Spreading processes on networks

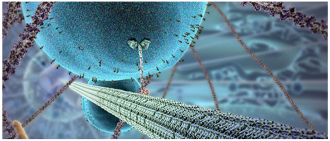
Liubov Tupikina, Marc Santolini

CRI, Bell labs

Big data course

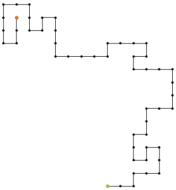
December 3, 2019

Spreading processes in nature



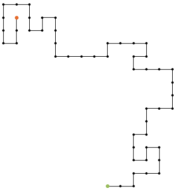
Transport of macromolecules, organelles and vesicles in living cells is a very complicated process that essentially determines and controls many biochemical reactions, growth and functioning of cells.

What we will talk about?

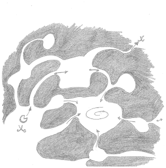


- **Part 1:**
  Spreading processes and diffusion

What we will talk about?

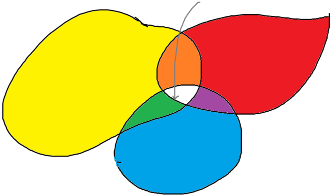


- **Part 1:**
  Spreading processes and diffusion



- **Part 2:**
  Spreading processes on networks

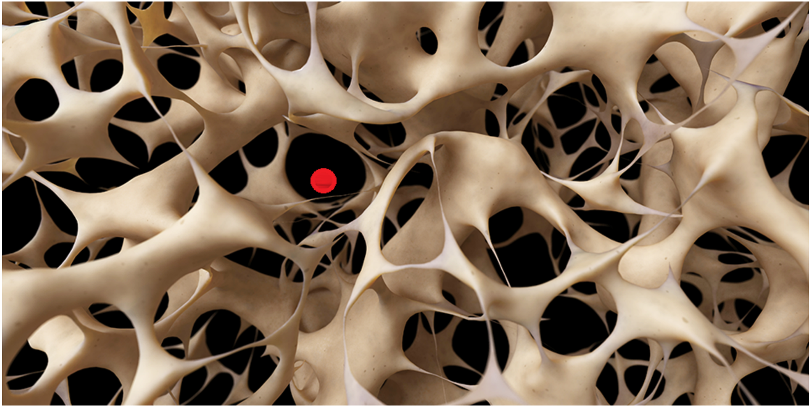
Scientists working on random walk theory:
Pearson, Smoluchowski, Langevin, Fermat, Bernoulli, Einstein, Kolmogorov, Pascal, Hughes, Brown, Wiener, Montroll, Weiss...



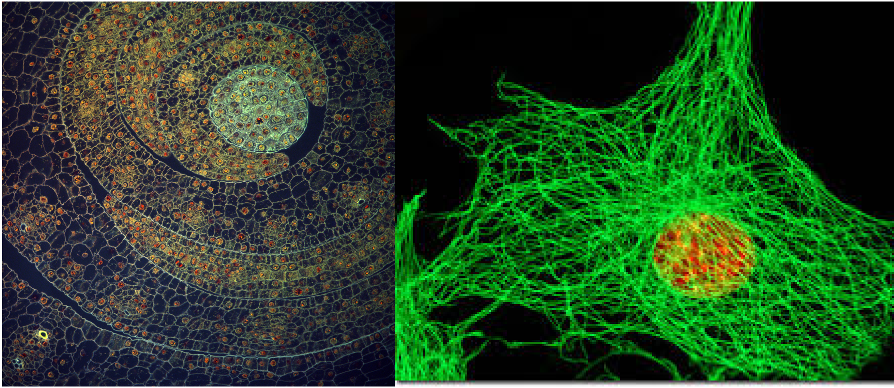
Intersection of Markov chain theory, Stochastic processes, Statistical physics, Disordered systems, Probability theory...

Real-world examples

Real-world examples



Real-world examples

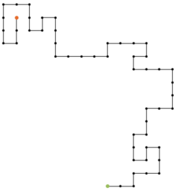


E.Katrukha,microtubules,imaginarycellrepresentation

`networkrepository/openflights`


©BabyRoutes.co.uk

Part 1



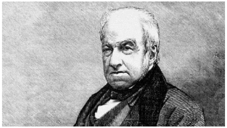
- **Spreading processes and diffusion**
  history
  random walks
  diffusion models

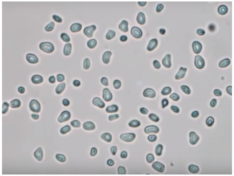
Spreading processes and diffusion: history

Experiment:

Brownian motion was discovered in 1785 by Jan Ingenhousz: irregular motion of coal dust particles on the surface.

In 1827 Robert Brown, a British botanist, is observing a suspended pollen grain in water.





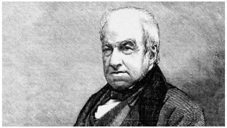
Spreading processes and diffusion: history

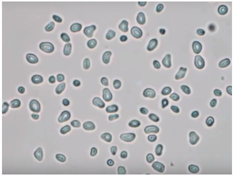
Experiment:

Brownian motion was discovered in 1785 by Jan Ingenhousz: irregular motion of coal dust particles on the surface.

In 1827 Robert Brown, a British botanist, is observing a suspended pollen grain in water. (video)

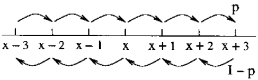
How to explain the experiment?

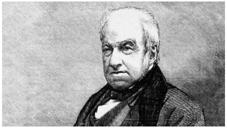


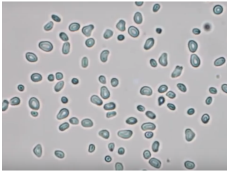


Random walks: How to explain the experiment?

Assumptions:



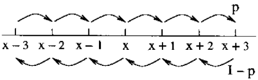


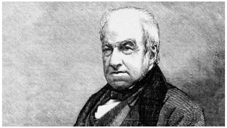


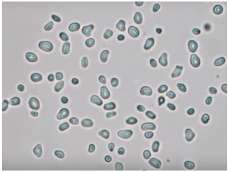
Random walks: How to explain the experiment?

Assumptions:

1. Random walk in 1 dimension
2. Memoryless: random walk forgets what he did last step
3. Symmetry: walk to the left of right with equal probability



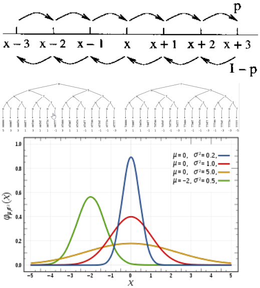




Diffusion: building a model

Assumptions:

1. Random walk in 1 dimension
2. Memoryless: random walk forgets what he did last step
3. Symmetry: random walk jumps to the left of right with equal probabilities p, q

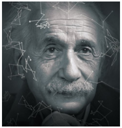


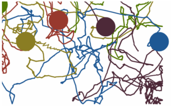
Examples of diffusion

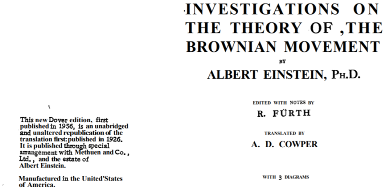
In 1905 the paper "About particles suspended in the liquid(...)" appeared. Molecular motion was suggested the explanation.

Diffusion equations:

$$\frac{dT}{dt} = D\frac{d^2T}{dx^2} \rightarrow D(x,t)\frac{d^2T}{dx^2}$$







This new Dover edition, first published in 1956, is an unabridged and unaltered republication of the translation first published in 1926. It is published through special arrangement with Methuen and Co., Ltd., and the estate of Albert Einstein.

Manufactured in the United States of America.

INVESTIGATIONS ON THE THEORY OF THE BROWNIAN MOVEMENT

BY

ALBERT EINSTEIN, PH.D.

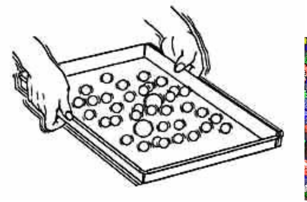
EDITED WITH NOTES BY

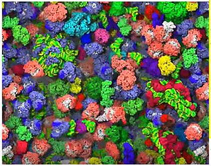
R. FÜRTH

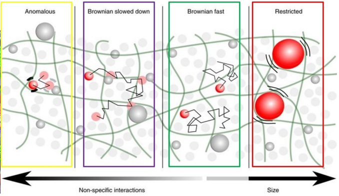
TRANSLATED BY

A. D. COWPER

WITH 3 DIAGRAMS



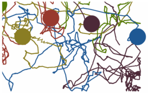




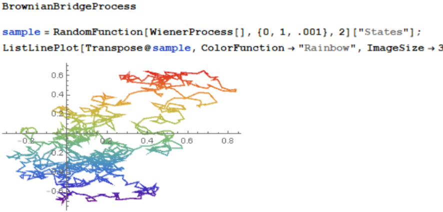
• Simulations of the crowded cell environment [McGuffee, Elcock, 2010] (video)

• Estimation of the properties for biological cells (first passage times distribution for the medical substance, etc.) [Klafter, Sokolov, 2010]

Numerical simulations



One can validate your hypothesis using numerical simulations.



Wolfram Mathematica random walks

Random walks properties

Probability density $P(x, t)$ is probability that we will find random walk in x at time t

First-passage time $F(x, t)$ is probability that at time t random walk is at x for the first time [Metzler et al., Redner, 2002]

Network measures based on random walks

Check reversed classroom topics (something, which is NOT in wikipedia)

Random walks: discrete space and time

Master equation describing probability

$P(x, n) = pP(x-1, n-1) + qP(x+1, n-1)$

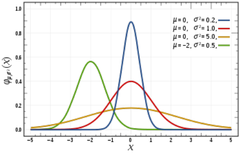
using the generating function and Fourier **transformation**

$P(k, z) = \sum_{n=0}^{\infty} z^n \sum_{x=-\infty}^{\infty} e^{ikx} P(x, n)$

We get $P(k, z) = 1/(1 - z(pe^{ik} + qe^{ik}))$

Then the final probability distribution for DTRW

$P(x, n) = \frac{N!}{(N+x)/2!(N-x)/2!} p^{(N+x)/2}(1-p)^{(N-x)/2}$



Gaussian in the long-time limit

$P(x, n) \to \frac{1}{\sqrt{2\pi npq}} e^{-(x-np)^2/2npq}$

Why do we need numerical simulations?

Continuous time random walk in 1D

At each x, where CTRW jumps, RW waits for τ distributed with $\psi(t)$.

Probability that no event occurred till time t:

$P^0(t) = \int_t^\infty \psi(t')dt'$

Probability that one event occurred till time t:

$P^1(t) = \int_t^\infty \psi(t')p(0, t - t')dt'$

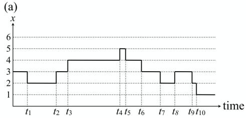
Apply trick to time: $t \to s$

$\tilde{P^1}(s) = \tilde{\psi}(s)\tilde{P^0}(s)$

...

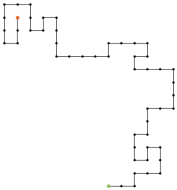
Probability that n **events** occurred till time t:

$\tilde{P^n}(s) = \tilde{\psi}^n(s)\tilde{P}^{n-1}(s) = \tilde{\psi}^n(s)\frac{1-\tilde{\psi}(s)}{s}$



[N.Masuda et al. 2017]

What we will talk about?



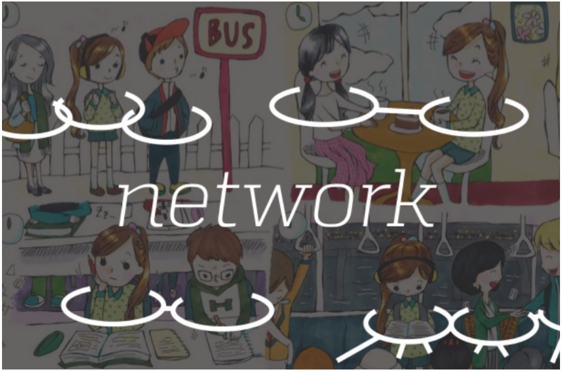
- **Part 1:**
  Spreading processes

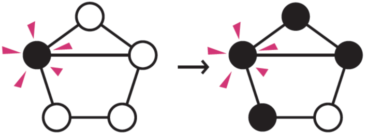
- **Part 2:**:
  Spreading processes on networks

Where can we find spreading processes?

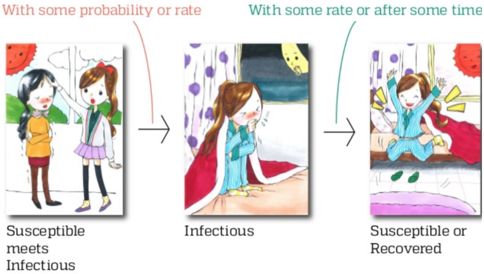
Spreading everywhere

P.Holme petterhol.me





P.Holme, et al, NJP (2018)
Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, Alessandro Vespignanis, Epidemic processes in complex networks
https://arxiv.org/abs/1408.2701



P.Holme `petterhol.me`

Reminder from the previous lecture

Types of epidemics spreading models:

$$S \xrightarrow{\beta} I$$

Types of epidemics spreading models

Types of epidemics spreading models:

$$S \xrightarrow{\beta} I$$

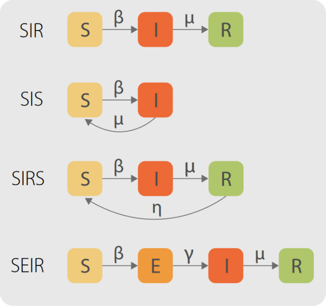
$$S \xrightarrow{\beta} I \xrightarrow{\mu} S$$

Nodes states:
S (susceptible), **I** (infected), **R** (recovered)
SI, SIS, SIR, SIRS, SIER...
Wang et al. SRDS 2003, Gómez et al. EPL 2010

Epidemics spreading SIR model

Model equations for:

SI (viral spreading), SIS, SIR model see notebook

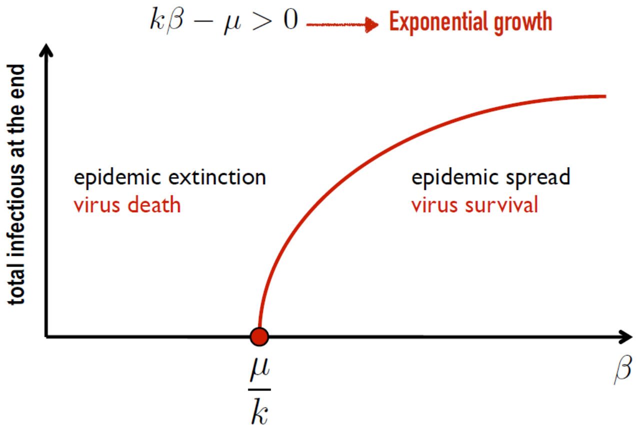


$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$

$$\frac{dR}{dt} = \gamma I.$$

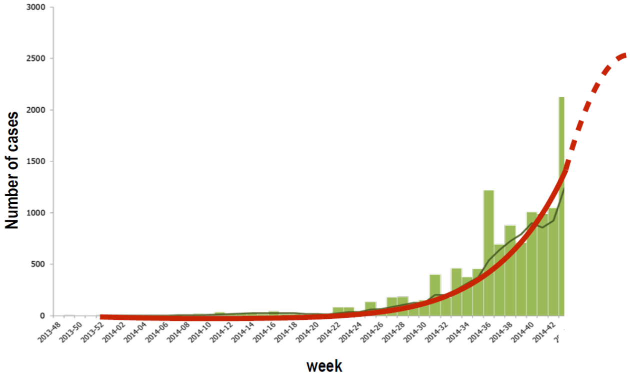
Why do we need to study mathematical models?



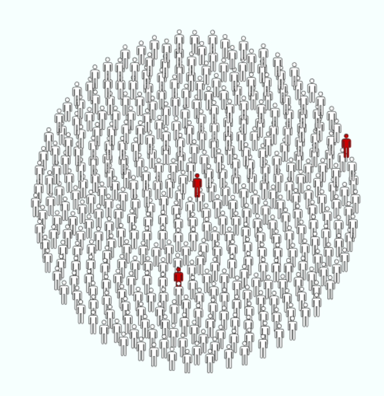
Epidemics threshold calculation in the case of SIR model

Why do we need to study mathematical models?

$k\beta - \mu > 0$ → Exponential growth

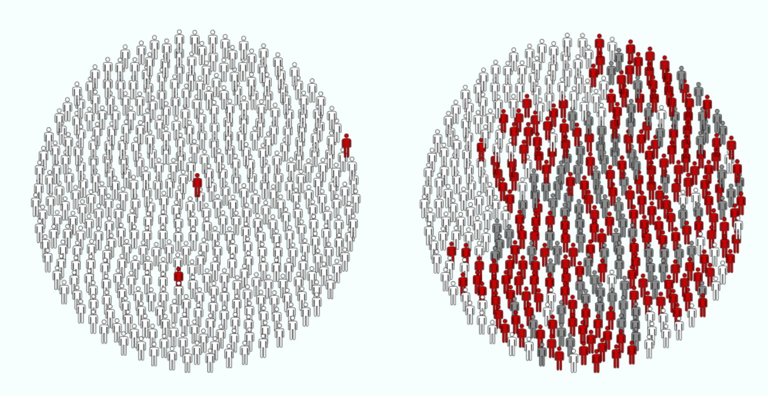


Epidemics outbreak predictions: Predicting Epidemic Risk from Past Temporal Contact Data: E. Valdano, C.Poletto, et al. Plos One 2015



Play yourself with different parameters...
`http://www.complexity-explorables.org/explorables/epidemonic/`

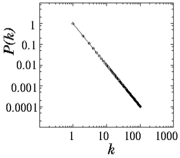
What if we epidemics spreads on a network?



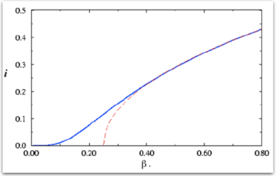
http://www.complexity-explorables.org/explorables/epidemonic/

What if we epidemics spreads on a network?

many real networks have
fat tail distribution



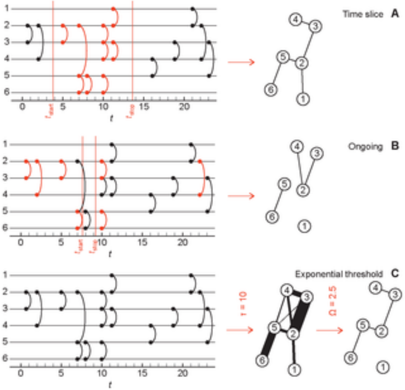
$$\beta_c \simeq \mu \frac{\langle k \rangle}{\langle k^2 \rangle} \longrightarrow \mathbf{0}$$



http://www.complexity-explorables.org/explorables/epidemonic/

Spreading on temporal networks

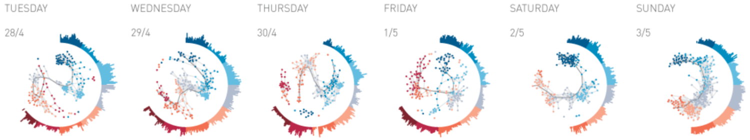
What if epidemics spreads on a temporal network?
Will infection spread faster?



Temporal networks, Saramaki, Holme,
`https://arxiv.org/abs/1108.1780`

Spreading on temporal networks

What if epidemics spreads on a temporal network? Will infection spread faster?

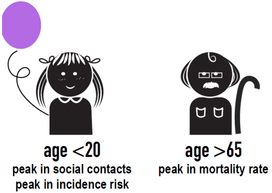


`sociopatterns.org`

Open questions

What are other topics of research in epidemics spreading?

Vaccination projects `jogl.io`



Influenza

- ~ billion cases annually (~3-5 million severe illness)
- ~300,000-500,000 deaths
- vaccine effectiveness:
23% (in 2014-15)
47% (in 2015-16)

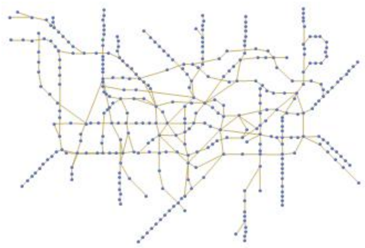
who should be vaccinated first?

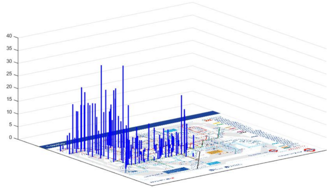
can we increase the impact of vaccination?

Applications: spreading on networks

Try notebooks https://github.com/Big-data-course-CRI/materials_big_data_cri_2019/

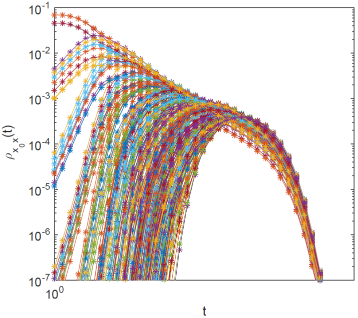
Applications: spreading on networks





Random walks properties $\rho_{x_0 x}(t)$ on a network of London metro.

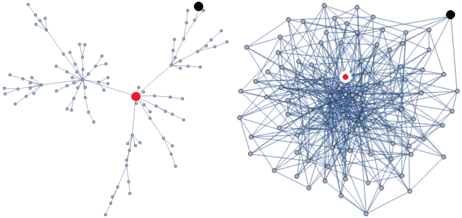
Topics for your network projects



- PageRank and other network measures
- Community detection algorithms
- Spreading processes on networks

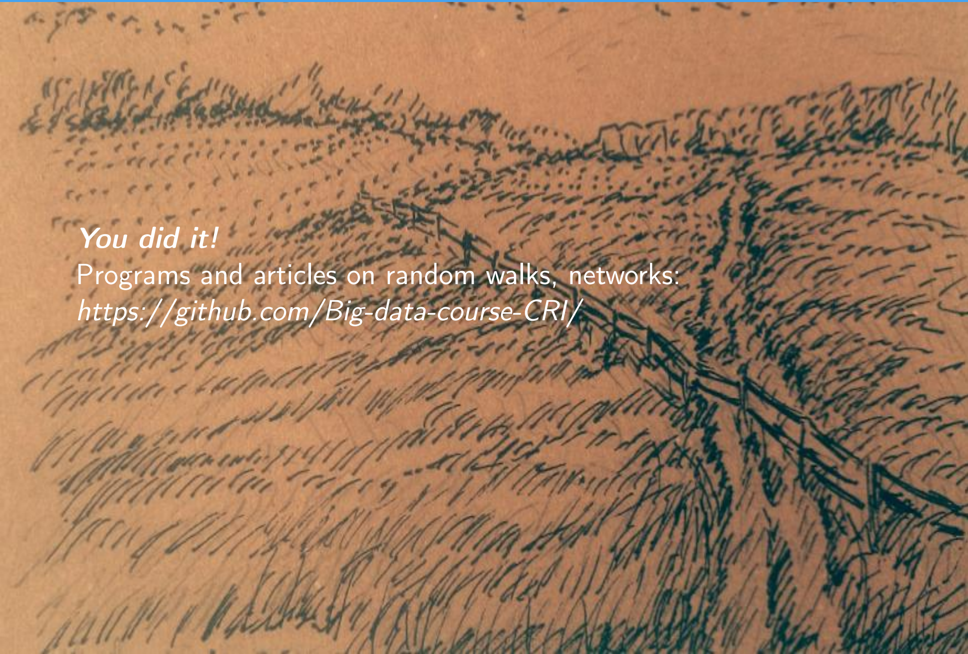
Hands-on part

Looking at notebooks



For notebooks go to:
`https://github.com/Big-data-course-CRI/`

Slide after all the maths



You did it!

Programs and articles on random walks, networks:

https://github.com/Big-data-course-CRI/

Appendix: Applications of random walk theory

What happens if we violate the assumptions?

1. Random walk in $R \to$ in dimension R^N
2. Memoryless $\to$ with memory: random walk remembers what he did last step
3. Non-Symmetry: random walk jumps to the left of right with different probabilities
4. New one??

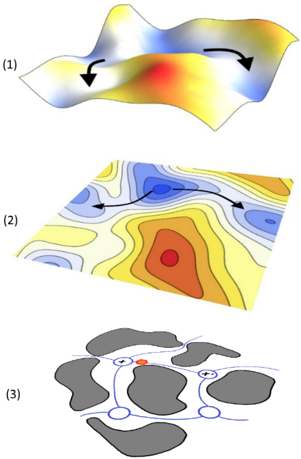
Fractal trajectories of random walk in $2D$

Main trick

$$P(x,t) = \frac{e^{-(x-vt)^2/4Dt}}{\sqrt{4\pi D t}} \qquad\qquad P(k,t) = e^{-(ikv + Dk^2 t)}$$

$$\begin{array}{ccc}
(x,t) & \overset{\text{Fourier}}{\Longleftrightarrow} & (k,t) \\
\Updownarrow & & \Updownarrow \text{ Laplace} \\
(x,s) & \Longleftrightarrow & (k,s)
\end{array}$$

$$P(x,s) = \frac{e^{-(v \mp \sqrt{v^2 + 4Ds})\,|x|/2D}}{\sqrt{v^2 + 4Ds}} \qquad\qquad P(k,s) = \frac{1}{s + ikv + Dk^2}$$


(1)
(2)
(3)

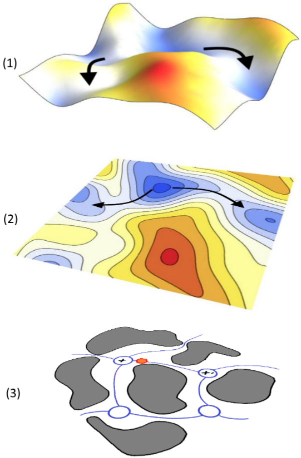
Heterogeneous Continuous Time Random Walk model

on a graph:

- graph with transition **matrix** Q,
- heterogeneous **travel time distributions** $\psi_{xx'}(t)$ between nodes x, x'.

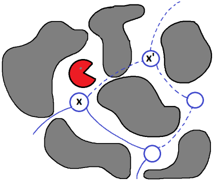
The generalized transition matrix

$$Q_{xx'}(t) = Q_{xx'}\psi_{xx'}(t)$$



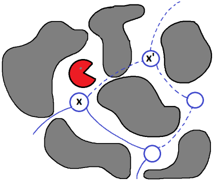
HCTRW is a formalism for studying diffusion in heterogeneous structures (1),(2).

Analytical results for HCTRW model



HCTRW on a graph:
1. graph with transition matrix Q,
2. travel times $\psi_{xx'}(t)$ between nodes x, x'

Analytical results for HCTRW model



HCTRW on a graph:

1. graph with transition matrix Q,
2. travel times $\psi_{xx'}(t)$ between nodes x, x'

Analytic formula for HCTRW propagator $\tilde{P}_{x_0 x}(s)$:

$$\boxed{\tilde{P}_{x_0 x}(s) = \frac{1 - \sum_{x'} \tilde{Q}(s)_{xx'}}{s} [(I - \tilde{Q}(s))^{-1}]_{x_0 x}}, \qquad (1)$$

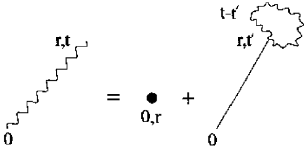
where $\tilde{Q}_{xx'}(s) = Q_{xx'} \tilde{\psi}_{xx'}(s)$.

Long-time behaviour for HCTRW propagator $P_{x_0 x}(t)$:

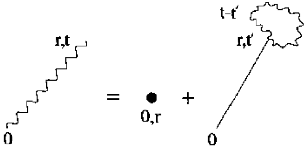
$$\boxed{\tilde{P}_{x_0 x}(s) \simeq \underbrace{p_{st}(x)}_{stat.dist.} + t(x) \sum_{k>0} \frac{u^{\dagger}_{0k}(x)\, v_{0k}(x_0)}{\lambda_{0k} + s\lambda_{1k}}} \qquad (2)$$

λ_{0k} are eigenvalues, u_{0k}, v_{0k} eigenvectors of $I - Q$.
$\lambda_{0k} + s\lambda_{1k}$ is the 1^{st} order correction for $I - Q + sT$,
$T_{xx'} = Q_{xx'} \langle \mathcal{T}_{xx'} \rangle$, $t(x) = \sum_{x'} T_{xx'}$.

First passage time

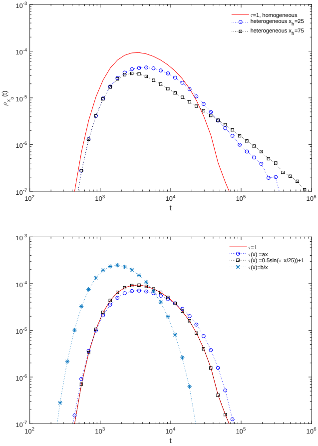


First passage time



$F(x,t)$ - first passage time probability, then using renewal approach, [Redner, 2002]:
$P(x,t) = \delta_{x,0}\delta_{t_0} + \sum_{t' \neq t} F(r,t')P(0,t-t')$

Results for Heterogeneous random walks



The FPT density of HCTRW on an interval, absorbing $x = 100$: (top) hetero-nodes $x_h = 25, 75$ with heavy-t.distr. $\alpha = 0.5$; (bottom) $\tau_{\pm} = 1$; $\tau_{xx\pm\mathbf{1}} = ax$; $\tau_{xx\pm\mathbf{1}} = 0.5\sin(\pi x/25) + 1$; $\tau_{xx\pm\mathbf{1}} = b/x$.

Analytic formula for HCTRW propagator $P_{x_{\mathbf{0}}x}(t)$ on a graph is derived.

Analytic framework of HCTRW links structural graph properties and dynamical RW properties

HCTRW framework allows of study asymptotic solutions, FPT for processes on graphs.

- Grebenkov, Tupikina, "Heterogeneous continuous time random walks", PRE 012148 (2018)

Random walks: discrete space and time

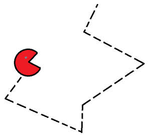
Will we get Gaussian for any hopping process, where $\langle x^2 \rangle$ is finite?

Random walks: discrete space and time

Will we get Gaussian for any hopping process, where $\langle x^2 \rangle$ is finite?
Yes! Probability distribution is independent from details of single-step hopping [Gnedenko, Kolmogorov 1954]

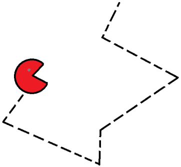
Types of random walks

- Continuous or discrete (networks) underlying system [Redner, 2002]
- Discrete or continuous time [Montroll, Weiss, 1965]
- Probability distributions of jumps from each node [Hudges, 2002]
- Random walks with resetting, self-avoiding random walk, adaptive random walks... [Masuda, Lambiotte et al.2017]



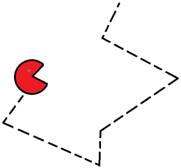
Random walks: continuous space and time

Continuous time random walk model (CTRW): CTRW waits at each x, where CTRW jumps, it waits time step for τ distributed with $\psi(t)$. The length of the steps are distributed with $f(x)$ function.



Random walks: continuous space and time

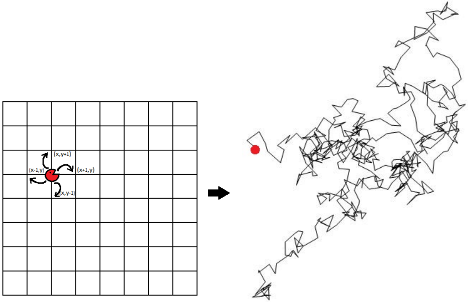
Continuous time random walk model (CTRW) $P(x, t)$ is probability to find RW at x coordinate at time t.
We can use the probability $P(x, n)$ from discrete time random walk
$P(x, t) = \sum_{n=0}^{\infty} P(x, n)p(n, t)$
Then using **renewal theory, Green function** we get the final equation
$P(k, s) = \frac{1-\tilde{\psi}(s)}{s} \frac{1}{1-f(k)\tilde{\psi}(s)}$



Some references

1. L. Tupikina, K. Rehfeld, N. Molkenthin, V. Stolbova, N. Marwan, and J. Kurths, "Characterizing the evolution of climate networks", Nonlin. Processes Geophys., 21, 705-711 (2014)
2. H. Bunina, L. Tupikina, "Authomorphisms of the semigroup of nonnegative invertible matrices of order 2 over rings", Journal of Mathematical Sciences, 183, 305-313 (2012)
3. N. Molkenthin, K. Rehfeld V. Stolbova L. Tupikina and J. Kurths, "On the influence of spatial sampling on climate networks", Nonlin. Processes Geophys., 21, 651-657 (2014)
4. J.F. Donges, J. Heitzig, B. Beronov, M. Wiedermann, J. Runge, Q.-Y. Feng, L. Tupikina, V. Stolbova, R.V. Donner, N. Marwan, H.A. Dijkstra, and J. Kurths, "Unified functional network and nonlinear time series analysis for complex systems science: The pyunicorn package", Chaos 25, 113101-1-25 (2015)
5. L.Tupikina, N. Molkenthin, C.Lopez, E.Hernandez-Garca, N.Marwan and J.Kurths, "Correlation networks from flows. The case of forced and time-dependent advection-diffusion dynamics", Plos One, 0153703 (2016)
6. H.Kutza, N. Molkenthin, L. Tupikina, N.Marwan, J.Donges, U.Feudel, J. Kurths, R.Donner, "A geometric perspective on spatially embedded networks. Quantification of edge anisotropy and application to flow networks", Chaos 27, 035802 (2017)
7. **L.Tupikina**, "Dynamics on networks: case of Heterogeneous opinion state model (on graphs)", arxiv 1708.01647
8. P.Holme, L.Tupikina "Explicit solutions for SIS model" arxiv.org 1802.08849
9. R.Donner, M.Lindner, N.Molkenthin, **L.Tupikina** "From nonlinear dynamics to complex systems: A Mathematical modeling approach" Springer book chapter, Cham, pp. 197-226 (2019)
10. D.Grebenkov, **L.Tupikina**, "Heterogeneous continuous time random walk" Phys. Rev. E 97, 012148 (2018)
11. **L.Tupikina**, D.Grebenkov "Continuous limits of Heterogeneous continuous time random walk", (in prep.)
12. M. Kang, **L. Tupikina**, G. Berthelot, C. Nicolaides, D. Grebenkov, B. Sapoval "Universality of resistor scale-free networks" (in prep.)
13. M.Stella, M.Porter, L.Tupikina, C. Cramer, "Multilayer network between science and education", Proceedings of NetSci (2018)
14. F.Caravelli, T. Cui, **L.Tupikina** "A perspective on graph spectra", book chapter, lecture notes (online)

Random walks on other systems



Homogeneous random walk (RW) on a discrete lattice in discrete time

Homogeneous continuous time random walk (CTRW) in continuous space

R.Metzler, J.Klafter, Phys.Rep. [2000]

Main Collaborators

Nora Molkenthin,
Jonathan Donges
Cristóbal López,
Emilio Hernández-García,
Henk Dijkstra,
Norbert Marwan, Jobst Heitzig,
Jürgen Kurths, Petter Holme
Denis Grebenkov