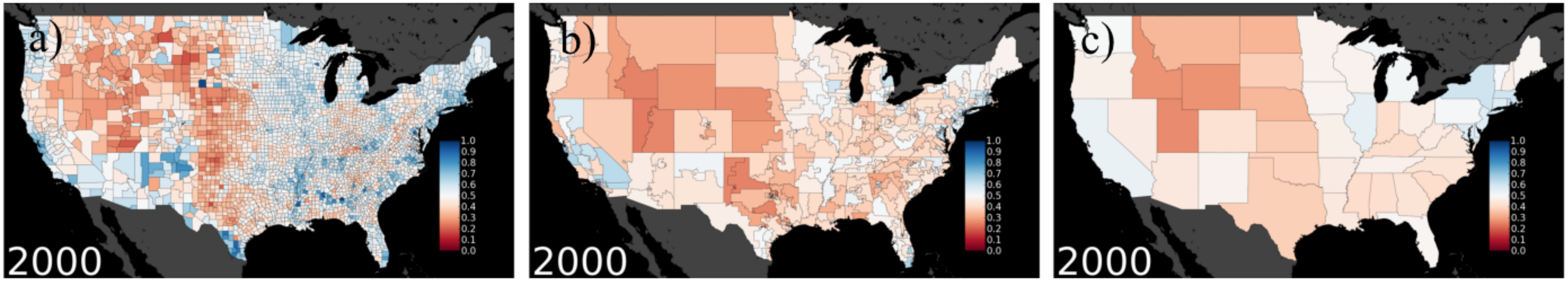

Data with <



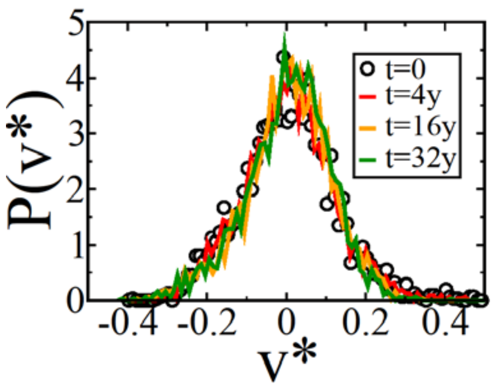




Model with <v>=0.5

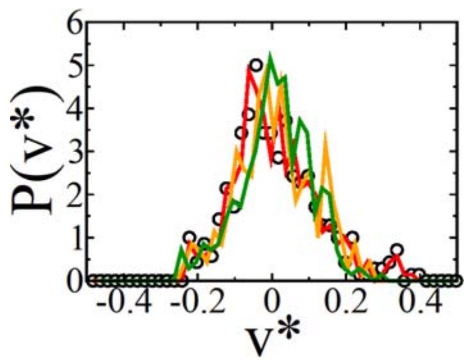
# Predictions at different length scales!!



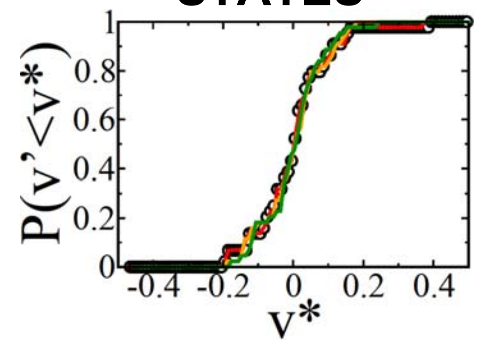

COUNTIES



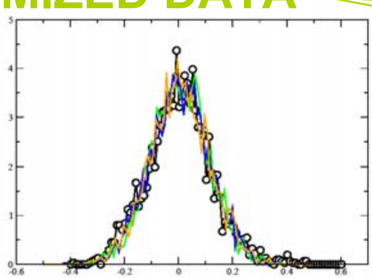

CONGRESSIONAL DISTRICTS

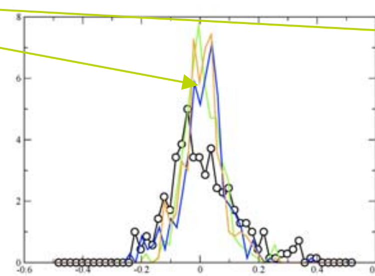


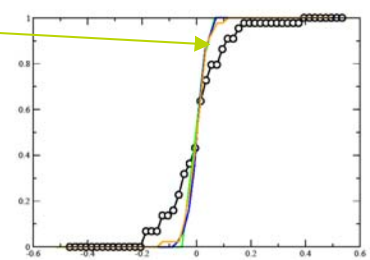

STATES



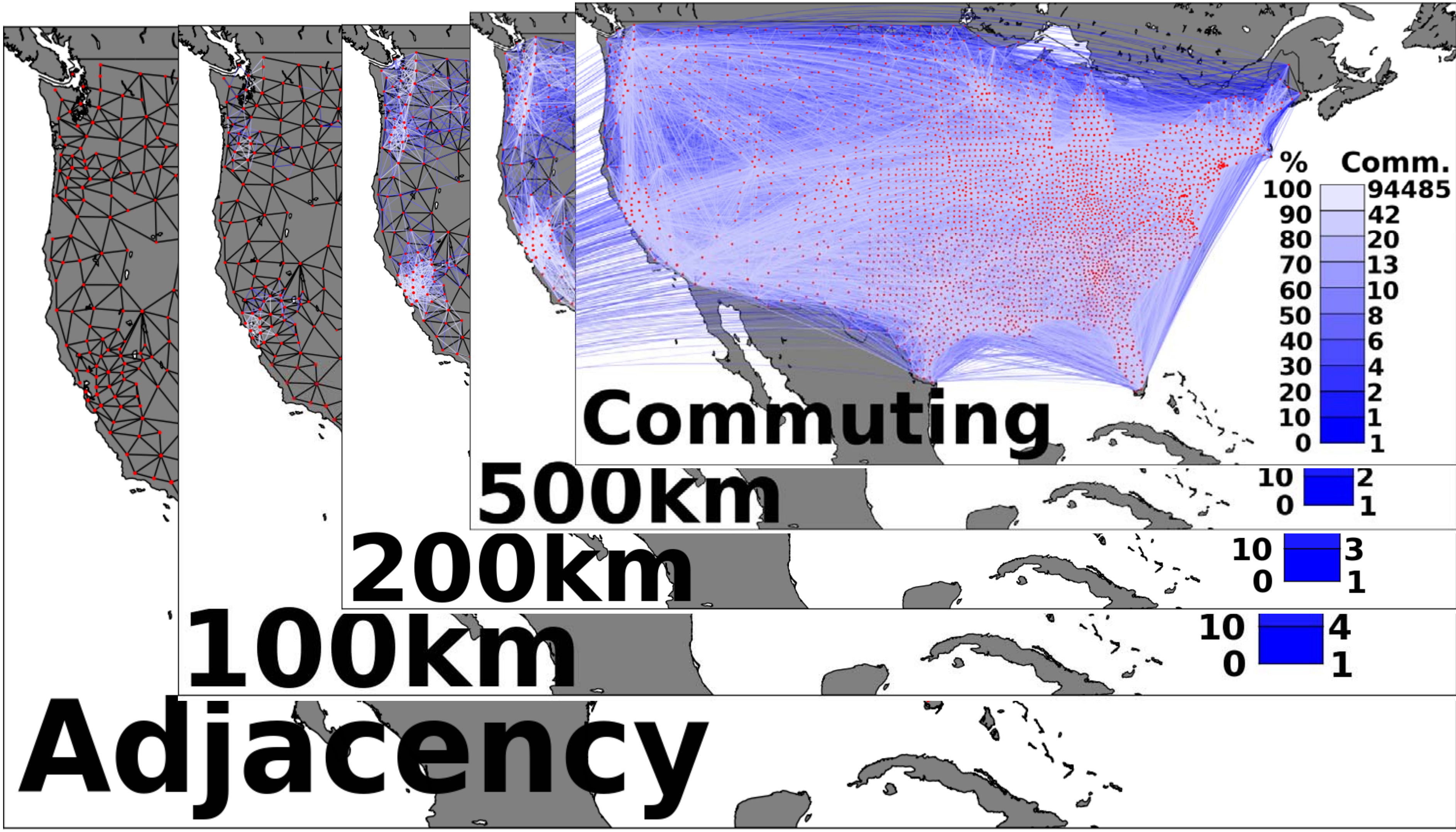

RANDOMIZED DATA

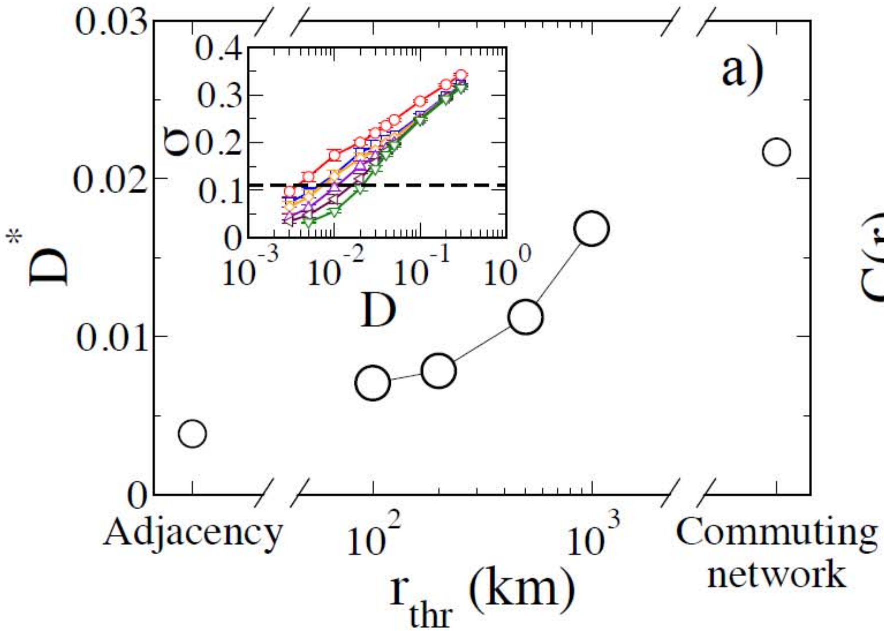


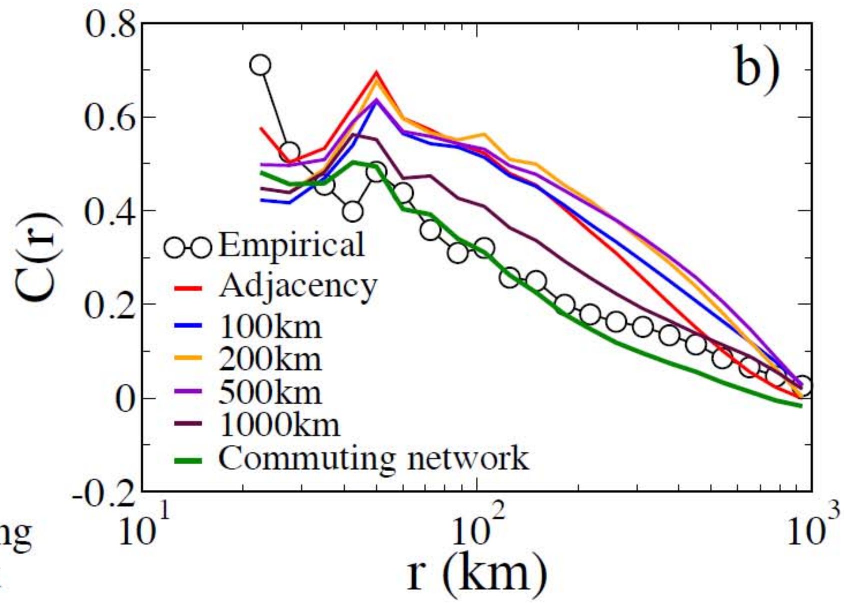




## MOBILITY RANGE






**Noise calibration for different mobility ranges**

**Spatial correlations in different networks:**

**LONG RANGE CORRELATIONS NEEDED TO REPRODUCE DATA**

## What have we learnt?

- **IBM implementation of a microscopic mechanism leading to *diffusive mesoscopic stochastic* dynamics reproducing statistical regularities of election data.**
- Data Based Modeling: Input parameters from census data for populations and commuting fluxes.
- Single calibrated model parameter: D, the noise intensity. Also calibration of time scale.
- **What do we explain?**
  -Two generic features in the background of election results:
  i)Stationarity of the dispersion of vote shares and ii) the time persistent logarithmic decay of spatial correlations.
  -Spatiotemporal fluctuations in electoral results at **different length scales**
  -***No attempt to predict electoral results***
- **Open question:** Role of network (dimensionality, range and weights of links), vs noise (imperfect imitation) in spatial log correlations.