Maximal Entropy Random Walk

the most random among random walks (maximizing entropy production)

https://en.wikipedia.org/wiki/Maximal_entropy_random_walk

RW/**information flow** for minimal information about a system
in agreement with the **maximum entropy principle**
strong localization property, **scale-free, nonlocal**

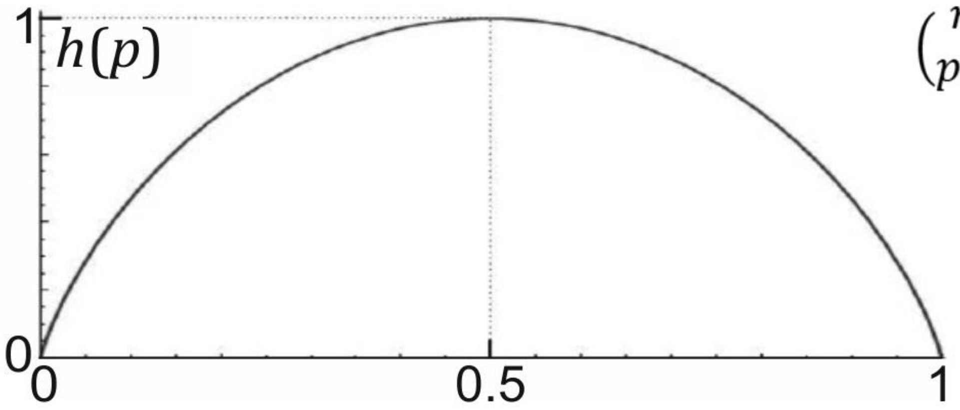
Some applications:

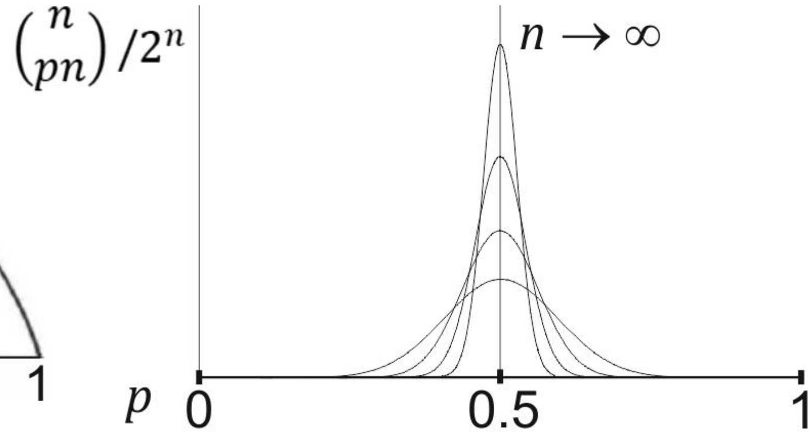
- **maximizing informational capacity** of channel under some constraints (data storage/transmission, maybe linguistics (?)),
- corrections to **diffusion models** to get agreement with quantum predictions (diffusion, conductance, molecular dynamics),
- **metrics for complex networks, data mining** (e.g. centrality measure, tractography, saliency regions, PageRank, SimRank, community detection)

Jarosław Duda, Kraków, UJ, 5. XII. 2016

We need n bits of information to choose one of 2^n possibilities.

For length n 0/1 sequences with pn of "1", how many bits we need to choose one?





$$\binom{n}{pn} \approx 2^{nh(p)} \qquad \text{for} \qquad h(p) := -p\lg(p) - (1-p)\lg(1-p)$$

A sequence of symbols with $(p_s)_{s=0..m-1}$ probability distribution contains asymptotically $\boldsymbol{H = \sum_s p_s \lg(1/p_s)}$ **bits/symbol** ($H \le \lg(m)$)

Seen as weighted average:

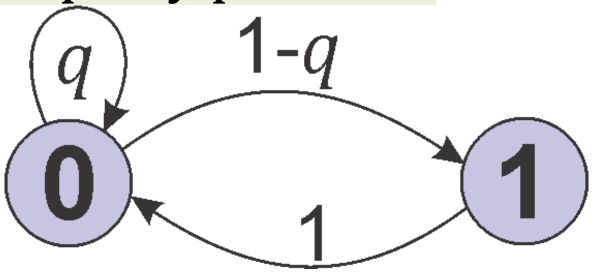
symbol/event of probability p contains $\lg(1/p)$ bits of information.

(Jaynes) principle of maximum entropy: while limited knowledge, the best assumption is probability distribution which maximizes entropy.

Fibonacci coding – as a bit sequence with **constraints**: no two neighboring ‘1’s
e.g. 0010101000010101001001 – each sequence should be equally probable

What about statistics of a single step?

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad S = \begin{pmatrix} q & 1-q \\ 1 & 0 \end{pmatrix} \qquad q = ?$$



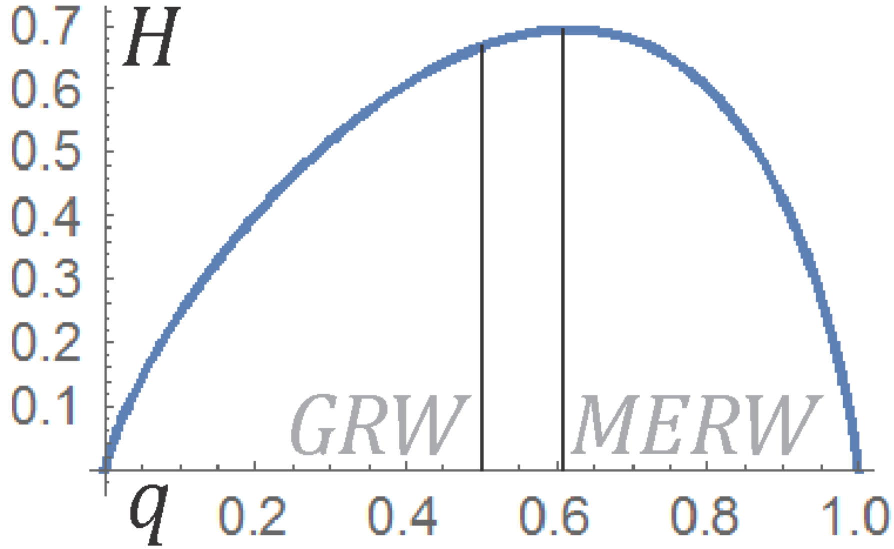
What q should we choose to maximize informational capacity?

Stationary probability: $\pi = (\Pr(0), \Pr(1))^T$

$$\pi S = \pi \qquad \pi = \left(\frac{1}{2-q}, 1 - \frac{1}{2-q}\right)$$

Entropy – informational content:

$$H = \sum_i \pi_i \sum_j S_{ij} \lg(1/S_{ij}) = \pi_0 \cdot h(q)$$



$H_{max} \approx 0.694241913$ bits/node

for $q = \frac{\sqrt{5}-1}{2} \approx 0.618034$ for MERW

($H = 0.6666\ldots$ for GRW)

($q = 0.5$ for GRW)

My original MERW motivation: **maximizing capacity under constraints** for 2D analogue of Fibonacci coding ("hard square": no two neighboring '1's)

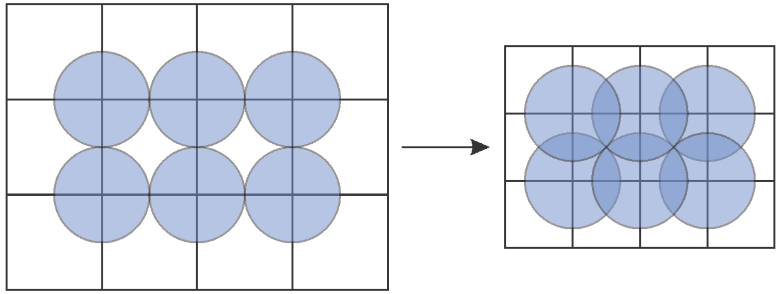
We get $\boldsymbol{H} \approx \mathbf{0.58789}$ bits/node

Some application:

use magnetic dots (twice) more densely,

at cost of constraints – two dots cannot overlap.

$$2 \cdot 0.58789 \approx 1.\mathbf{176}$$



We get 17.6% capacity increase due to better positioning!

(e.g. using 1D MERW on the space of possible successive lines)

Approximate with finite width stripe $\infty \times m$

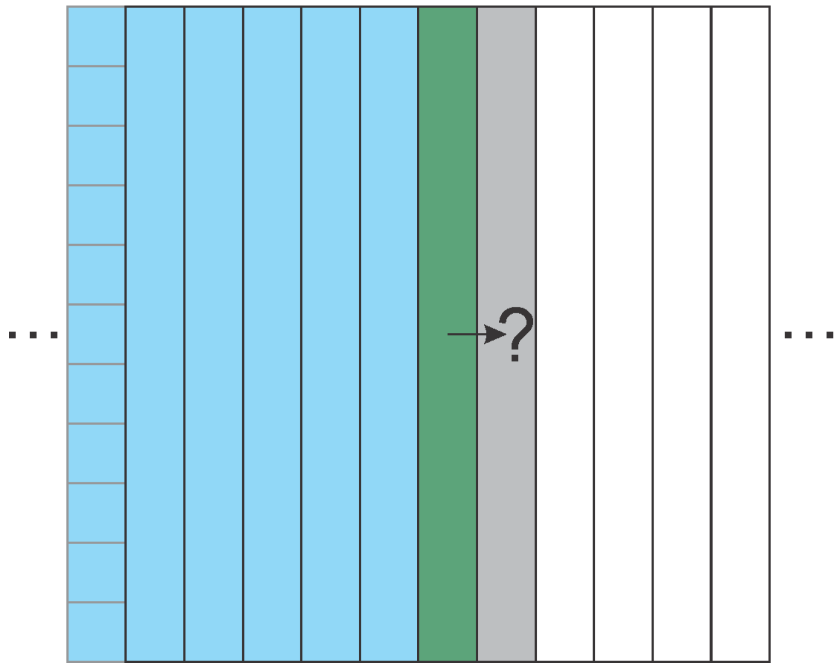
(large) **alphabet**: allowed slices

Adjacency matrix: possible neighbors

… **find MERW** for adjacency matrix… ?

→ translate into local transition probability rules

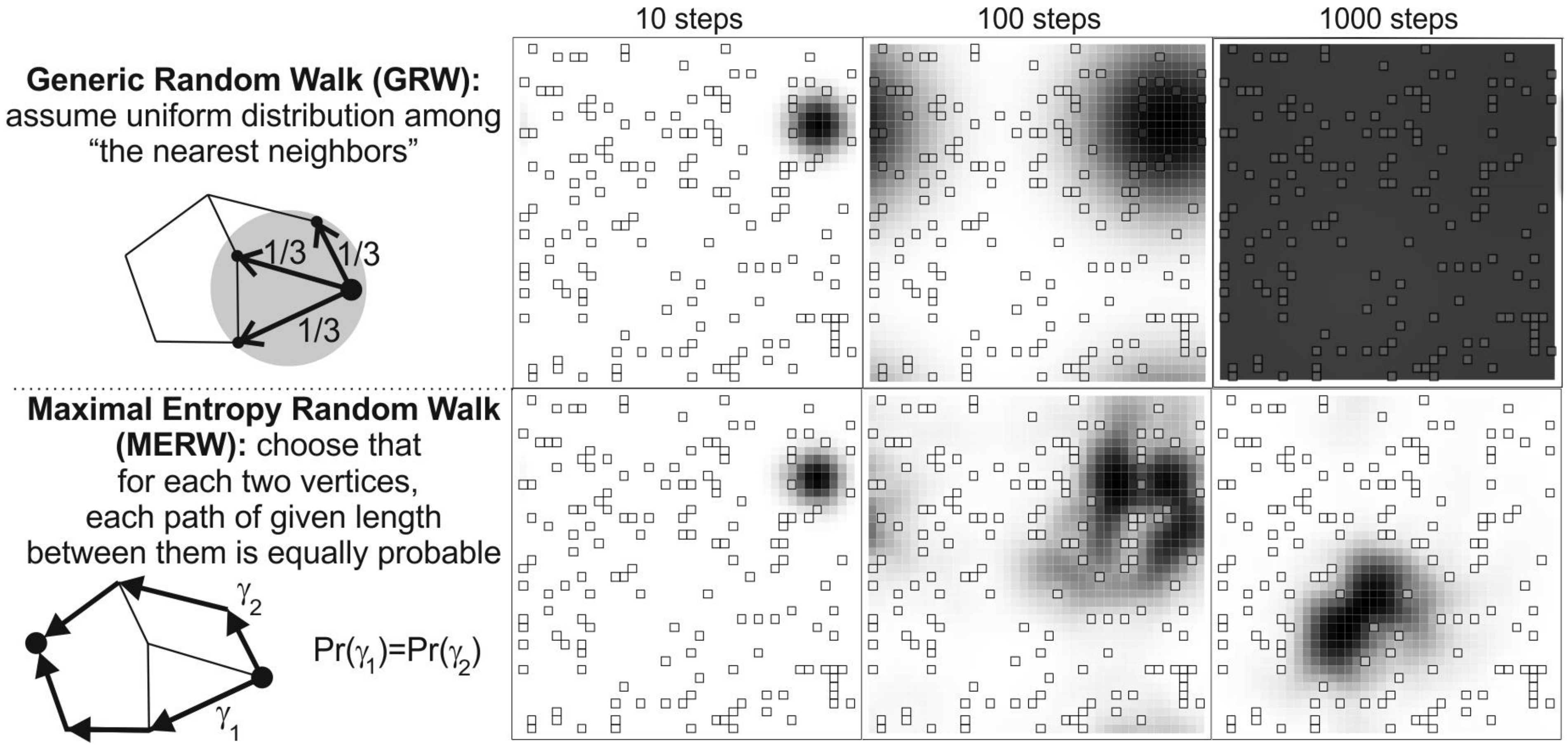
scanning lines, a few neighbors determine probability



Graph (M) $M_{ab} \in \{0,1\}$ → **stochastic matrix** (S) $0 \le S_{ab} \le M_{ab}$, $\forall_a \sum_b S_{ab} = 1$ → **stationary probability** (π) $\sum_a \pi_a S_{ab} = \pi_b$

Average **entropy** production per step: $\sum_a \pi_a \sum_b S_{ab} \lg(1/S_{ab})$

GRW and MERW are equal on regular graphs, but e.g. on defected 2D lattice:

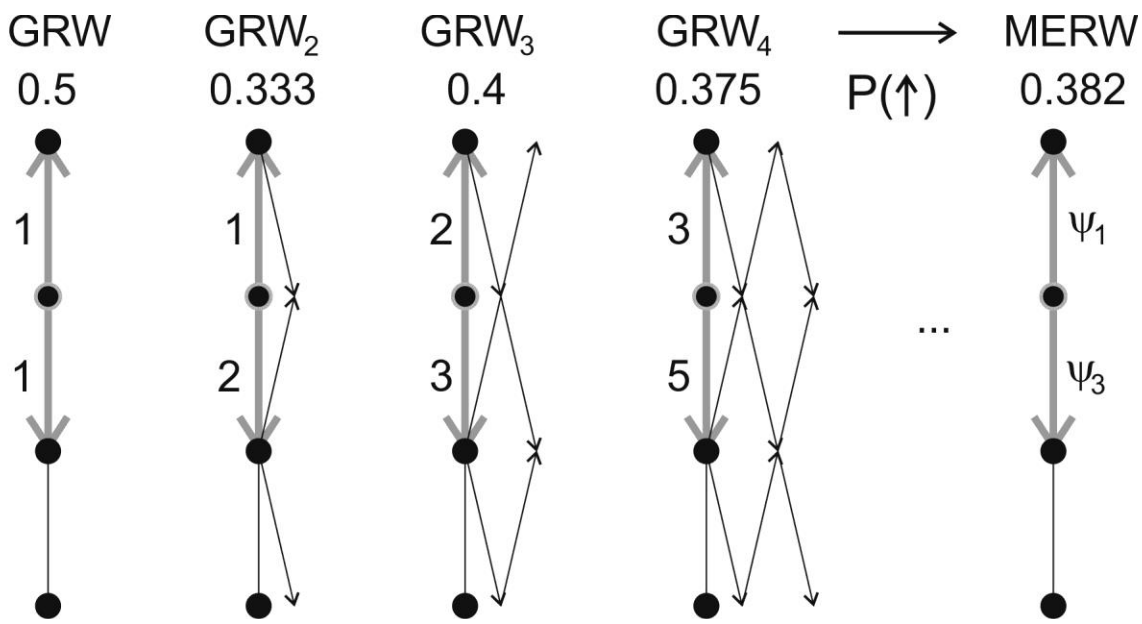


GRW assumes we know exactly the used probabilistic algorithm, has characteristic length

MERW assumes only there are no hidden local probabilistic rules, is scale-free limit of GRW

MERW as scale-free limit of GRW $S_{ab}^{GRW_k} \propto M_{ab} \sum_c (M^{k-1})_{bc}$

GRW: each outgoing **edge** is equally probable$(k=1)$



GRW_k – each outgoing **length** k **path is equally probable.**

In the limit, the number of paths starting with $a \to b$ **is proportional** to coordinate (ψ_b) of the **dominant eigenvector** of M :

$$M\psi = \lambda\psi$$

Frobenius-Perron theorem for connected graph: real, nondegenerated $\lambda > 0$, $\forall_a\, \psi_a > 0$

Normalization for vertex a: $\sum_b M_{ab}\psi_b = (M\psi)_a = \lambda\psi_a$

Finally: **while being in a, probability of jumping to b is:** $\boxed{S_{ab} = \frac{M_{ab}}{\lambda}\frac{\psi_b}{\psi_a}}$

(symmetric M:)

For which stationary probability distribution $(\pi S = \pi)$ is $\pi_a \propto \psi_a^2$ nonlocality

$$(\pi S)_b = \sum_a \psi_a^2 \cdot \frac{M_{ab}}{\lambda}\frac{\psi_b}{\psi_a} = \sum_a \psi_a M_{ab} \cdot \frac{\psi_b}{\lambda} = \lambda\psi_b \frac{\psi_b}{\lambda} = \psi_b^2 = \pi_b$$

$$\boxed{(S^k)_{ab} = \frac{(M^k)_{ab}}{\lambda^k}\frac{\psi_b}{\psi_a}}$$

MERW – uniform probability distribution among **paths**

GRW – all length 1 paths equally probable, **MERW** – all length $t \to \infty$

Assume **unique dominant eigenvalue** (Frobenius-Perron for connected aperiodic)

$M\psi = \lambda\psi \quad$ for largest λ

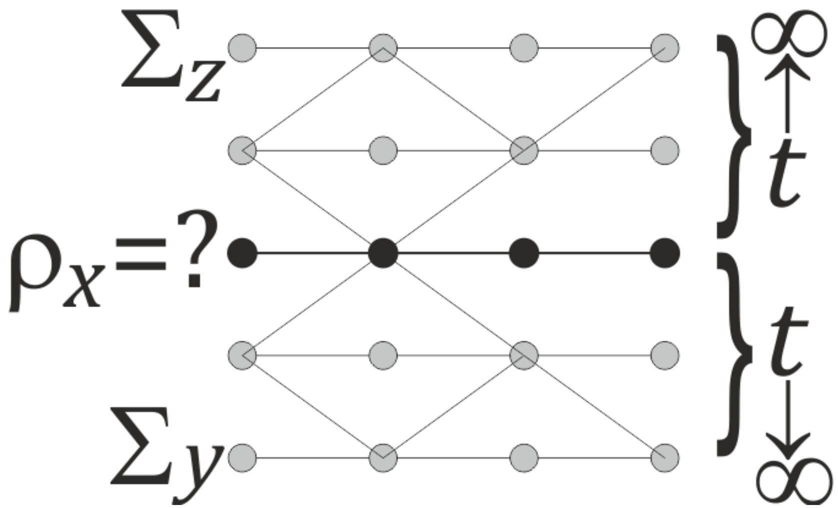
Hubbard: $\mathcal{H} = -\sum_{\text{edge } xy} a_y^+ a_x \equiv -M$

$M^t \approx \lambda^t \, |\psi\rangle\langle\psi| \qquad$ for $\quad t \to \infty$

The number of length $2t$ paths

with x in the center:

$$\sum_{yz} (M^t)_{yx}(M^t)_{xz} \approx \lambda^{2t}\psi_x^2$$



$$\boldsymbol{\rho_x \equiv \Pr(x) \propto \psi_x^2} \qquad \textbf{– Born rule}$$

Analogously $\Pr(xy) \propto \psi_x M_{xy} \psi_y \Rightarrow \boldsymbol{S_{xy} = \frac{M_{xy}}{\lambda}\frac{\psi_y}{\psi_x}}$ stochastic matrix

Generally: $\qquad \Pr(\gamma_0 \dots \gamma_l) = \frac{1}{\lambda^l}\frac{\psi_{\gamma 0}}{\psi_{\gamma l}} \qquad\qquad (S^l)_{xy} = \frac{(M^l)_{xy}}{\lambda^l}\frac{\psi_y}{\psi_x}$

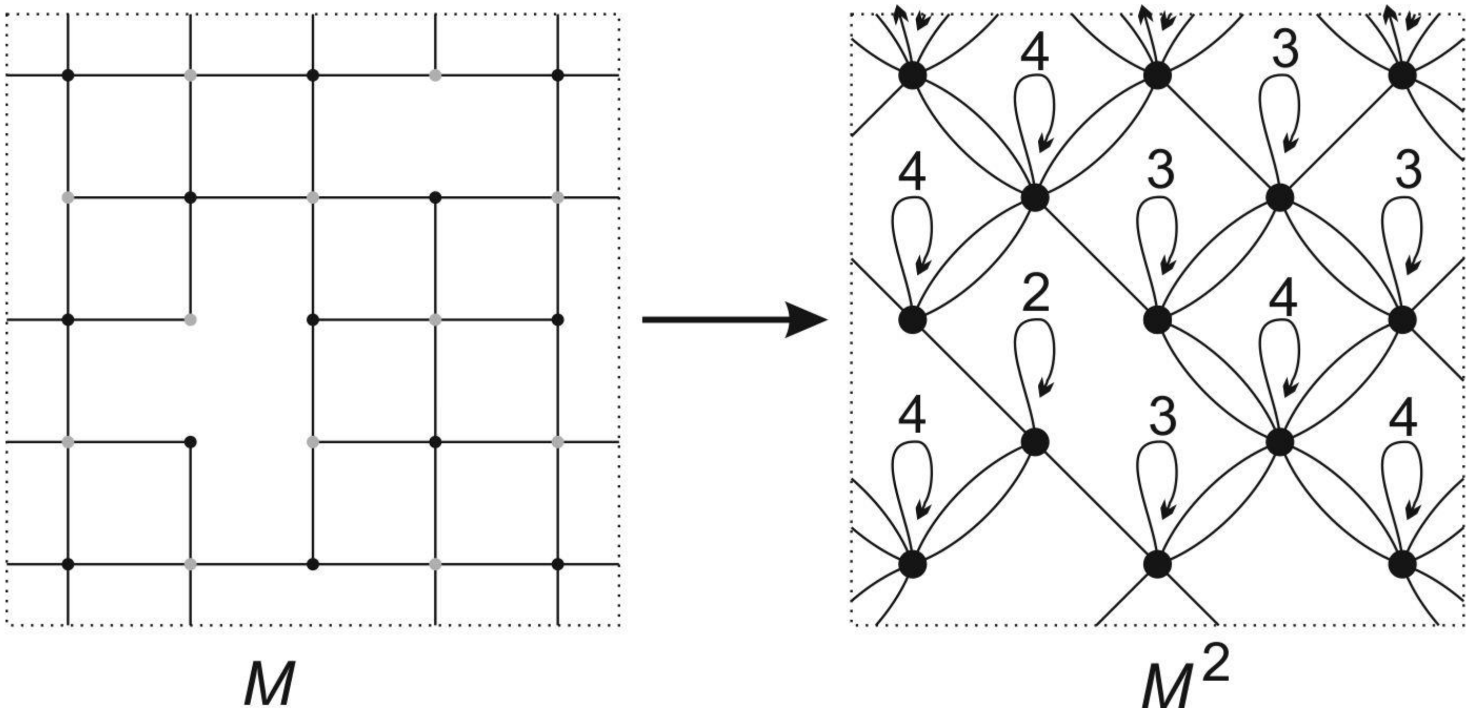
Renormalization (being **scale-free:** discretization independent**)**

Previously: time scale-free, now: space

Spatial: powers of the graph should have the same random walk

$$\left(\left(S^{\mathrm{MERW}(M)}\right)^l\right)_{ij} = \sum_{\gamma_1,\ldots,\gamma_{k-1}} \frac{M_{i\gamma_1}}{\lambda}\frac{\psi_{\gamma_1}}{\psi_i}\cdot\frac{M_{\gamma_1\gamma_2}}{\lambda}\frac{\psi_{\gamma_2}}{\psi_{\gamma_1}}\cdot\ldots\cdot\frac{M_{\gamma_{k-1}\gamma_k}}{\lambda}\frac{\psi_j}{\psi_{\gamma_{k-1}}} = \frac{(M^l)_{ij}}{\lambda^k}\frac{\psi_{\gamma_k}}{\psi_{\gamma_0}} = \left(S^{\mathrm{MERW}(M^l)}\right)_{ij}$$

Usually not true for GRW



Approximating MERW for short range knowledge (GRW_k)

Sinatra, R., Gomez-Gardenes, J., Lambiotte, R., Nicosia, V. & Latora, V. *Maximal-entropy random walks in complex networks with limited information*, Phys. Rev. E 83, 030103 (2011)

	GRW	GRW_2	GRW_3	$MERW$
	$\frac{h(\pi^0)}{h(\pi)}$	$\frac{h(\pi^1)}{h(\pi)}$	$\frac{h(\pi^2)}{h(\pi)}$	$h_{\max} = h(\pi)$
Regular lattice	1.000	1.000	1.000	1.79
Random regular graph	1.000	1.000	1.000	1.79
ER random graph	0.954	0.993	0.998	1.98
Uncorrelated scale-free $\gamma = 1.5$	0.886	0.992	0.996	2.36
BA model	0.825	0.976	0.996	2.52
Assortative scale-free $\gamma = 1.5$	0.876	0.991	0.999	2.44
Disassortative scale-free $\gamma = 1.5$	0.937	0.990	0.997	2.18
Regular lattice (1% defects)	0.996	0.997	0.998	1.38
Regular lattice (10% defects)	0.967	0.978	0.981	1.34
Regular lattice (20% defects)	0.931	0.955	0.963	1.29
Internet autonomous system [22]	0.744	0.900	0.980	4.10
U.S. Airports [18]	0.879	0.990	0.997	3.88
E-mail [23]	0.881	0.983	0.997	3.03
SCN (cond-mat) [24]	0.694	0.867	0.946	3.17
SCN (astro-ph) [24]	0.784	0.941	0.973	4.41
PGP [25]	0.597	0.920	0.976	3.75

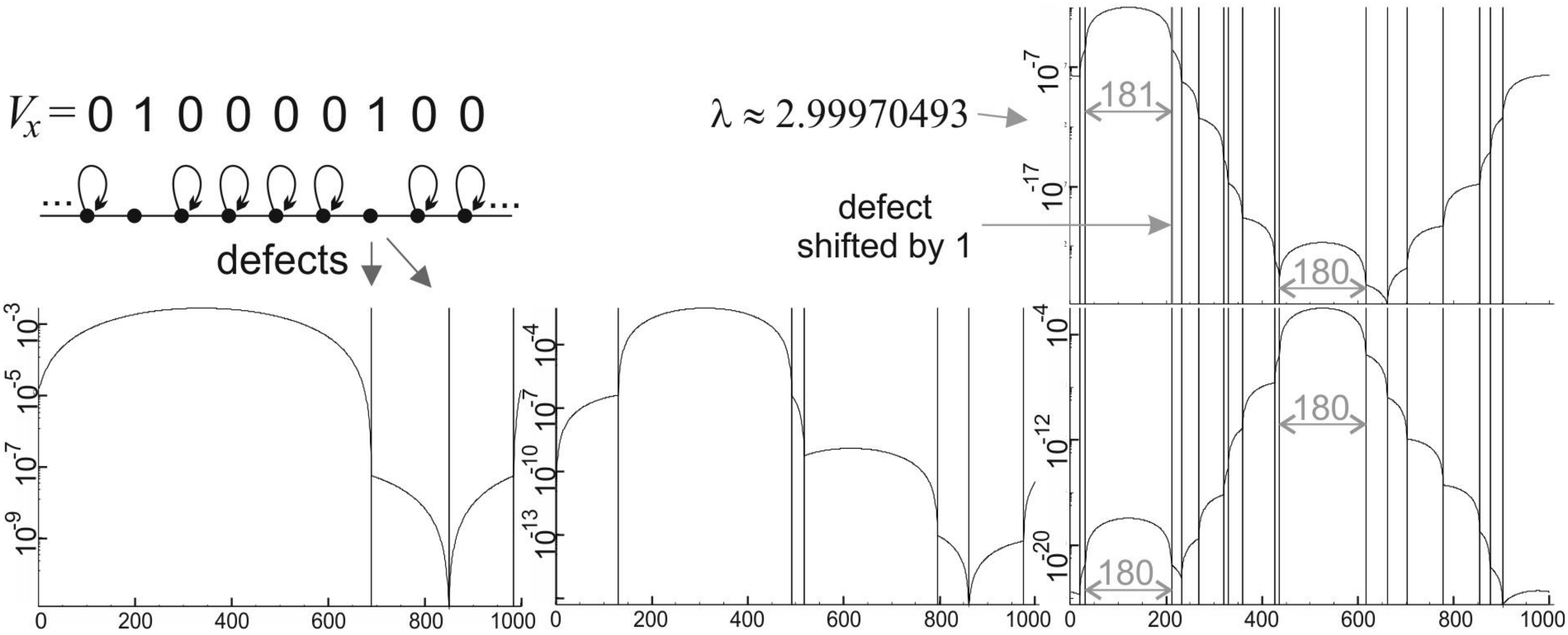
$h \to h_{max}$ but the behavior can be qualitatively different

GRW: stationary probability for undirected graph $\boxed{\propto d_i = \sum_j M_{ij}}$

MERW: stationary probability $\boxed{\rho \propto \psi^2}$ where $\boxed{M\psi = \lambda\psi}$ for largest λ

Defected	$(\lambda\psi)_x = (M\psi)_x = \psi_{x-1} + (1 - V_x)\psi_x + \psi_{x+1}$	$/-3\psi_x$	$/\cdot -1$
1D lattice	$E\psi_x = -(\psi_{x-1} - 2\psi_x + \psi_{x+1}) + V_x\psi_x$	for smallest	$E = 3 - \lambda$

Discrete stationary **Schrödinger equation**, **Nonlocal** – depends on the entire graph



p_d = 0.002 λ ≈ 2.99998025 p_d = 0.01 λ ≈ 2.99992549 p_d = 0.03 λ ≈ 2.99970172 (**diffusion**) A basic question for many complex systems:

what stationary probability density should we expect?

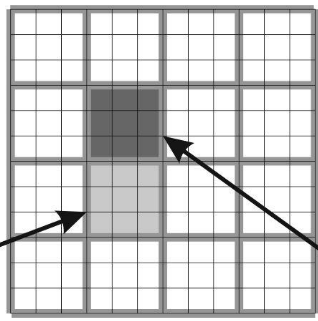
For example for electrons or solitons "hopping" between atoms in a lattice

two answers (should agree in intersection of applicability):

Quantum mechanics:	Diffusion:
Define energy density for given system: **Hamiltonian** $(\widehat{H})$, find its dominant eigenvector/eigenfunction (ψ):	Choose transition probabilities – - stochastic matrix/operator $(\hat{S})$, and ask for its stationary density: dominant eigenvector/eigenfunction
$\widehat{H}\psi = \lambda\psi, \quad \rho = \lvert\psi\rvert^2$	$\rho\hat{S} = \rho$
Strong localization property (e.g. Anderson's)	Usually weak localization property

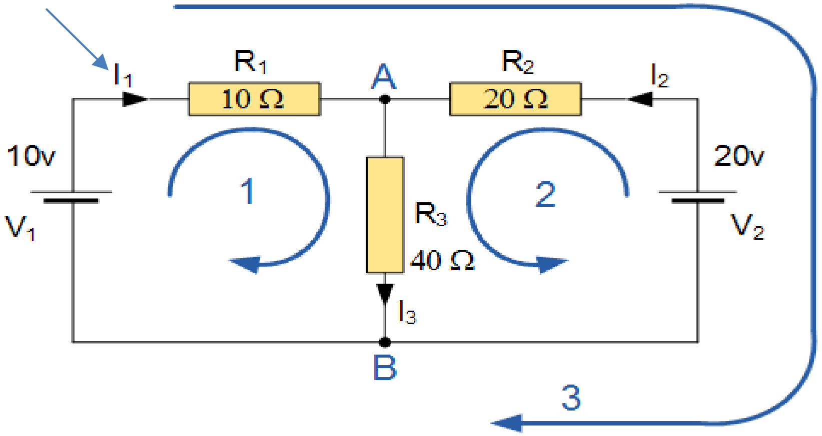
"Stochastic" questions available for **macroscopic** situations: (Heisenberg uncertainty influence microscopic ones)

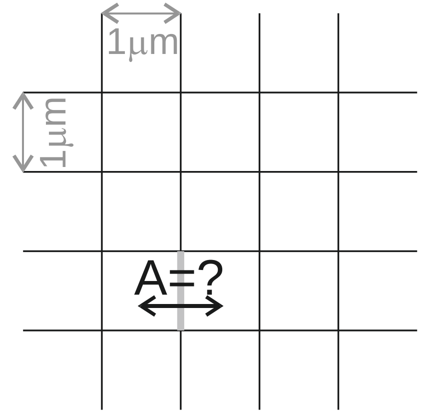
insert single electron here



What is the probability of finding it here after 1ns ?

Kirchhoff laws to find **currents** – **percentage** of electrons going given wire



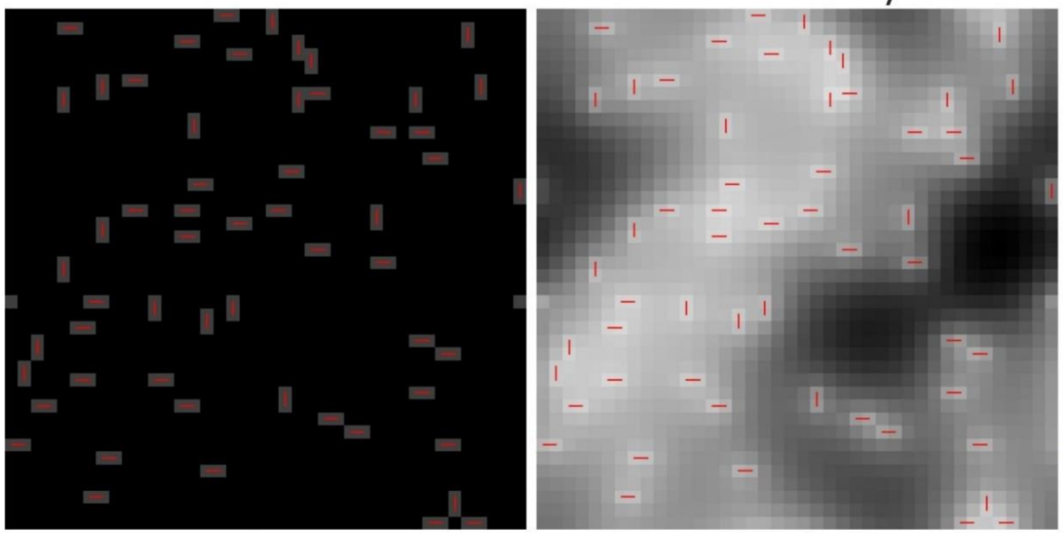


Imagine **semiconductor** as lattice and **ask for currents** – **stochastic** model

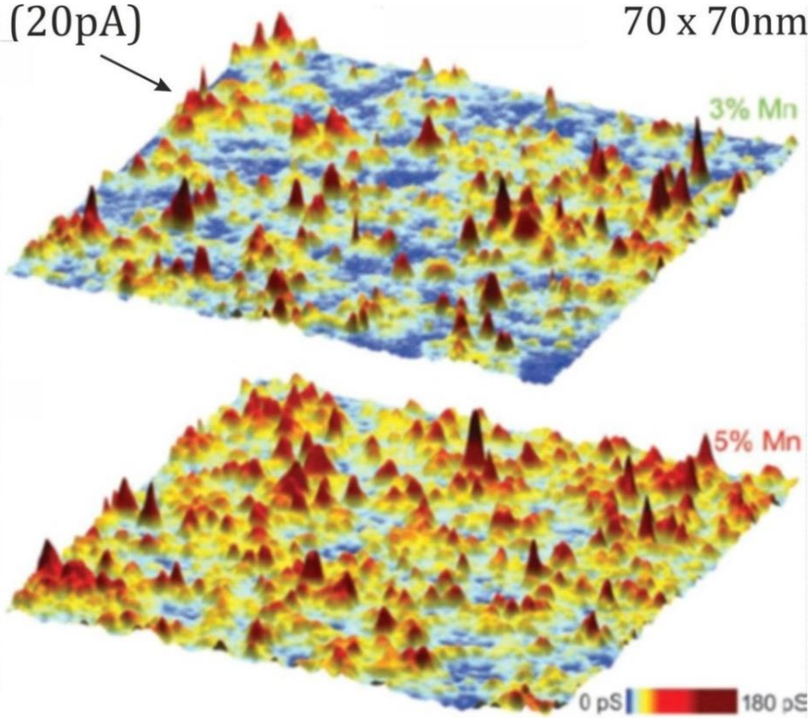
STM measurenments of electron density for $Ga_{1-x}Mn_xAs$ (20pA)

GRW

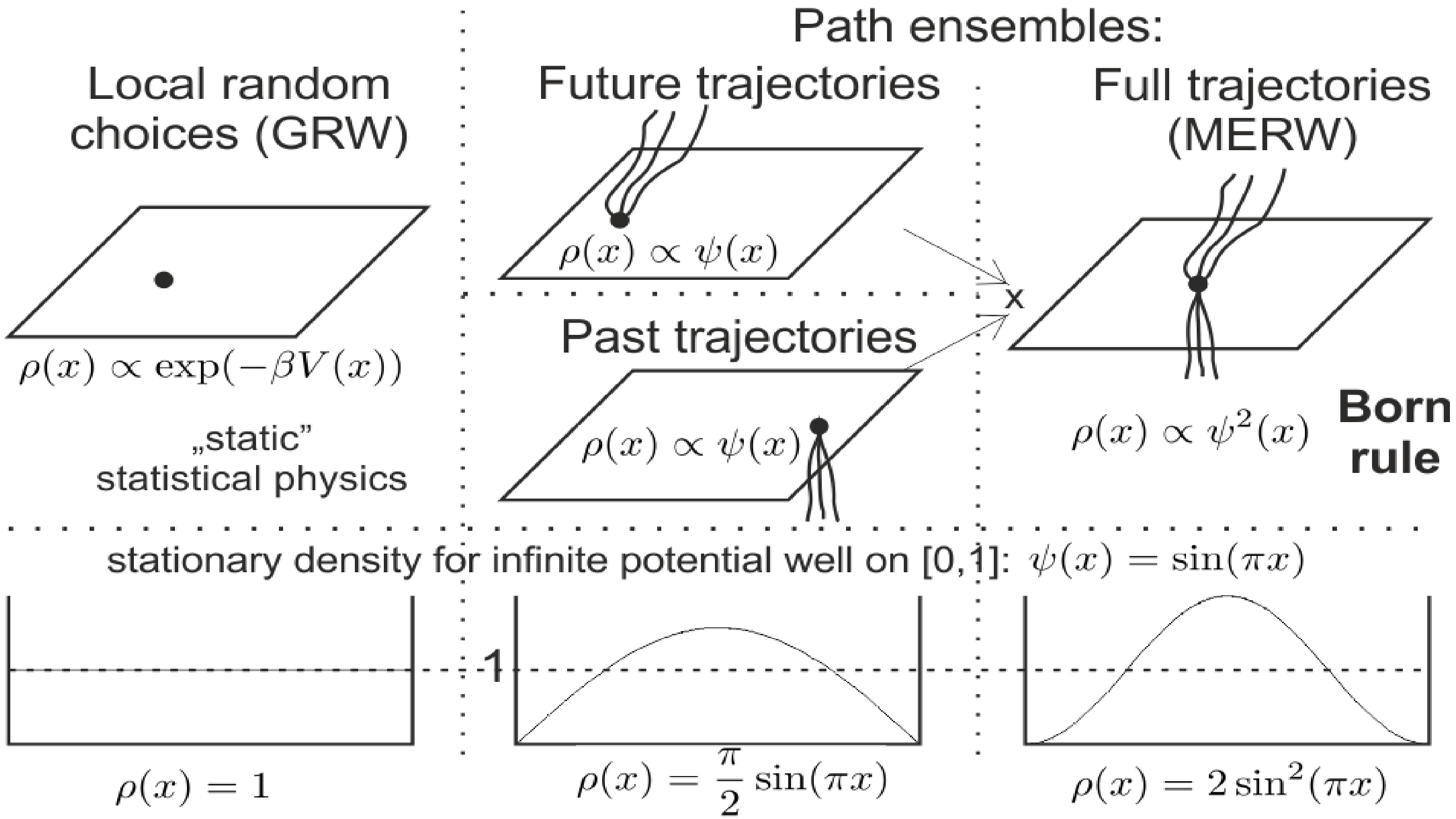
ground state density
of Bose-Hubbard / MERW



conductor insulator



(Link) Discrepancy source: **GRW only approximates maximal uncertainty principle**



Squares in **Born rule**
[lead to violation of](#)
[**Bell inequalities**](#)

$$\mathrm{P}(A=B) \;+\; \mathrm{P}(A=C) \;+\; \mathrm{P}(B=C) \;\geq\; 1$$

$$\begin{array}{c} A=0 \\ A=1 \end{array}
\begin{array}{|c|c|c|c|} \hline 1 & & 1 & \\ \hline & 1 & & 1 \\ \hline \end{array}
+
\begin{array}{|c|c|c|c|} \hline 1 & 1 & & \\ \hline & & 1 & 1 \\ \hline \end{array}
+
\begin{array}{|c|c|c|c|} \hline 1 & & & 1 \\ \hline 1 & & & 1 \\ \hline \end{array}
=
\begin{array}{|c|c|c|c|} \hline 3 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 3 \\ \hline \end{array}$$

$$\begin{array}{ccccc} B= & 0 & 1 & 0 & 1 \\ C= & 0 & 0 & 1 & 1 \end{array}$$

probabilities

amplitudes $\quad \psi = \begin{array}{|c|c|c|c|} \hline 0 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 0 \\ \hline \end{array}$

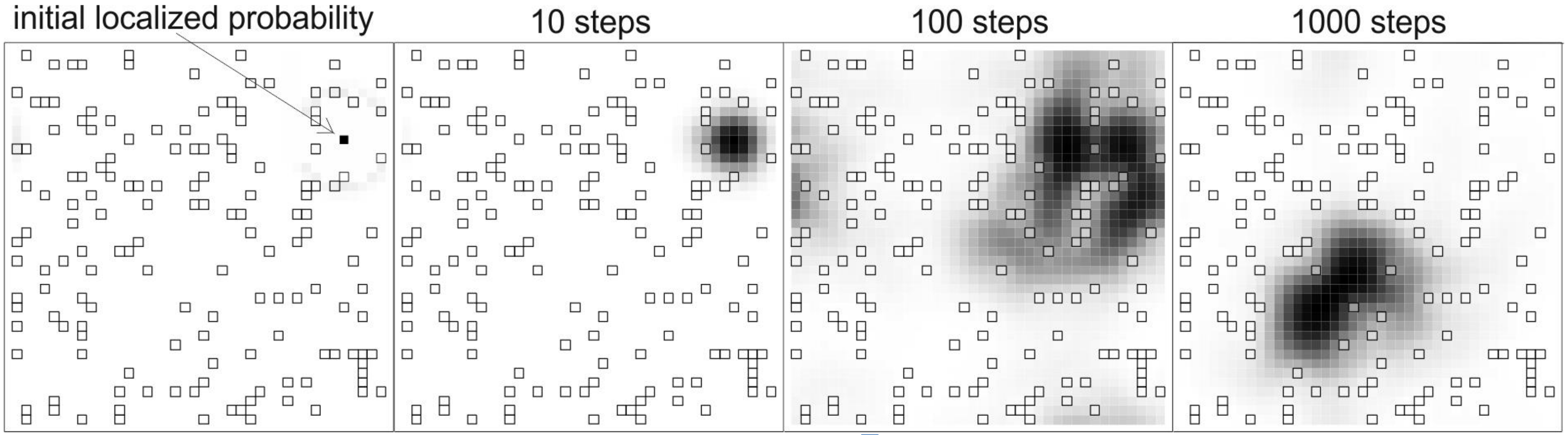
$$\mathrm{P}(A=B) = \frac{(\psi_{000}+\psi_{001})^2 + (\psi_{110}+\psi_{111})^2}{\sum_{A,B\in\{0,1\}} (\psi_{AB0}+\psi_{AB1})^2} = \frac{2}{10}$$

$$\mathrm{P}(A=B) \;+\; \mathrm{P}(A=C) \;+\; \mathrm{P}(B=C) \;= 0.6$$

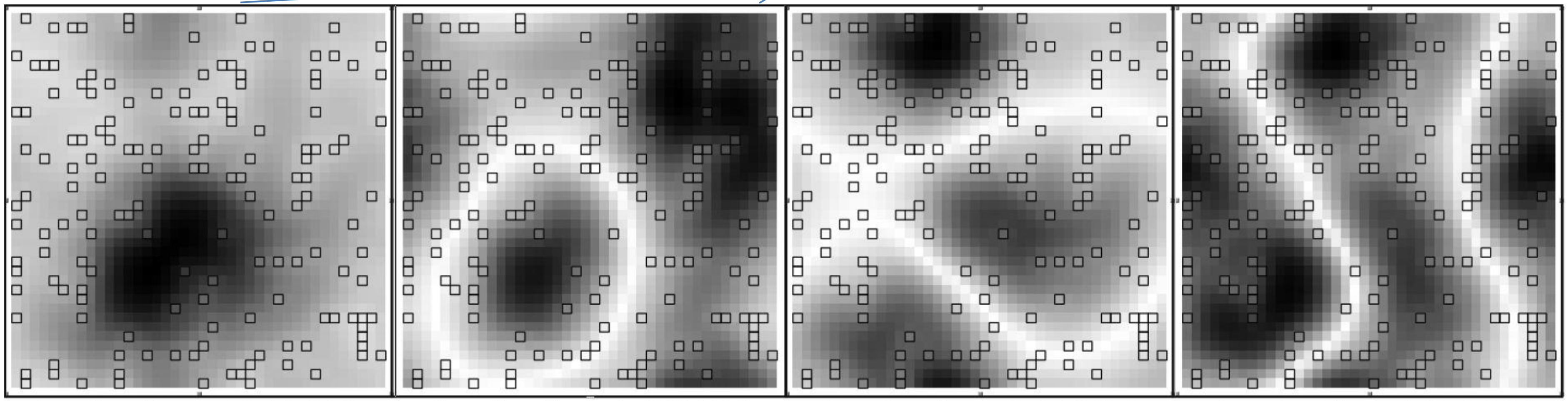
MERW evolution:

$$\left(S^{\mathrm{M}}\right)_{ij}^{t} = \frac{(M)_{ij}^{t}}{\lambda_0^t}\frac{\psi_{0,j}}{\psi_{0,i}} = \left(\sum_k \left(\frac{\lambda_k}{\lambda_0}\right)^t \varphi_{k,j}\psi_{k,i}\right)\frac{\psi_{0,j}}{\psi_{0,i}}$$

First "stochastic shift" toward **near** (overlapping) eigenvectors (sub-diffusion), then "deexcitate" toward nearer **ground state** (super-diffusion)

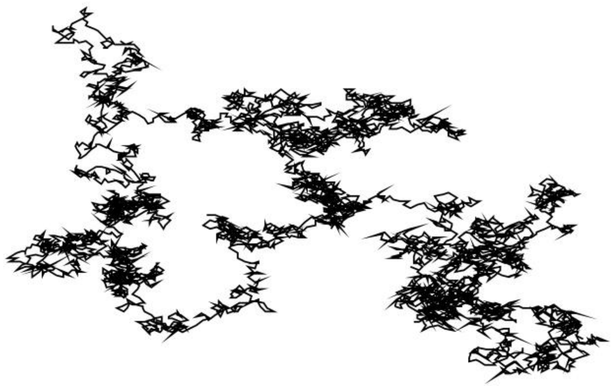


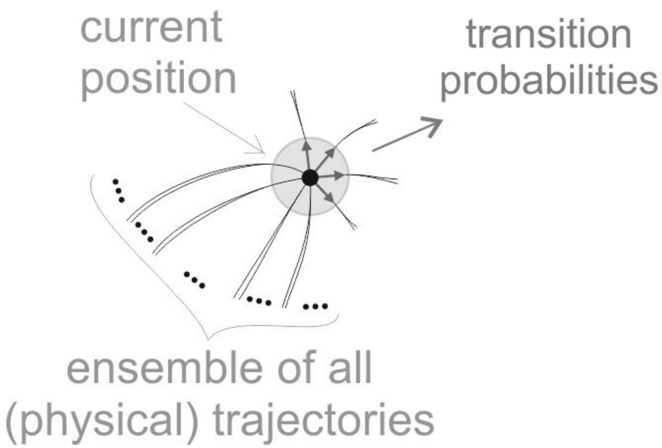
Eigenvectors $|\psi_k|$:

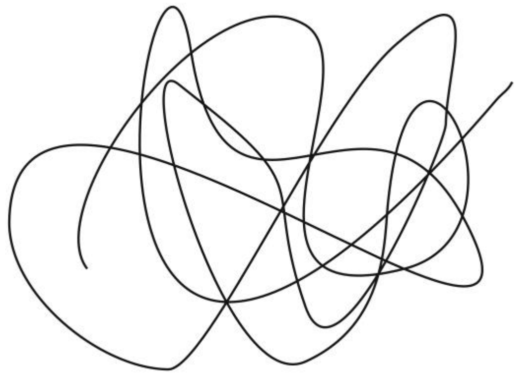


$k = 0 \quad \lambda \approx 4.942458 \qquad k = 1 \quad \lambda \approx 4.922788 \qquad k = 2 \quad \lambda \approx 4.918159 \qquad k = 3 \quad \lambda \approx 4.906692$

stochastic picture thermodynamical picture ergodic picture







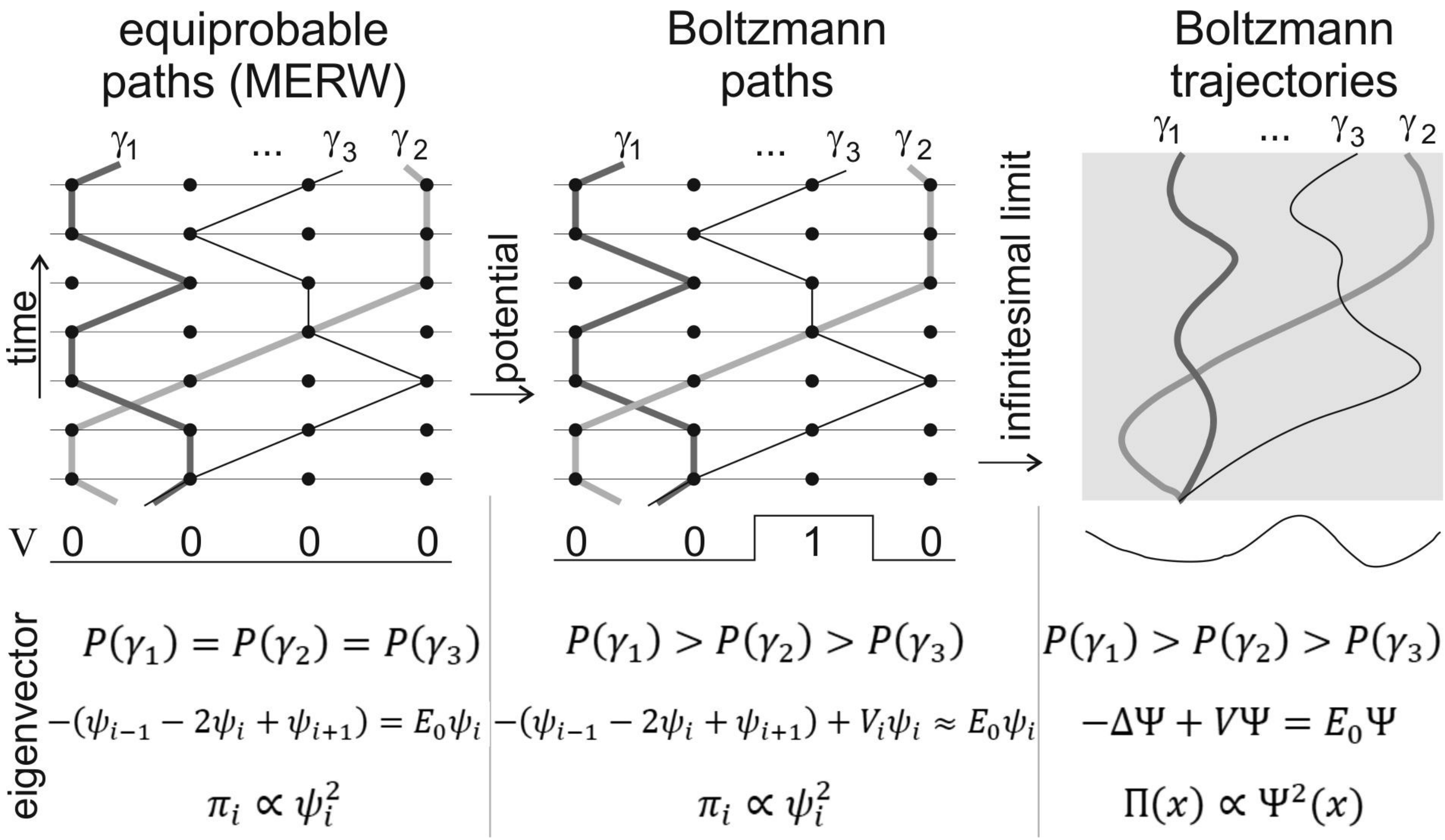
Stochastic picture **–** the evolution is indeed successive random decisions, accordingly to chosen by us transition probabilities – locally maximizing entropy, no **localization property**
Ergodic picture – evolution is usually fully determined, but because of chaotic behavior we introduce densities by averaging over single trajectory (thermodynamical fluctuations?)

Thermodynamical picture**: system too complicated -** use maximal uncertainty principle/canonical ensemble to predict the most probable behavior only.
– transition probabilities calculated from canonical ensemble among possible trajectories going through a given point – fully optimizing entropy (free energy),
– object doesn't directly use these probabilities (**nonlocal** - depend on the whole space), but just somehow chooses a trajectory (**not imposing any local probabilistic rules!**)
Only we use the found probabilities to estimate the probability density of its position,
– stationary density has strong localization property – to thermal equilibrium predicted by quantum mechanics – ground state density of the corresponding Hamiltonian.

Add potential to emphasize some scenarios: **Boltzmann distribution**

maximizes entropy while fixed sum of energies (minimizes free energy)

$$\max_{(p_i):\sum_i p_i=1} \left(\sum_i p_i \ln(1/p_i) - \sum_i p_i E_i\right) = \ln\left(\sum_i e^{-E_i}\right) \quad \text{for} \quad p_i \propto e^{-E_i}$$



energy of path $(\gamma_t\gamma_{t+1} \ldots \gamma_s)$ is $V_{\gamma_t\gamma_{t+1}} + \cdots + V_{\gamma_{s-1}\gamma_s}$

Boltzmann distribution among paths – use matrix: $M_{ij} = A_{ij} e^{-\beta V_{ij}}$

$$S_{\gamma_0\gamma_1} S_{\gamma_1\gamma_2} \dots S_{\gamma_{l-1}\gamma_l} = \frac{M_{\gamma_0\gamma_1} \dots M_{\gamma_{l-1}\gamma_l}}{\lambda^l} \frac{\psi_{\gamma_l}}{\psi_{\gamma_0}} = \frac{e^{-\beta\left(V_{\gamma_0\gamma_1} + V_{\gamma_1\gamma_2} + .. + V_{\gamma_{l-1}\gamma_l}\right)}}{\lambda^l} \frac{\psi_{\gamma_l}}{\psi_{\gamma_0}}$$

Eigenequation for 1D lattice: ϵ – time step, δ – lattice constant

$$\lambda_\epsilon \psi_i = (M_\epsilon \psi)_i = e^{-\beta\epsilon\frac{V_{i-1}+V_i}{2}} \psi_{i-1} + e^{-\beta\epsilon V_i} \psi_i + e^{-\beta\epsilon\frac{V_i+V_{i+1}}{2}} \psi_{i+1}$$

$$\lambda_\epsilon \psi_i \approx \psi_{i-1} + \psi_i + \psi_{i+1} - \epsilon\beta \left(\frac{V_{i-1} + V_i}{2} \psi_{i-1} + V_i \psi_i + \frac{V_i + V_{i+1}}{2} \psi_{i+1} \right)$$

$$\lambda_\epsilon \psi_i \approx \psi_{i-1} + \psi_i + \psi_{i+1} - 3\epsilon\beta V_i \psi_i \qquad /-3\psi_i \qquad /\cdot \frac{-1}{3\beta\epsilon}$$

$$\frac{3 - \lambda_\epsilon}{3\beta\epsilon} \psi_i \approx -\frac{1}{3\beta} \frac{\psi_{i-1} - 2\psi_i + \psi_{i+1}}{\epsilon} + V_i \psi_i$$

$\epsilon \to 0$

$$\epsilon = \frac{\delta^2}{3\alpha}, \qquad E_\epsilon = \frac{3-\lambda_\epsilon}{3\beta\epsilon} \longrightarrow \boxed{E\Psi = \left(-\frac{\alpha}{\beta}\Delta + V\right)\Psi}$$

Going to normalized $\Psi^2(x)$ stationary probability density for the lowest possible E

Propagator: $$S^t(x,y) = \frac{\langle x \middle| e^{-t\beta\hat{H}} \middle| y \rangle}{e^{-t\beta E_0}} \frac{\Psi_0(y)}{\Psi_0(x)} = \frac{\sum_i e^{-t\beta E_i} \langle x|\Psi_i\rangle\langle\Psi_i|y\rangle}{e^{-t\beta E_0}} \frac{\Psi_0(y)}{\Psi_0(x)}$$

MERW philosophy: $S^t(x,y) = \frac{\langle x\left|e^{-t\beta\hat{H}}\right|y\rangle}{e^{-t\beta E_0}}\frac{\Psi_0(y)}{\Psi_0(x)} = \frac{\sum_i e^{-t\beta E_i}\langle x|\Psi_i\rangle\langle\Psi_i|y\rangle}{e^{-t\beta E_0}}\frac{\Psi_0(y)}{\Psi_0(x)}$

Euclidean path integrals (EPI)

Considered paths have time length τ.

$$U(x',\tau;x) = A(\tau)\int_{\text{all paths}} \exp\left[-\frac{1}{\hbar}S_E[x(\tau)]\right]$$

Discrete analogue would be GRW_τ.

Differences:

1) EPI paths end at the final time: **path ensemble depends on times**,

MERW paths go to infinity – **fixed probability distribution** among paths

2) EPI starts with mathematically **problematic continuous** case

MERW start with **understanding** of much simpler **discrete** case

3) EPI performs Wick rotation to **imaginary time** of **assumed** axioms,

MERW just applies universal mathematical **maximal entropy principle**

Time dependence – e.g. potential can vary with time: $M_{ij}^t = A_{ij} e^{-\beta V_{ij}^t}$

energy of path $(\gamma_t \gamma_{t+1} \dots \gamma_s)$ is $V_{\gamma_t \gamma_{t+1}}^t + \dots + V_{\gamma_{s-1}\gamma_s}^{s-1}$ where $V_{ij}^t \equiv V_{ij}(t)$

Generalized dominant eigenvectors: density on the end of **past** and **future** ensembles

$$\varphi_j^t := \lim_{l\to\infty} \frac{\sum_i \left(M^{t-l} M^{t-l+1} \dots M^{t-1}\right)_{ij}}{\widetilde{N}^t(l)} \qquad \psi_i^t := \lim_{l\to\infty} \frac{\sum_j \left(M^t M^{t+1} \dots M^{t+l-1}\right)_{ij}}{N^t(l)} \qquad (\geq 0)$$

$$((\varphi^t)^T M^t)_j = \lim_{l\to\infty} \frac{\sum_i \left(M^{t-l} M^{t-l+1} \dots M^t\right)_{ij}}{\widetilde{N}^t(l)} = \tilde{\lambda}^t \varphi_j^{t+1} \qquad \text{where} \qquad \tilde{\lambda}^t = \lim_{l\to\infty} \frac{\widetilde{N}^{t+1}(l+1)}{\widetilde{N}^t(l)}$$

$$(M^t \psi^{t+1})_i = \lim_{l\to\infty} \frac{\sum_i \left(M^t M^{t+1} \dots M^{t+l}\right)_{ij}}{\widetilde{N}^{t+1}(l)} = \lambda^t \psi_i^t \qquad \text{where} \qquad \lambda^t = \lim_{l\to\infty} \frac{N^t(l+1)}{N^{t+1}(l)}$$

$$\boxed{\text{Stationary probability: } p_i^t = \varphi_i^t \psi_i^t \qquad \text{propagator} \qquad S_{ij} = \frac{M_{ij}^t}{\lambda^t} \frac{\psi_j^{t+1}}{\psi_j^t}}$$

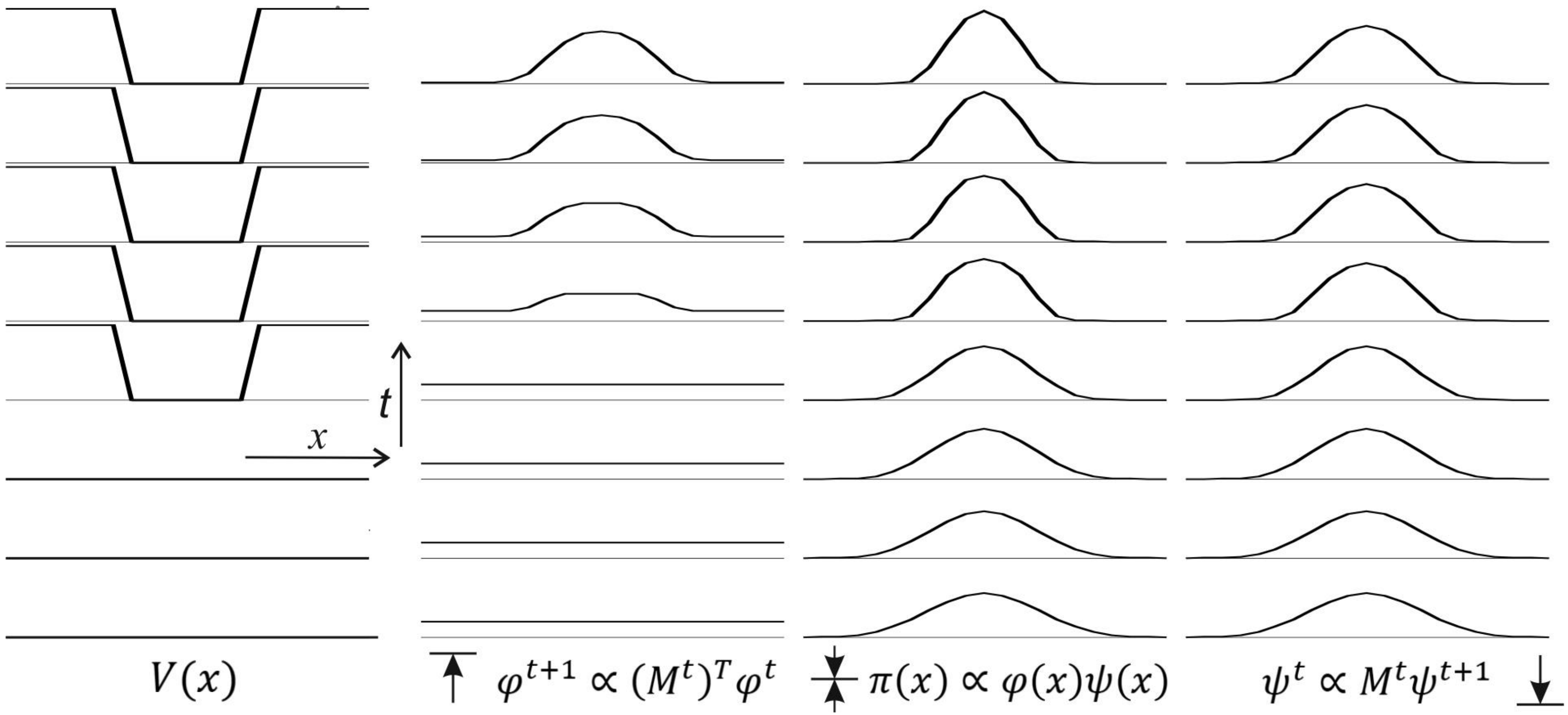
$$\boxed{(S^{ts})_{ij} := (S^t S^{t+1} \dots S^{s-1})_{ij} = \frac{(M^t M^{t+1} \dots M^{s-1})_{ij}}{\lambda^t \lambda^{t+1} \dots \lambda^{s-1}} \frac{\psi_j^s}{\psi_i^t}}$$

Conserved probability? $\quad (\varphi^t)^T \psi^t = (\varphi^t)^T \frac{M^t \psi^{t+1}}{\lambda^t} = (\varphi^t)^T M^t \frac{\psi^{t+1}}{\lambda^t} = \frac{\tilde{\lambda}^t}{\lambda^t} (\varphi^{t+1})^T \psi^{t+1}$

Continuity equation $\Leftrightarrow \lambda = \tilde{\lambda}$ (exact values only balance between φ and ψ)

$$\boxed{\text{Final evolution equation:} \qquad \lambda^t \varphi^{t+1} = (M^t)^T \varphi^t \qquad\qquad M^t \psi^{t+1} = \lambda^t \psi^t}$$

Adiabatic approximation: If V is locally constant, φ and ψ are approximately right and left dominant eigenvectors of M … but generally:



Boltzmann distribution among paths is time-symmetric

It is effective model: only represents our knowledge
We know about the change – that later particle should be in the well,
so earlier it should be nearby

Now for **1D lattice** there appears additional time derivative:

$$\lambda_\epsilon^t \psi_x^t = (M^t \psi^{t+1})_x \approx \psi_{x-1}^{t+1} + \psi_x^{t+1} + \psi_{x+1}^{t+1} - 3\epsilon\beta V_x^t \psi_x^t \qquad /-3\psi_x^{t+1} \qquad /\cdot \frac{-1}{3\epsilon\beta}$$

$$\frac{1}{\beta}\,\frac{\psi_x^{t+1}-\psi_x^t}{\epsilon} - E_\epsilon^t \psi_x^t \approx -\frac{1}{3\beta}\frac{\psi_{x-1}^{t+1}-2\psi_x^{t+1}+\psi_{x+1}^{t+1}}{\epsilon} + V_x^t\psi_x^t \qquad \text{for} \qquad E_\epsilon^t := \frac{3-\lambda_\epsilon^t}{3\epsilon\beta}$$

Finally choosing $\epsilon = \frac{\delta^2}{3\alpha}$ in infintesimal limit we get **evolution equations**:

$$\boxed{\frac{d}{dt}\Phi = \beta\left(E - \widehat{H}\right)\Phi \qquad \frac{d}{dt}\Psi = \beta\left(\widehat{H} - E\right)\Psi \qquad \text{for} \qquad \widehat{H} = -\frac{\alpha}{\beta}\Delta + V}$$

Φ should evolve forward in time (to be stable), Ψ backward

In **adiabatic approximation** $\Phi \approx \Psi$ for $E(t) = \langle \Phi(t) | \widehat{H}(t)\Psi(t)\rangle$

$$\frac{d}{dt}(\Phi\Psi) = \beta\left(\left(\left(E-\widehat{H}\right)\Phi\right)\Psi + \Phi\left(\widehat{H}-E\right)\Psi\right) = \alpha\big((\Delta\Phi)\Psi - \Phi(\Delta\Psi)\big) = \alpha\nabla\cdot((\nabla\Phi)\Psi - \Phi(\nabla\Psi))$$

$$\boxed{\textbf{Continuity equation}: \quad \frac{d}{dt}\rho = -\nabla\cdot J \quad \text{for} \quad J = \alpha(\Phi\nabla\Psi - \Psi\nabla\Phi)}$$

Quantum ($\psi \in \mathbb{C}$) : $j = \frac{\hbar}{2mi}(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi})$ substituting $\psi = \frac{e^{i\gamma}}{\sqrt{2}}(\Phi + i\Psi)$

$$j = \frac{\hbar e^{i\gamma}e^{-i\gamma}}{4mi}\big((\Phi - i\Psi)\nabla(\Phi + i\Psi) - (\Phi + i\Psi)\nabla(\Phi - i\Psi)\big) = \frac{\hbar}{2m}(\Phi\nabla\Psi - \Psi\nabla\Phi)$$

Suggesting to choose $\alpha = \frac{\hbar}{2m}$ $\qquad \beta = \frac{2m}{\hbar^2}\alpha = \frac{1}{\hbar}$ (fundamental noise?)

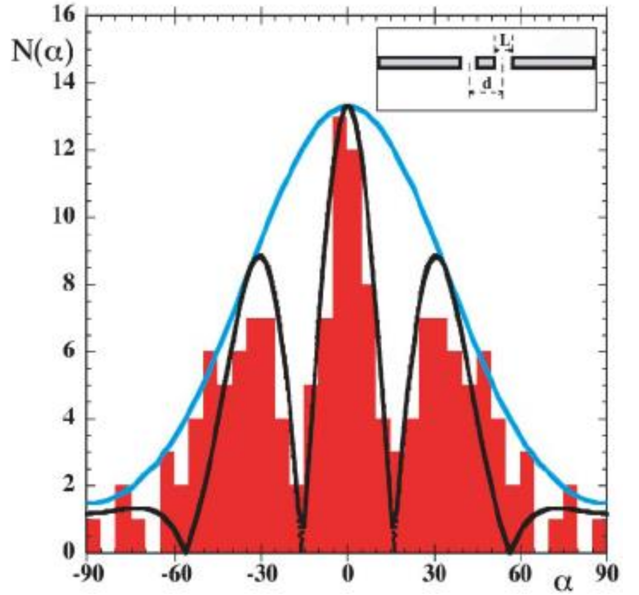
Macroscopic soliton model – oil droplet maintaining shape due to surface tension

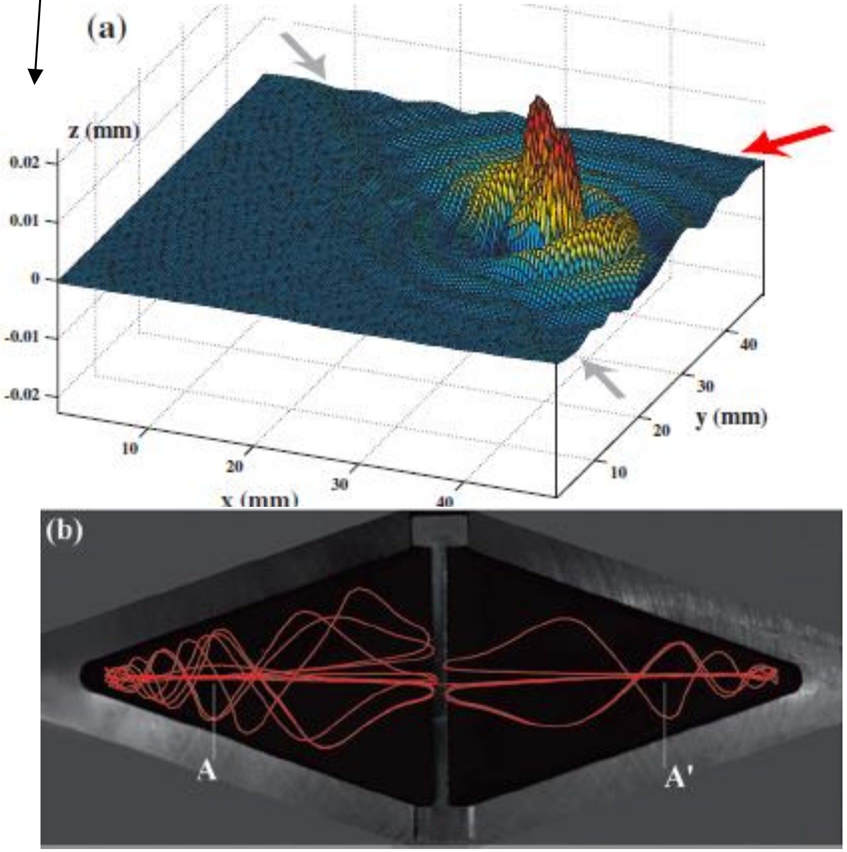
Bouncing droplet on vertically vibrated bath is **coupled to the surface waves it generates**. Becomes a "walker" moving at **constant velocity**.

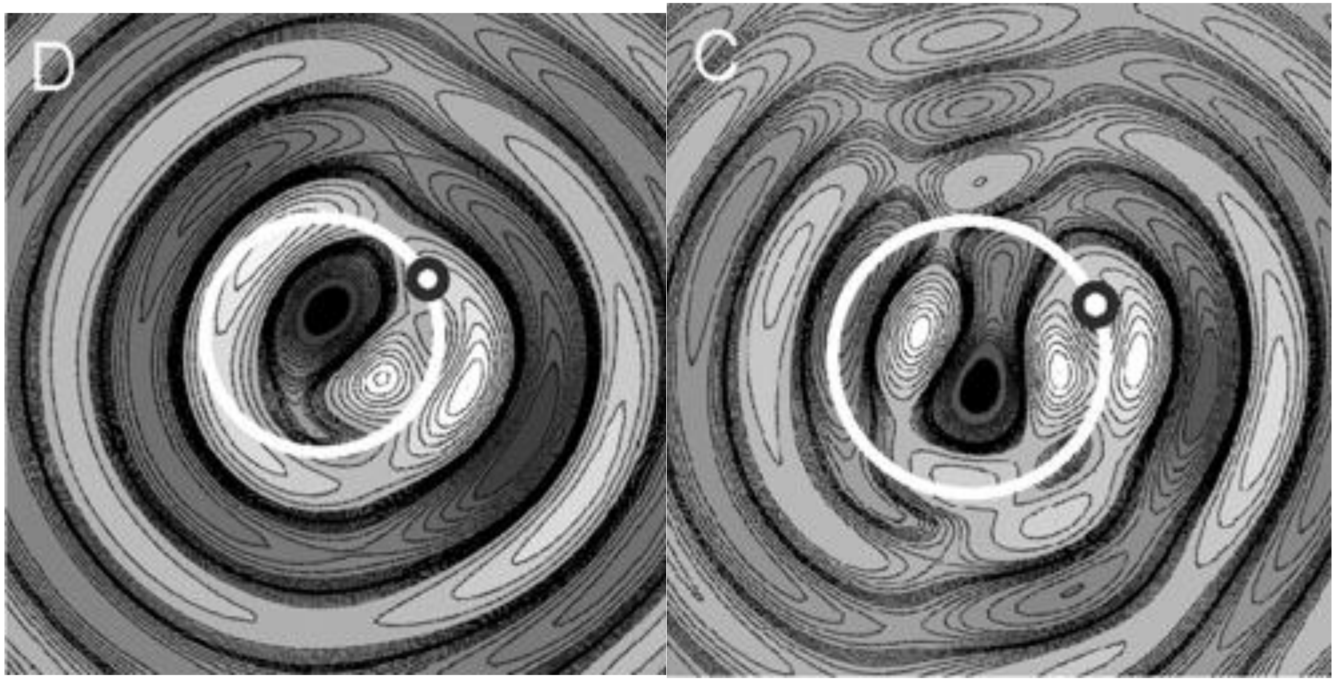
Y. Couder and E. Fort, **Single-Particle Diffraction and Interference at a Macroscopic Scale,** Phys. Rev. Lett. 97 (2006)

A. Eddi, E. Fort, F. Moisy, and Y. Couder, **Unpredictable Tunneling of a Classical Wave-Particle Association**, Phys. Rev. Lett. 102 (2009)

E. Fort, A. Eddi, A. Boudaoud, J. Moukhtar, and Y. Couder, **Path-memory induced quantization of classical orbits,** PNAS vol. 107 (2010)







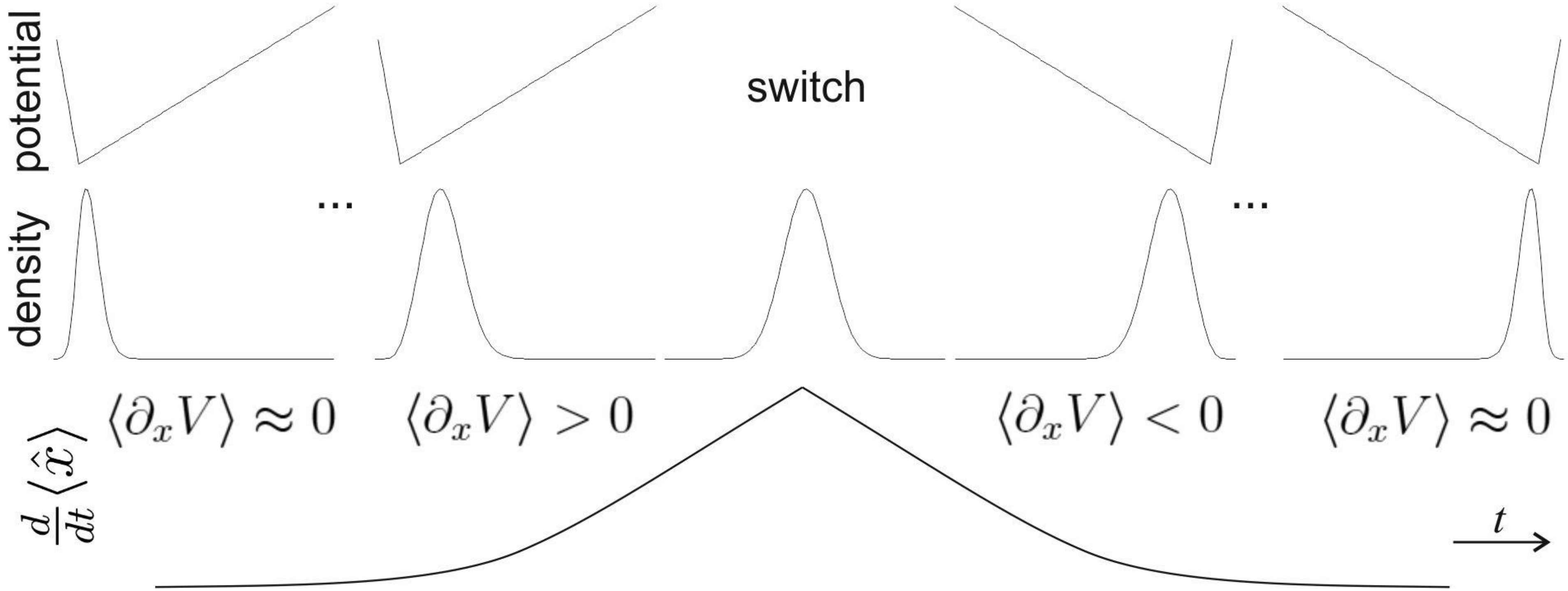
$$\frac{d}{dt}\langle\Phi|\hat{O}\Psi\rangle = \beta\langle\Phi|(E-\hat{H})\hat{O}\Psi\rangle + \langle\Phi|\frac{\partial\hat{O}}{\partial t}\Psi\rangle + \beta\langle\Phi|\hat{O}(\hat{H}-E)\Psi\rangle$$

Ehrenfest equation: $\langle\hat{O}\rangle = \langle\frac{\partial\hat{O}}{\partial t}\rangle + \beta\langle[\hat{O},\hat{H}]\rangle$

$[\hat{x},\hat{H}] = 2\frac{\alpha}{\beta}\nabla \quad \Rightarrow \quad \frac{\mathrm{d}\langle\hat{x}\rangle}{\mathrm{dt}} = \langle 2\alpha\nabla\rangle = \frac{\langle\hat{p}\rangle}{m}$ for $\hat{p} = 2\alpha\nabla = \hbar\nabla$

Now $[\hat{p},\hat{H}] = [\hbar\nabla, V] = \hbar\nabla V \quad \Rightarrow \quad \frac{d}{dt}\langle\hat{p}\rangle = \beta\langle\hbar\nabla V\rangle = \langle\nabla V\rangle = \int \rho(x)\nabla V(x)dx$

Getting opposite than expected: $m\frac{d^2}{dt^2}\langle\hat{x}\rangle = +\langle\nabla V\rangle$



In quantum mechanics ψ is complex function

$$\langle\psi|\psi\rangle = \text{const} \quad \text{because} \quad \langle\psi| \to e^{i\hat{H}t/\hbar}\langle\psi| \quad \text{while} \quad |\psi\rangle \to e^{-i\hat{H}t/\hbar}|\psi\rangle$$

In MERW Φ and Ψ are real nonnegative functions

$$\langle\Phi|\Psi\rangle = \text{const} \quad \text{because} \quad \langle\Phi| \to e^{-\beta t(\hat{H}-E)}\langle\Phi| \quad \text{while} \quad |\Psi\rangle \to c^{\beta t(\hat{H}-E)}|\Psi\rangle$$

This time momentum operator is not self-adjoined:

$$\hat{p} = \hbar\nabla \qquad\qquad \hat{p}^\dagger = -\hbar\nabla$$

$\hat{p}^2$ also is not self-adjoined, so we have to use $\hat{p}^\dagger\hat{p}$ instead

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V = \frac{\hat{p}^\dagger\hat{p}}{2m} + V$$

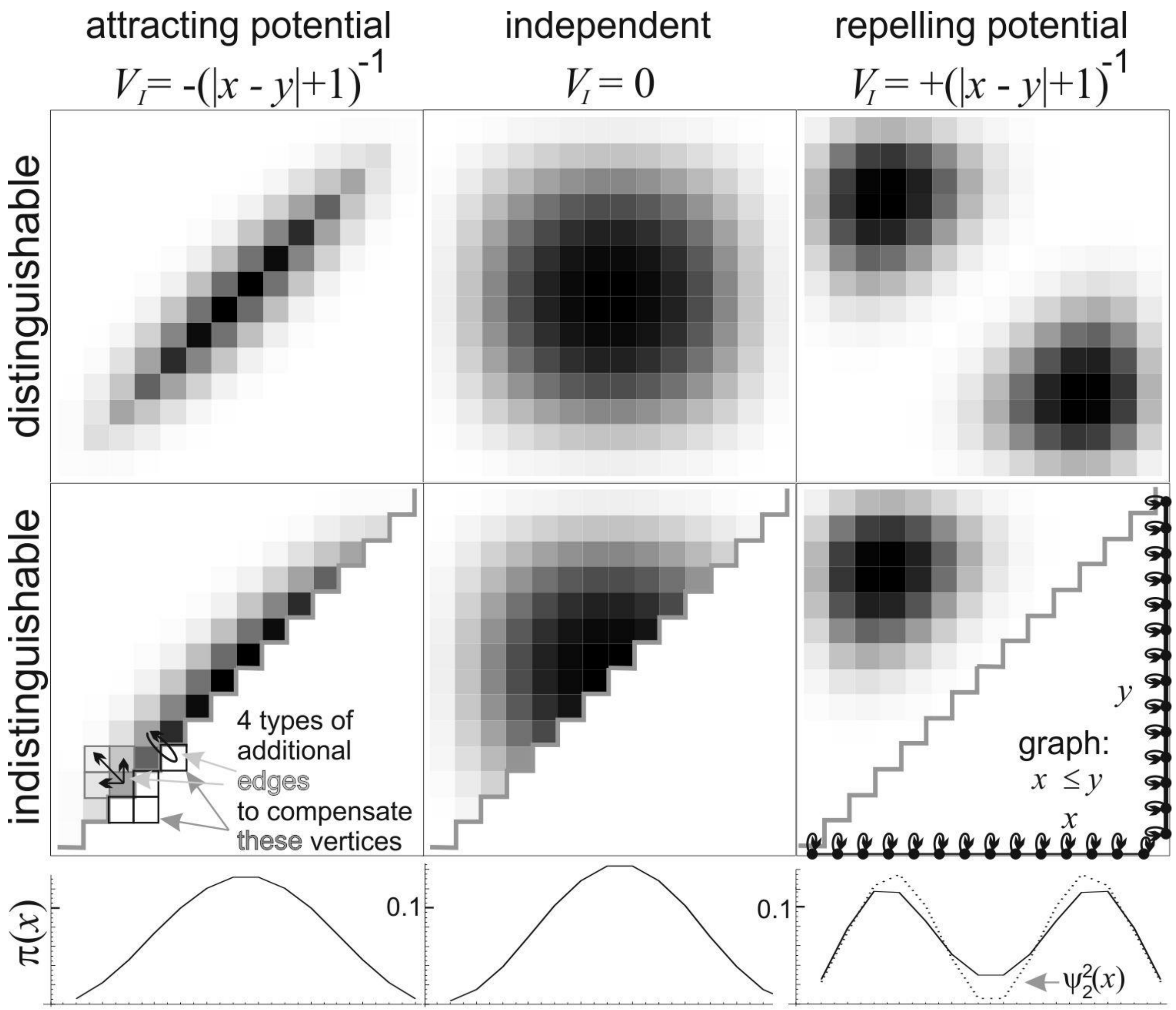
For adiabatic approximation ($\Phi = \Psi$) we get **Heisenberg principle** analogue:

$$0 \le \langle(\hat{x} + \lambda\hat{p})\Psi|(\hat{x} + \lambda\hat{p})\Psi\rangle = \langle\Psi|(\hat{x} - \lambda\hat{p})(\hat{x} + \lambda\hat{p})\Psi\rangle = \langle\hat{x}^2\rangle + \lambda^2\langle\hat{p}^\dagger\hat{p}\rangle - \lambda\hbar$$

Discriminant ≤ 0:

$$\sqrt{\langle\hat{x}^2\rangle}\sqrt{\langle\hat{p}^\dagger\hat{p}\rangle} \ge \frac{\hbar}{2}$$

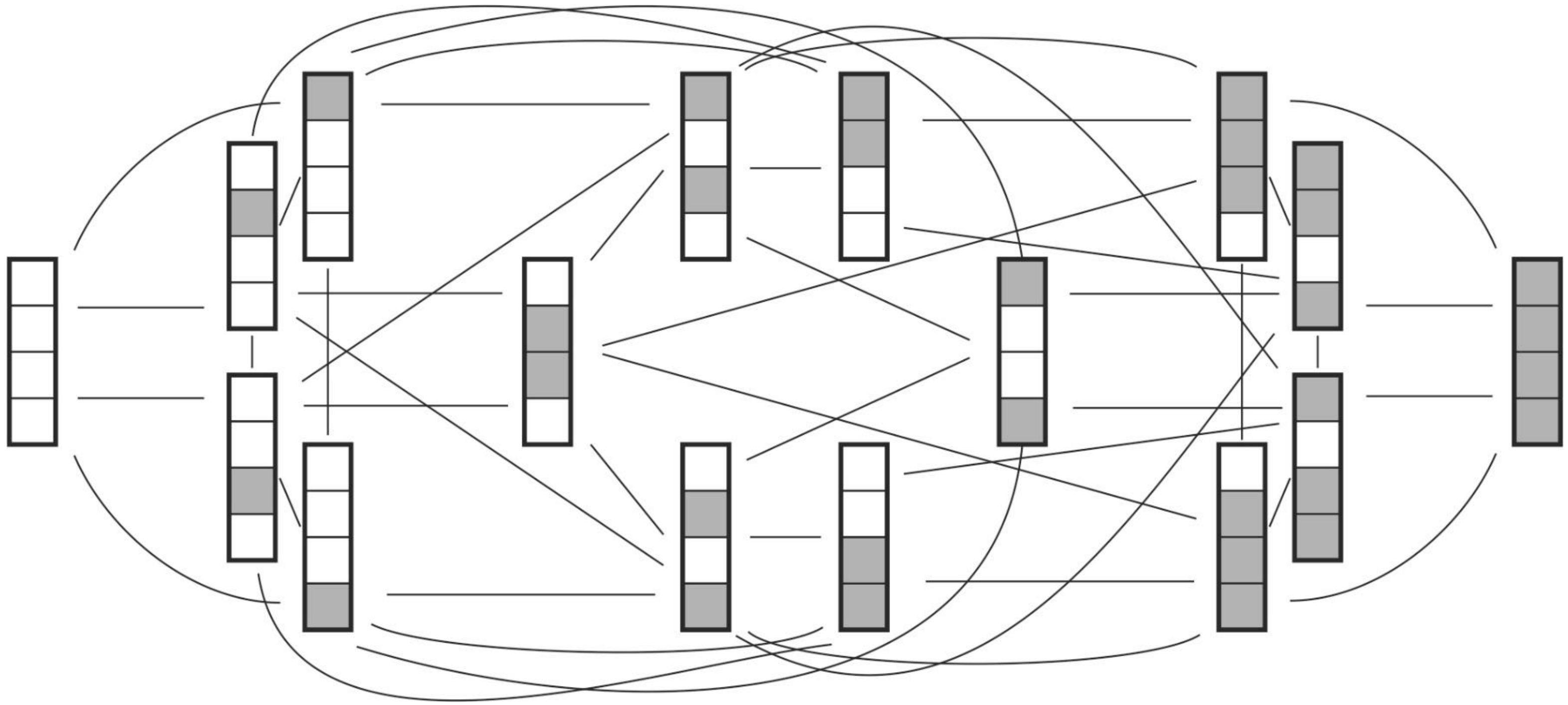
Two particles – consider trajectory in the space of pair configurations



Thermodynamical **Pauli exclusion principle**:
repelling particles choose separate dynamical equilibrium **states**

Various number of particles: vertex $\equiv$ configuration

For example adjacency matrix for fermions on length 4 segment graph



$|\bar{n}\rangle$ - sum of all $n!$ permutations

$\hat{a}|\bar{n}\rangle = n|\overline{n-1}\rangle$ $\quad$ $\hat{a}^\dagger|\overline{n-1}\rangle = |\bar{n}\rangle$ $\quad$ $\hat{a}^\dagger\hat{a}|\bar{n}\rangle = n|\bar{n}\rangle$ $\quad$ $[\hat{a}, \hat{a}^\dagger] = 1$

Standard normalization: $|n\rangle = |\bar{n}\rangle/\sqrt{n!}$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \qquad \hat{a}^\dagger|n-1\rangle = \sqrt{n}|n\rangle \qquad \left(\hat{a}^\dagger\right)^n|0\rangle = \frac{1}{\sqrt{n!}}|n\rangle$$

Bose-Hubbard model – repulsing bosons on lattice

$$\hat{H}_{BH} = -t \sum_{(i,j)\in\mathcal{E}} \hat{a}_j^\dagger \hat{a}_i + \frac{U}{2}\sum_{i\in\mathcal{V}} \hat{n}_i(\hat{n}_i - 1) \qquad \ldots + \sum_{i\in\mathcal{V}} V(i)\hat{n}_i + \sum_{i,j\in\mathcal{V}} V_I(i,j)\hat{n}_i\hat{n}_j$$

Accordingly to **MERW**: diagonal terms ≡ self-loops ("paying for staying")

$$\hat{H}_{MERW} \propto -\sum_{(i,j)\in\mathcal{E}} \hat{a}_j^\dagger \hat{a}_i\, e^{-\epsilon\beta V(\text{configuration before and after transition})} \approx$$

$$\approx -\sum_{(i,j)\in\mathcal{E}} \hat{a}_j^\dagger \hat{a}_i + \epsilon\beta d \sum_{i\in\mathcal{V}} V(\text{configuration after transition})\, \hat{a}_i^\dagger \hat{a}_i$$

Three ϵ order approximations used exactly as for lattices: $e^{-\epsilon\beta V} \approx 1 - \epsilon\beta V$, that for neighboring vertices, V and coordinates of dominant eigenvector are nearly equal $\left(\hat{a}_i^\dagger \hat{a}_j \approx \hat{a}_i^\dagger \hat{a}_i \right)$.

Both Hamiltonians are practically equivalent for **single particle without potential** and **in continuous limit**, but generally they only approximate each other.

Another question: why only one particle can transit at once?

Summary of diffusion part: If instead of guessing the stochastic propagator (assuming that the walker indeed uses these probabilities), we assume maximal entropy/uncertainty principle (only we use these probabilities), the predictions are no longer in disagreement with QM.

The main "quantum corrections to stochastic models": localization, e.g. in semiconductor … where else it is essential?

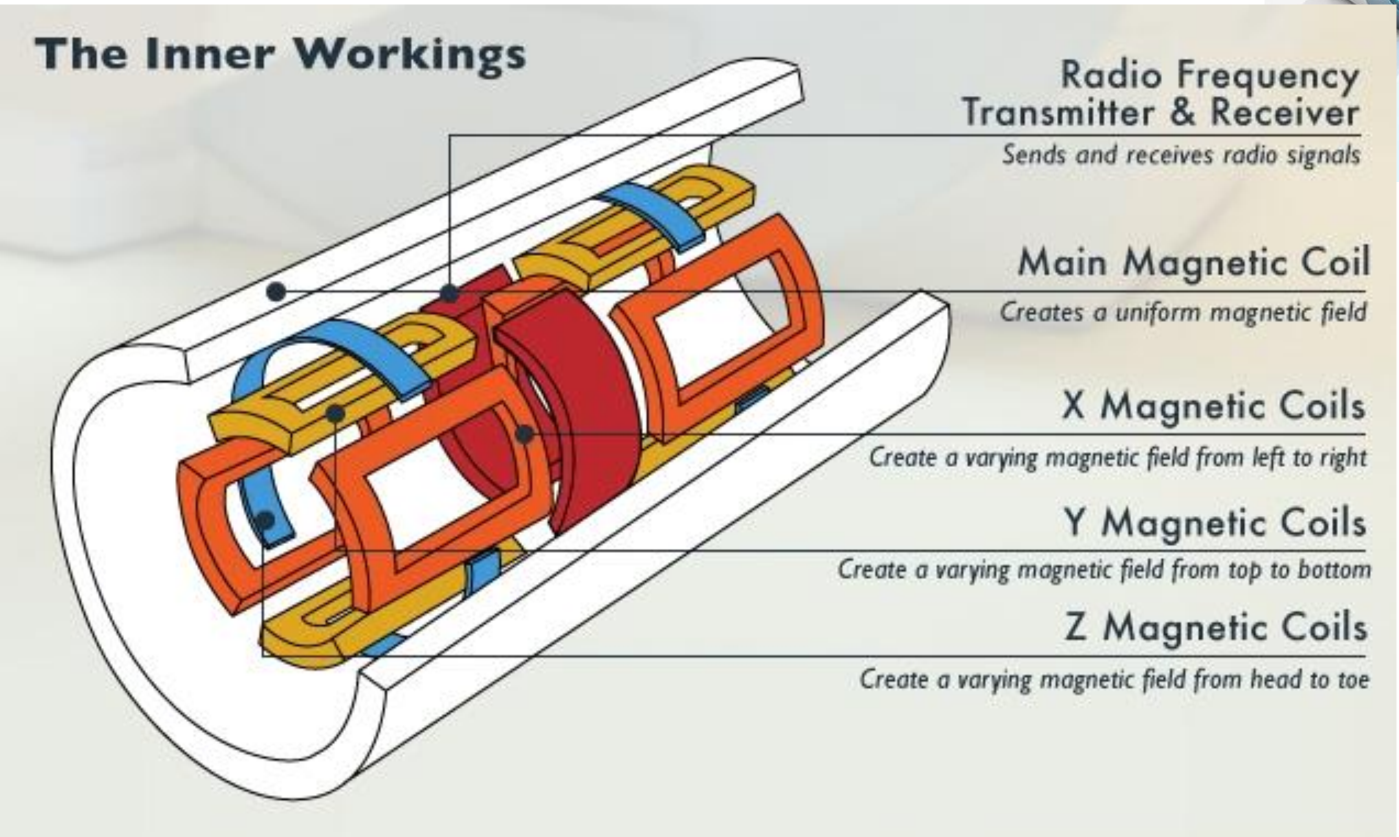
some further work:

- improving mathematical formalism,
- try to motivate, derive Levy parameters from deeper dynamics,
- see $S^t(x,y) = \frac{\sum_i e^{-t\beta E_i}\langle x|\Psi_i\rangle\langle\Psi_i|y\rangle}{e^{-t\beta E_0}}\frac{\Psi_0(y)}{\Psi_0(x)}$ propagator as **"stochastic shift toward quantum eigenstate"** of perturbed trajectories,
- add velocity into consideration in analogy to Langevin equation,
- add other internal degrees of freedom like direction of spin,
- find deeper understanding of quantum mechanics,
- **find more quantum corrections to standard diffusion models**.

MERW in tractography ...

MRI – Magnetic Resonance Imaging

- Align spins by strong magnetic field (3T)
- Use frequency of Larmor precession to select nuclei (e.g. proton) and positions
- Measure relaxation times:
  T2: spin-spin (synchronization)
  T1: spin-lattice (alignment)



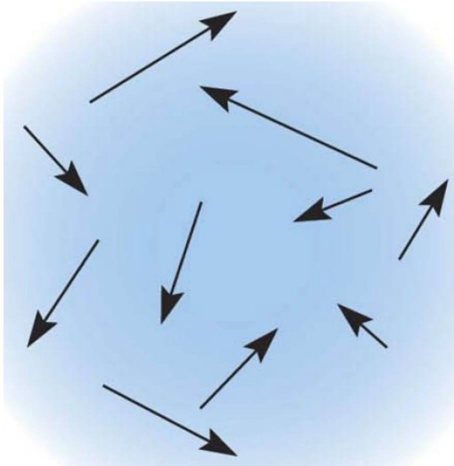
Fast relaxation e.g. in fats
Slower in water ...
Distinguish types of tissues

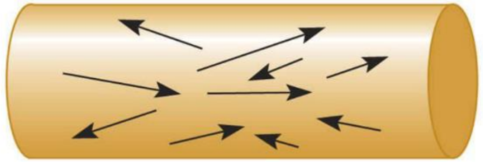
fMRI: look at Fe in heme:
in deoxyhemoglobin $S = 2$
in oxyhemoglobin $S = 0$
find tissues requiring oxygen

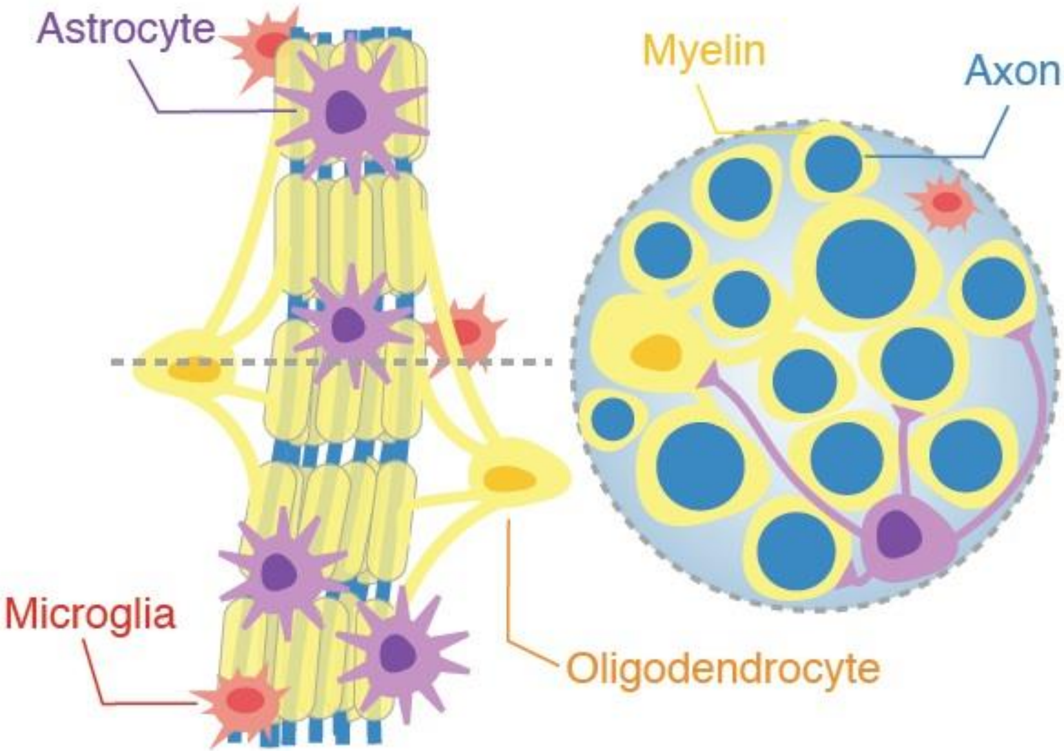
Strong magnetic gradient - then opposite one
Produces phase difference – then cancels it

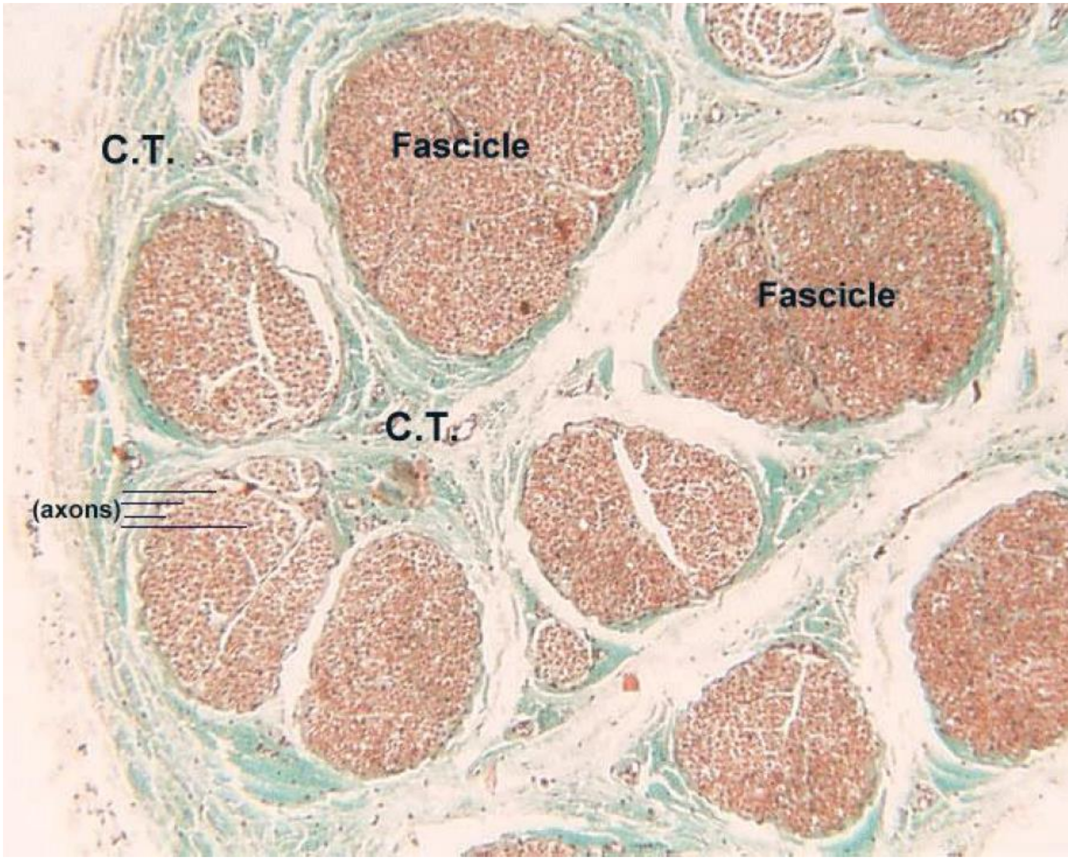
... **unless diffusion** ...

First: **DWI** – Diffusion Weighted Imaging
Maps diffusion coefficients (e.g. strokes)
Now: **DTI** – Diffusion Tensor Imaging:
Allows to track neural pathways

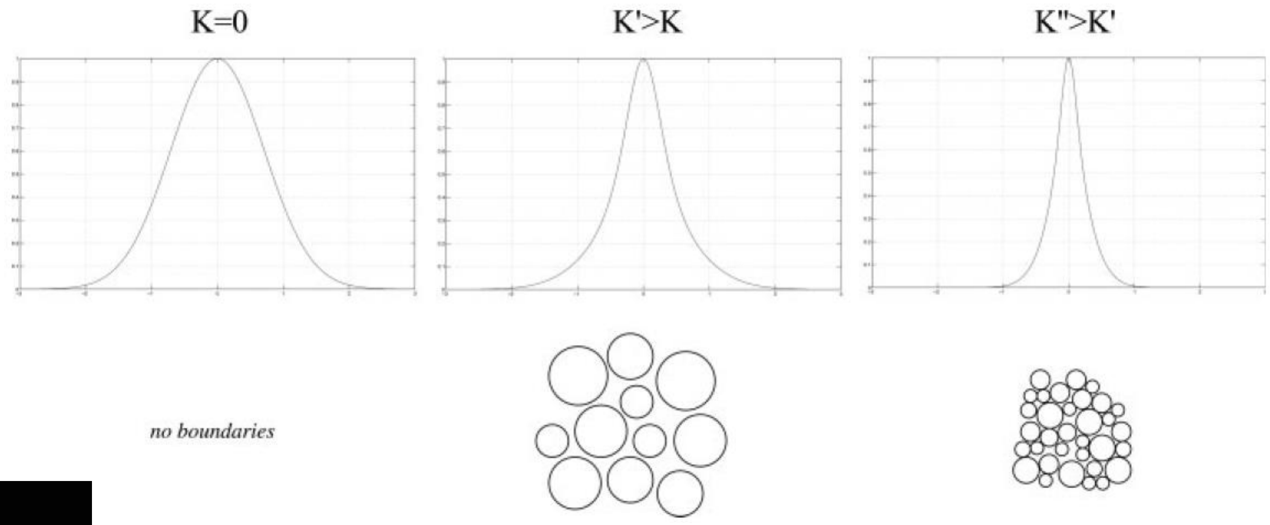


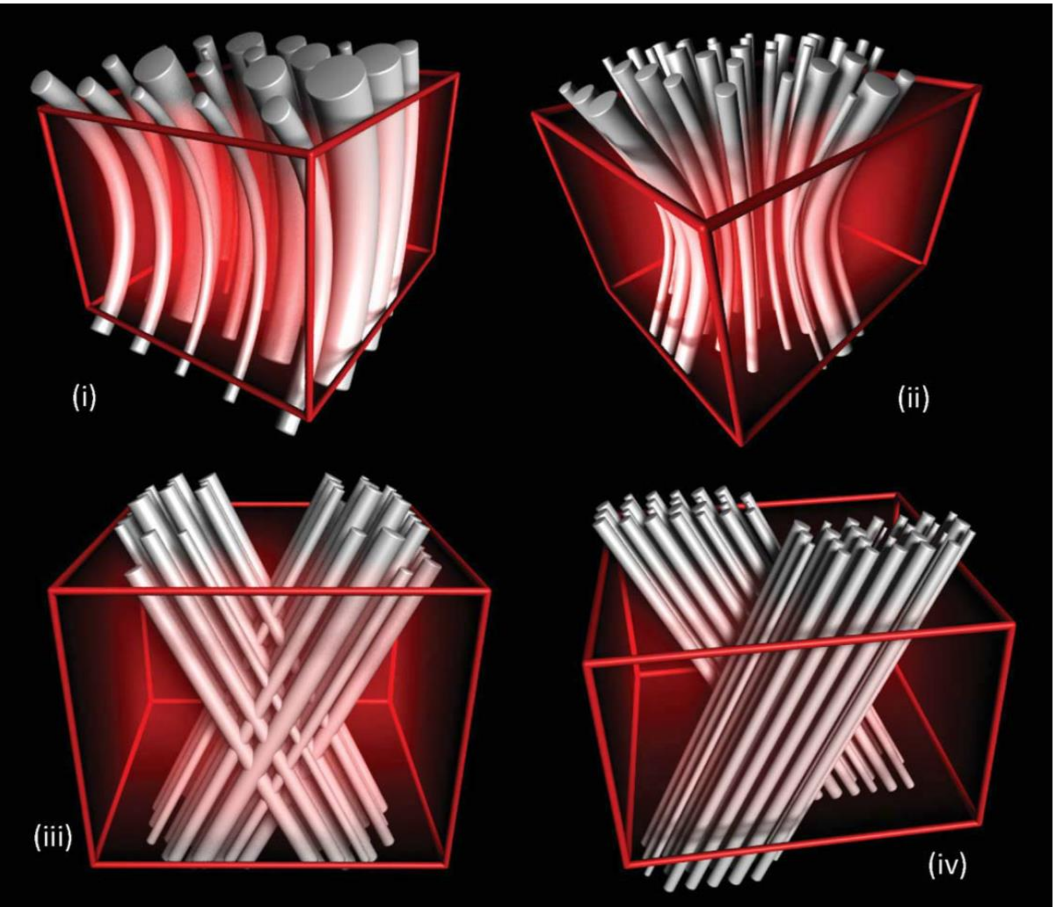






Measuring for
two different parameters
gives us kurtosis:





HARDI – High Angular Resolution Diffusion Imaging

Could allow to handle:

- Fiber crossings
- Parallel fibers

Difficult for **local tractography**:
Integrating field of directions

Global tractography:
Find best tracts basing on the entire data

DTI – represent local diffusion as an ellipse,

... but ...

We measure diffusion in many (e.g. 256) directions

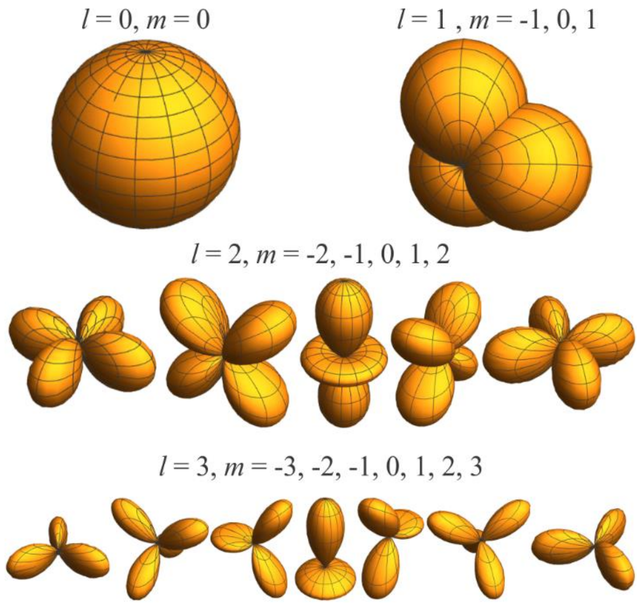
e.g. twice to get also kurtosis (different gradients)

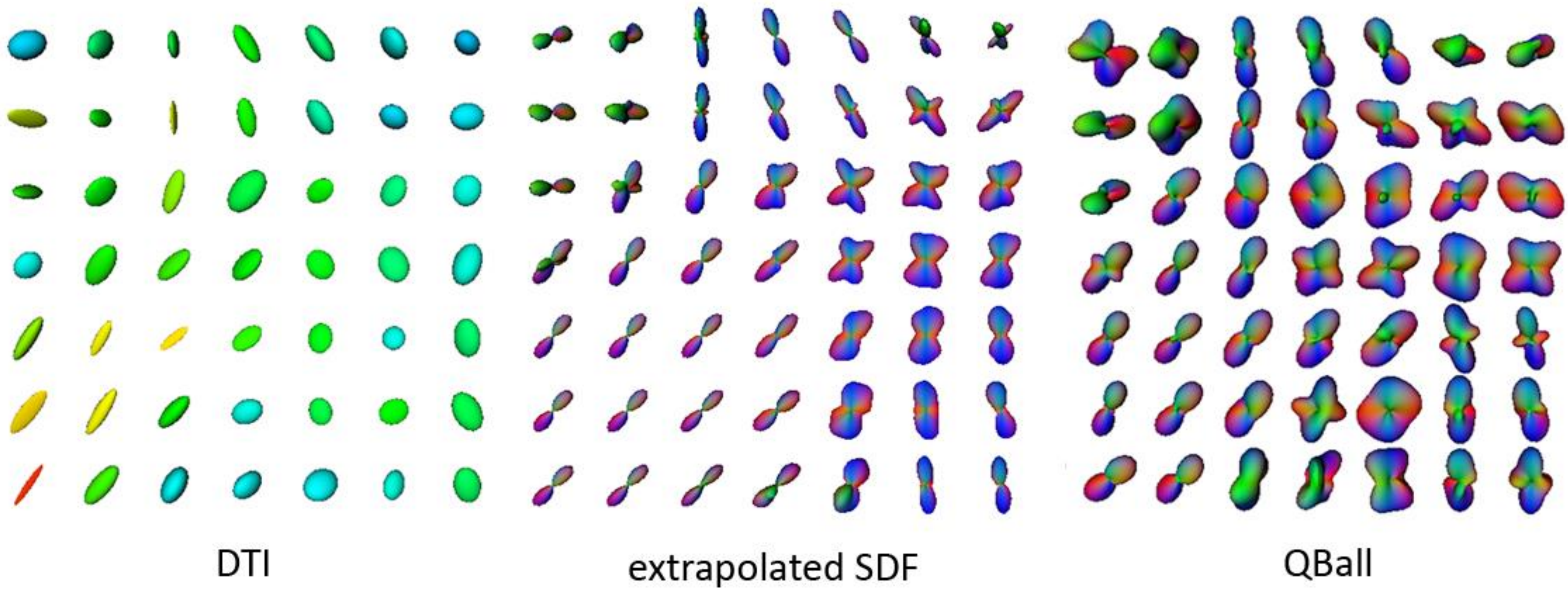
HARDI – high angular resolution diffusion imaging:

SDF – spherical diffusion function

QBall – smoothen on icosahedron (20 directions)

Spherical harmonics – orthogonal for $\int_S f g d\Omega$



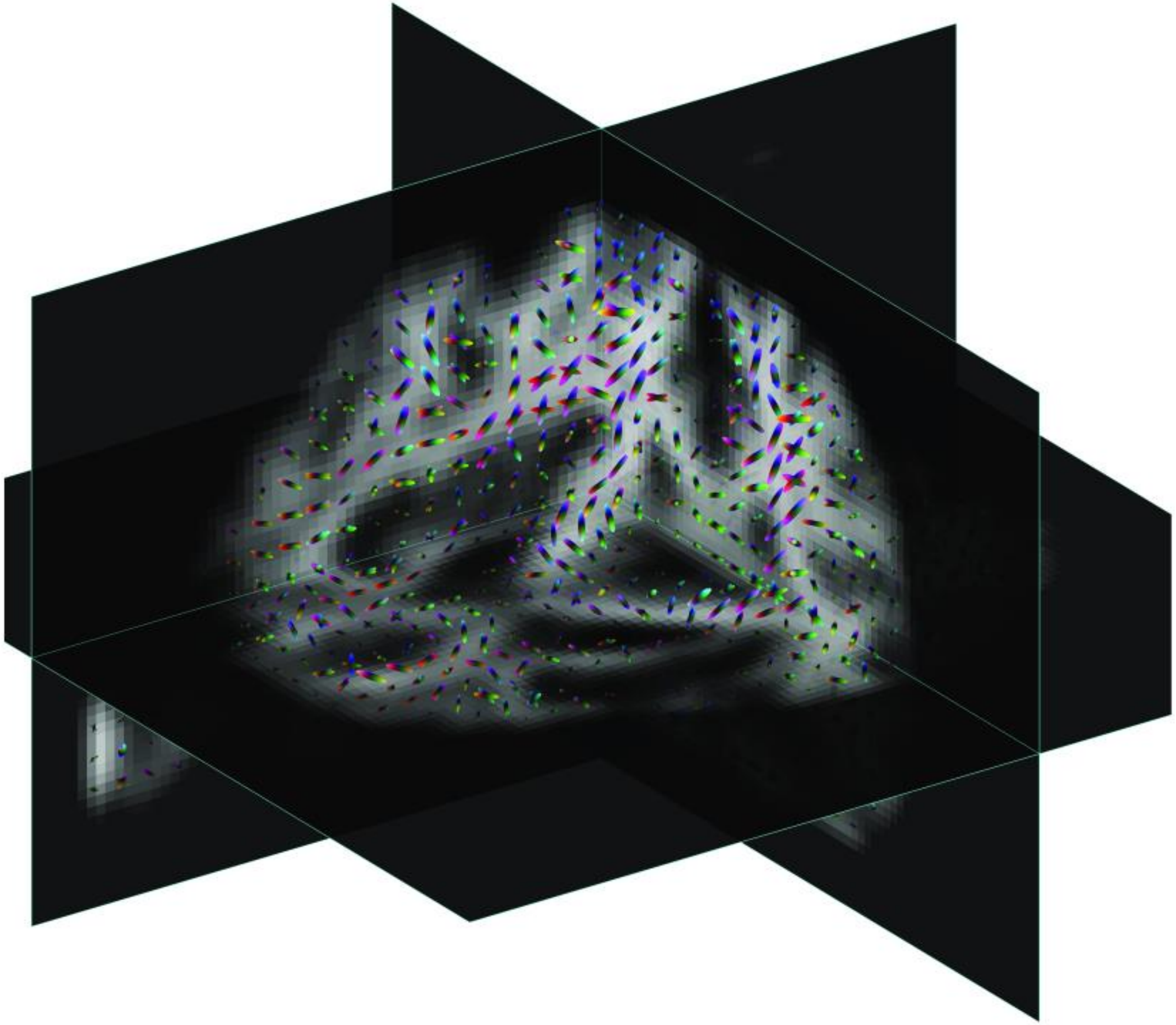


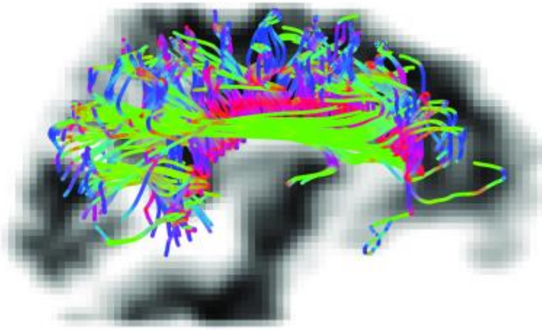
Finding continuous neural track structure from discretization

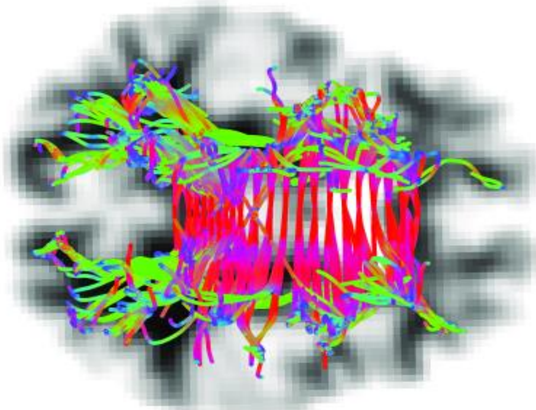
Information pathways in a disordered lattice,

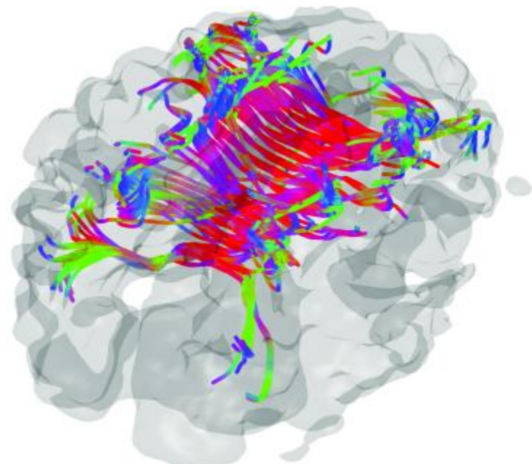
Lawrence R. Frank 1,2,* and Vitaly L. Galinsky, Phys. Rev. E (2014)

Simultaneous Multi-Scale Diffusion Estimation an Tractography Guided by Entropy Spectrum Pathways Vitaly L. Galinsky and Lawrence R. Frank, IEEE Transactions on Medical Imaging (2014)







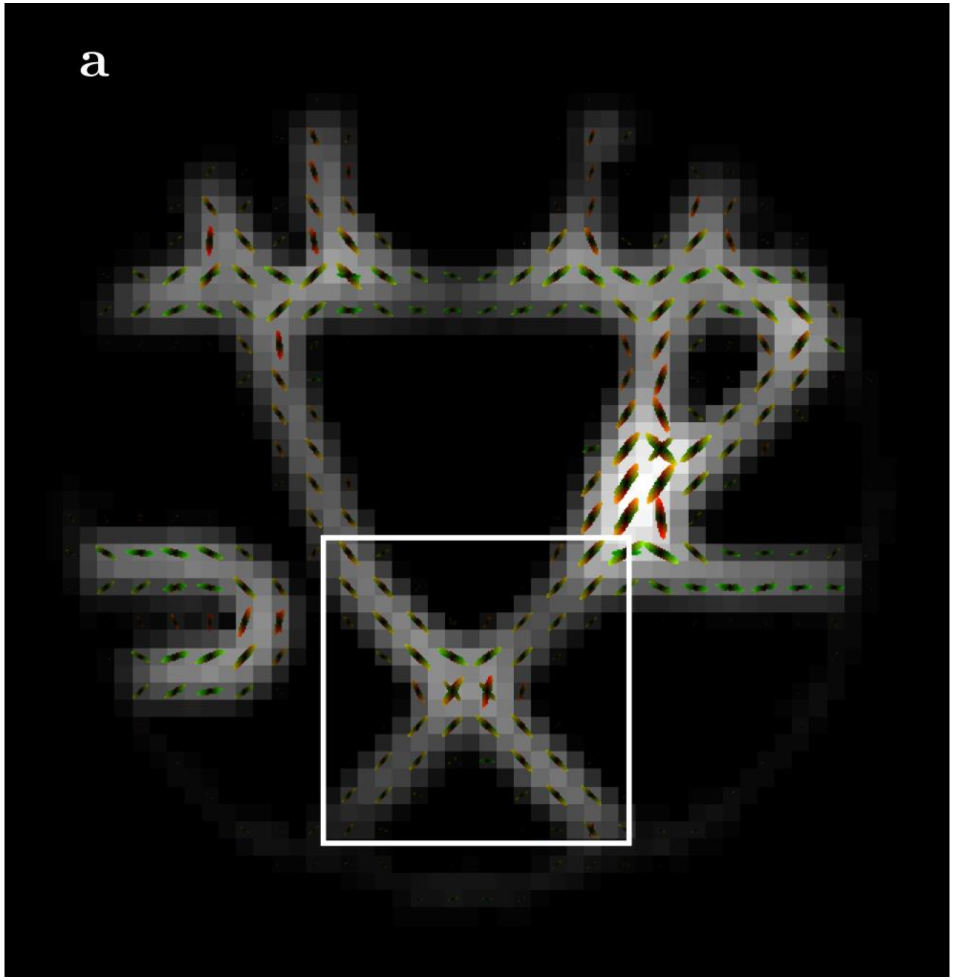


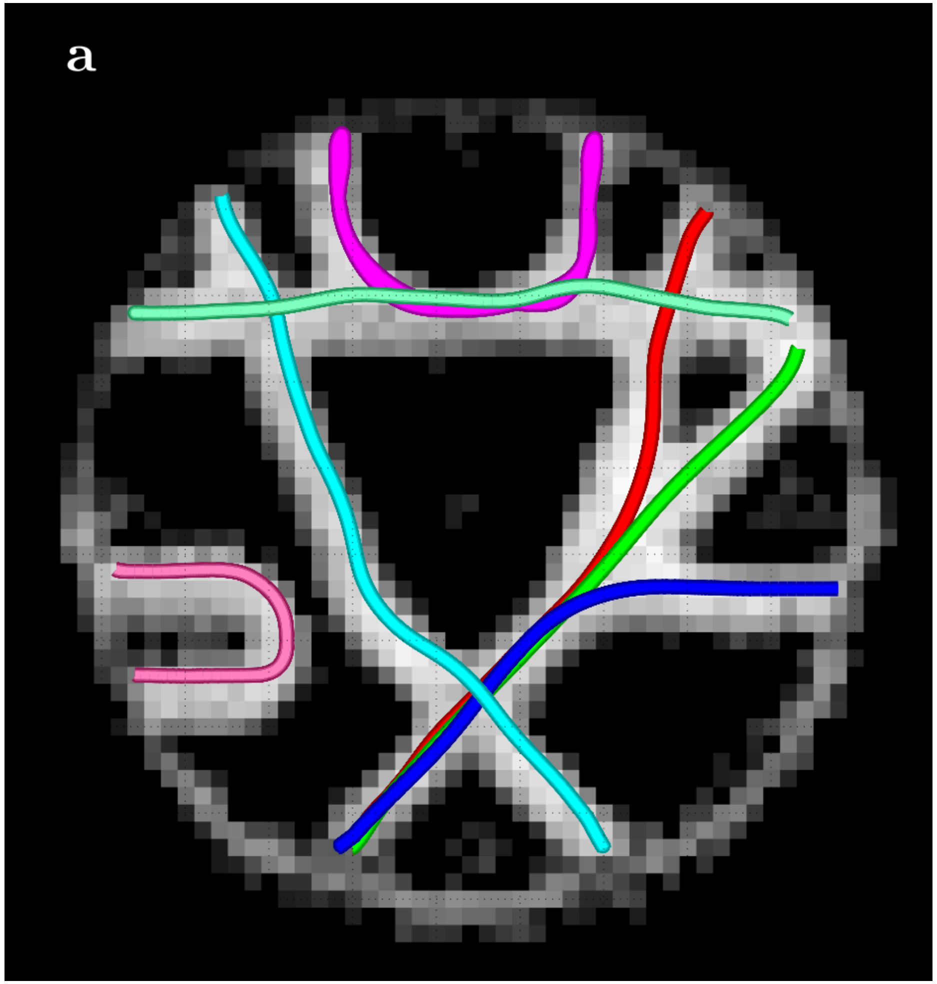
locally track **possible** → **globally** extract **most probable** pathways

(crossings?) **maximizing entropy/information flow**

Entropy Spectrum Pathways (ESP): use many largest eigenvectors

$$\rho_i^{(k)} = \left(\psi_i^{(k)}\right)^2 \qquad p_{ij}^{(k)} = \frac{M_{ij}}{\lambda_k} \frac{\psi_j^{(k)}}{\psi_i^{(k)}}$$



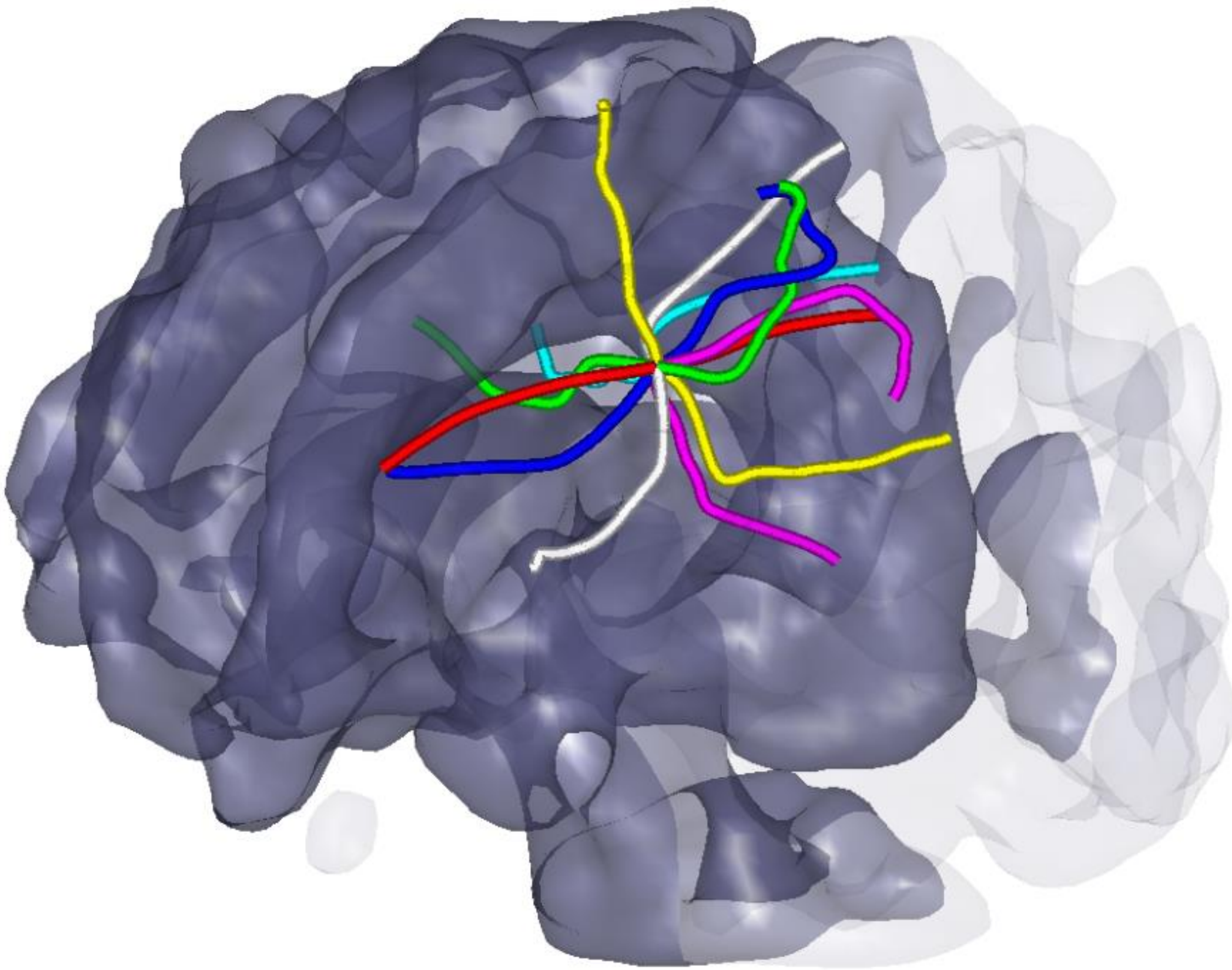


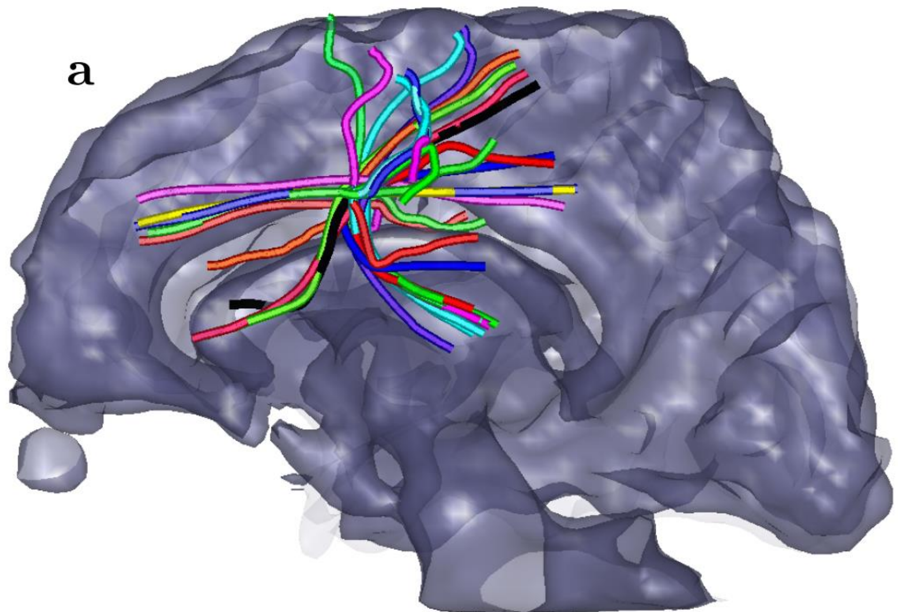
Global handling with crossings

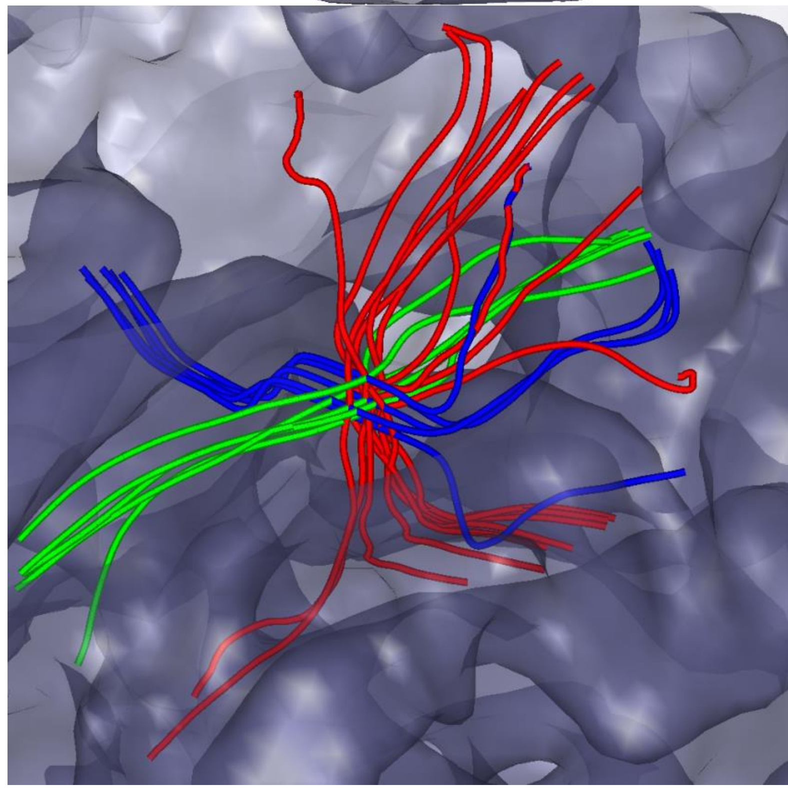
Doesn't use anatomical assumptions

(problem with individual differences)

Standard method: white track below only

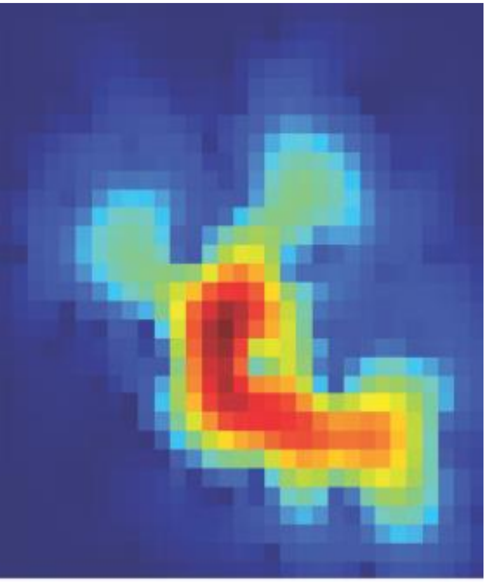




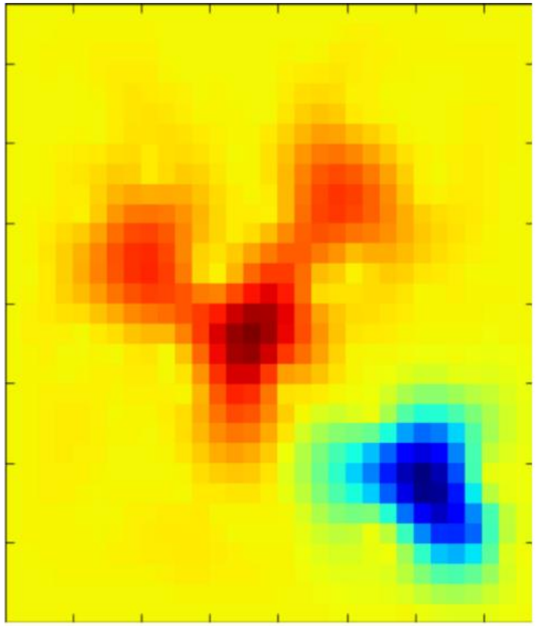


Stochastic process → tract

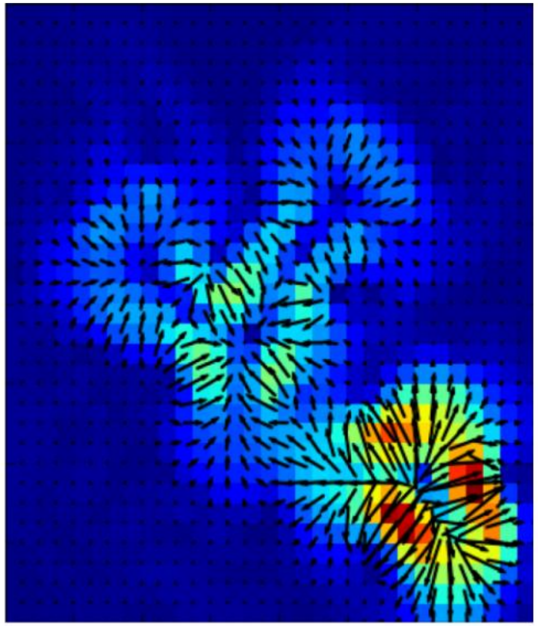
Continuity equation: $\partial_t P + \nabla \cdot J_I = 0$
Information flux: $J_I = J_d + J_c$
Diffusion (equilibrating density): $J_d = -D\nabla P$
convection (for entropy S): $J_c = -LP\nabla S$
getting $V = S$ Fokker-Planck equation: $\partial_t P + L\nabla \cdot (P\nabla S) = D\Delta P$
entropy $S_{ij} = -\sum_{\gamma:i\to j} \Pr(\gamma) \ln(\Pr(\gamma))$



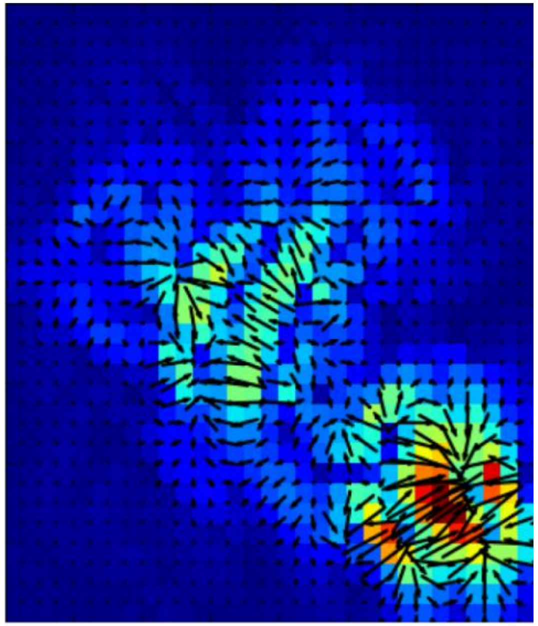
(a) e_1



(a) $S(x, y)$



(b) $\nabla S(x, y)$

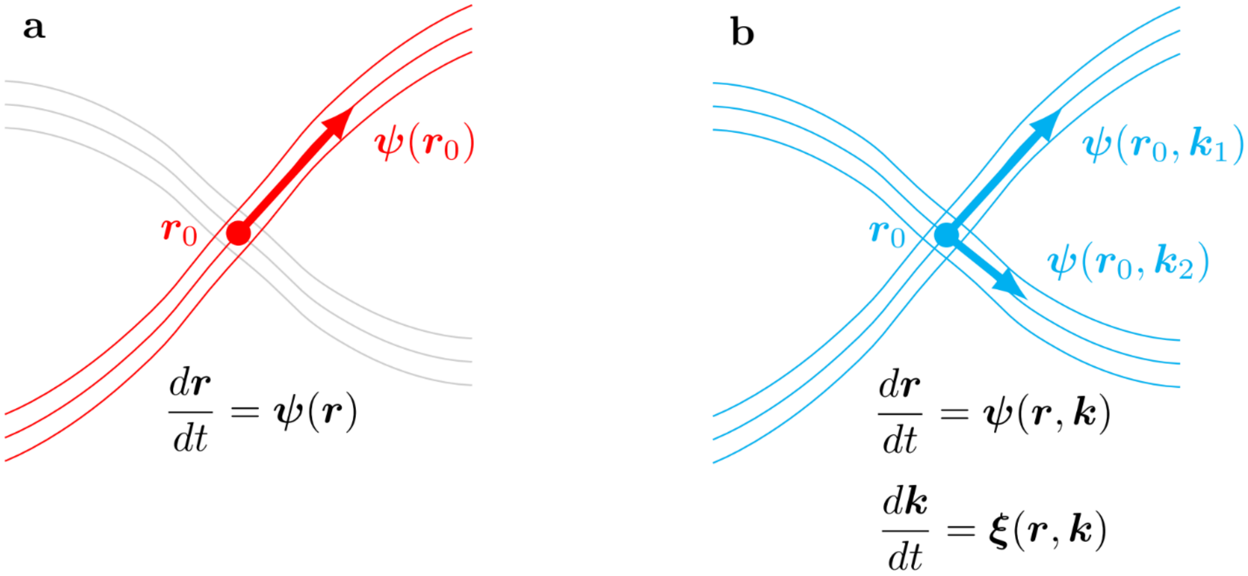


(c) $\nabla\nabla S(x, y)$

$$P = P_0 + P_1$$

Where $P_1 \ll P_0$ is a small perturbation – **we trace a wave**

$$P_1(\boldsymbol{r}, \boldsymbol{R}, t) = A(\boldsymbol{r}, \boldsymbol{R}, t)e^{i\Psi(\boldsymbol{r},t)},$$
$$\Psi(\boldsymbol{r}, t) = \boldsymbol{k} \cdot \boldsymbol{r} - \omega t,$$

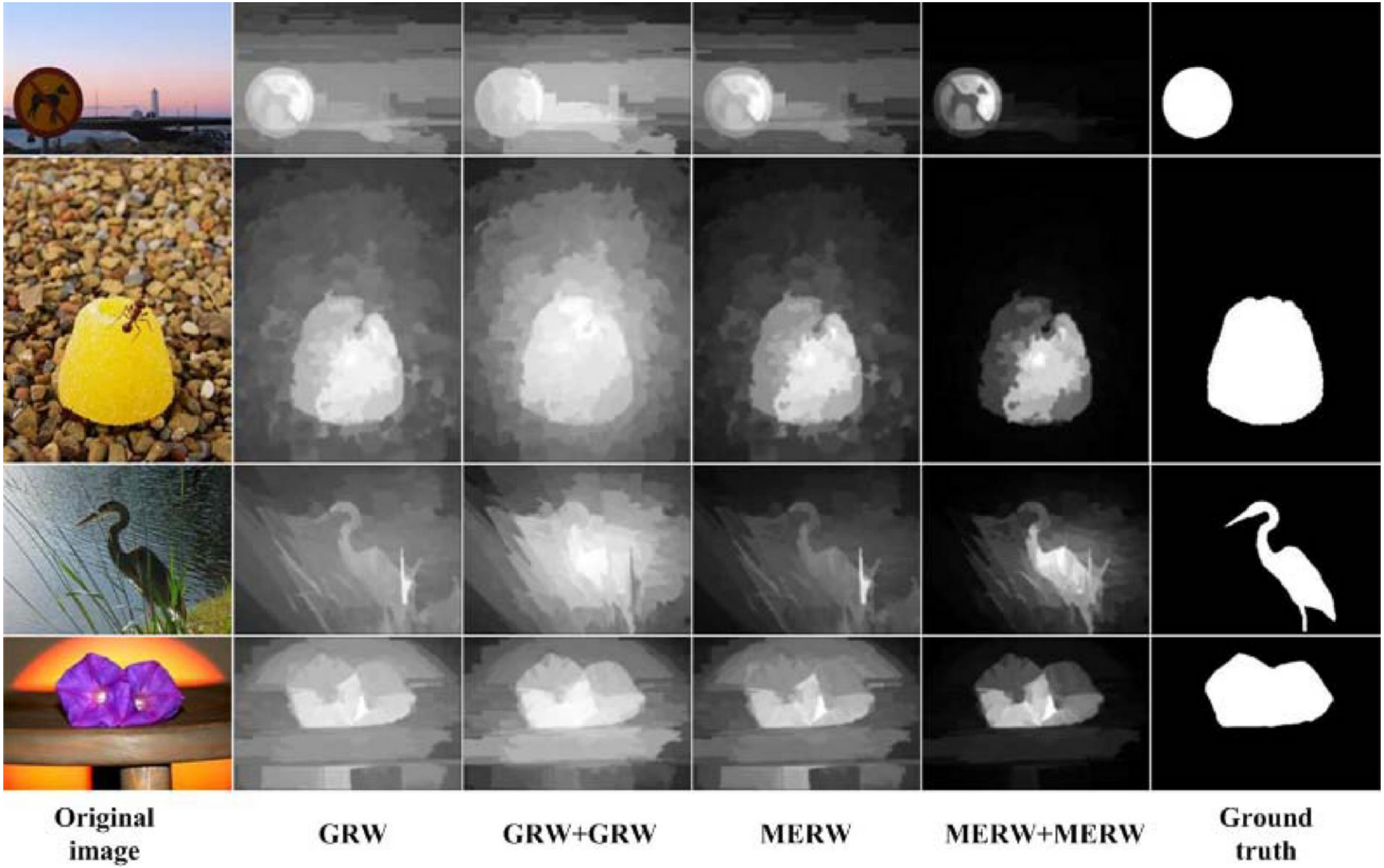


$$\frac{d\boldsymbol{r}}{dt} = \frac{2}{V_{\mathbf{R}}} \int \left[\boldsymbol{X}(\boldsymbol{k} \cdot \boldsymbol{X}) + 2\boldsymbol{k}Y(Y|\boldsymbol{k}|^2 - Z)\right] d\boldsymbol{R}, \quad (33)$$

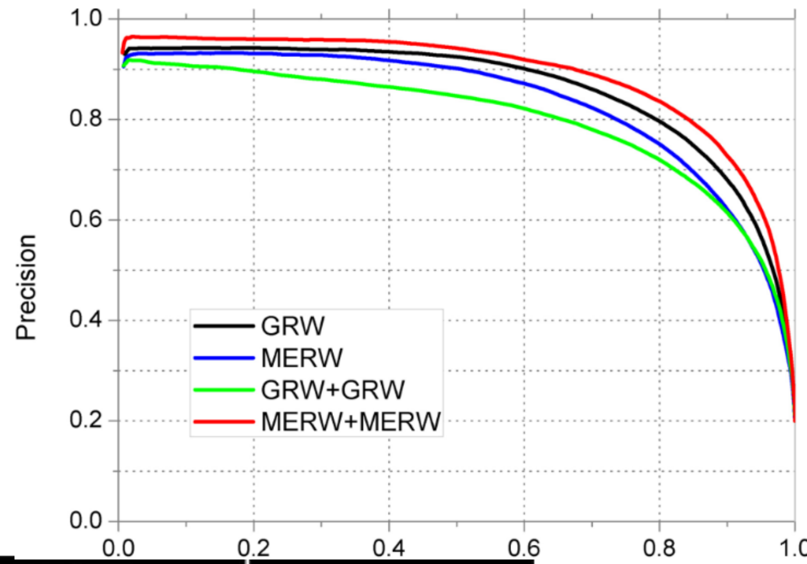
$$\frac{d\boldsymbol{k}}{dt} = -\frac{2}{V_{\mathbf{R}}} \int [(\boldsymbol{k} \cdot \boldsymbol{X})\nabla(\boldsymbol{k} \cdot \boldsymbol{X}) + (|\boldsymbol{k}|^2 Y - Z)(|\boldsymbol{k}|^2 \nabla Y - \nabla Z)] \, d\boldsymbol{R}. \quad (34)$$

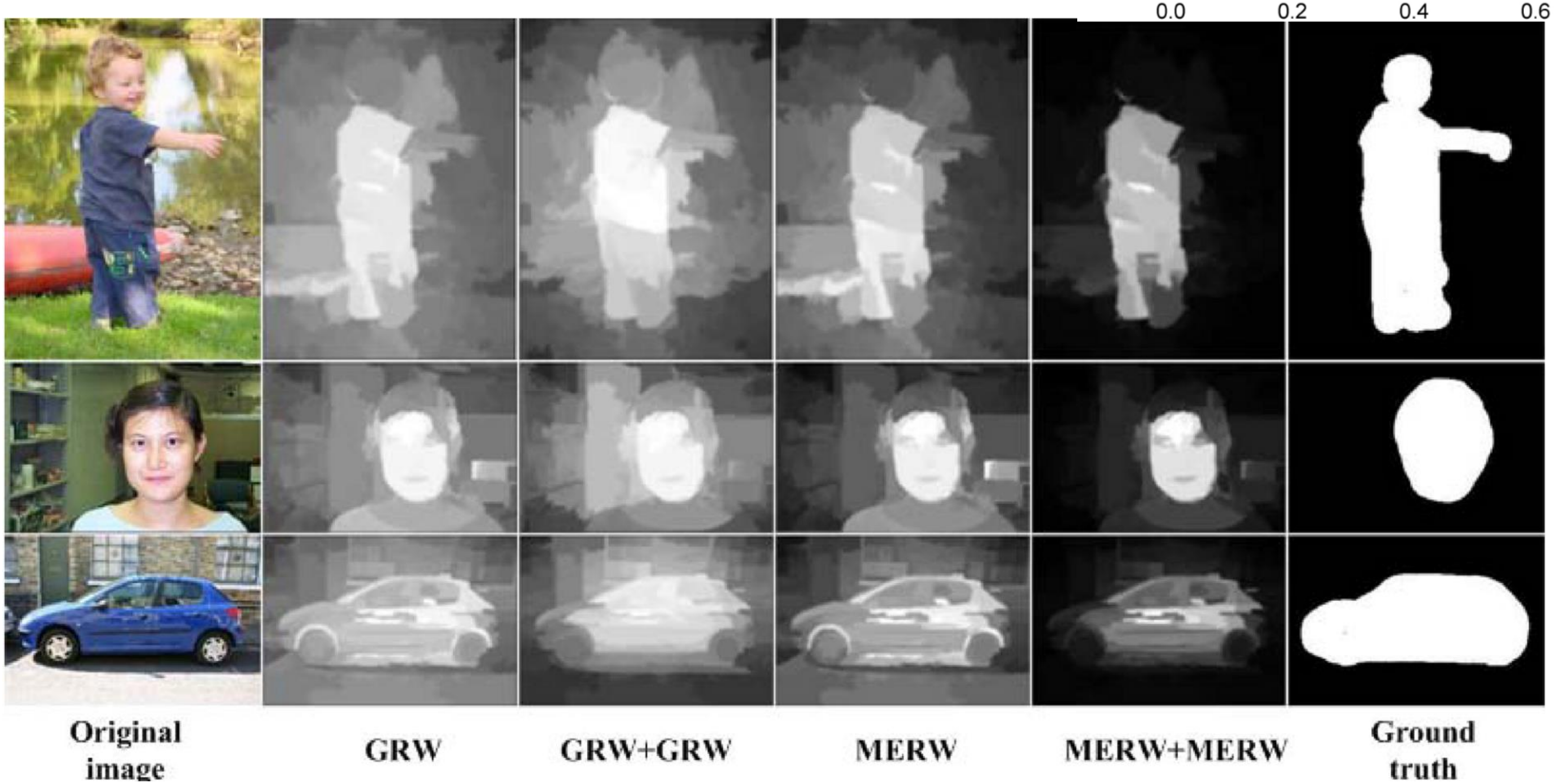
where $\boldsymbol{X} = \nabla P_0(2 + \ln P_0) + \nabla(P_0(1 + \ln P_0))$, $Y = P_0(1 + \ln P_0)$, $Z = \nabla \cdot \nabla P_0(2 + \ln P_0)$, and, again we averaged over all dynamic displacement scales.

JG Yu, J Zhao, J Tian, Y Tan, *Maximal Entropy Random Walk for Region-Based Visual Saliency* (IEEE, 2014)



- divide picture into regions (8x8 blocks, "superpixels")
- create graph among regions using similarities as weights ($w_{ij} = \exp(-d(r_i, r_j))$),
- saliency map is the stationary probability distribution of GRW or MERW



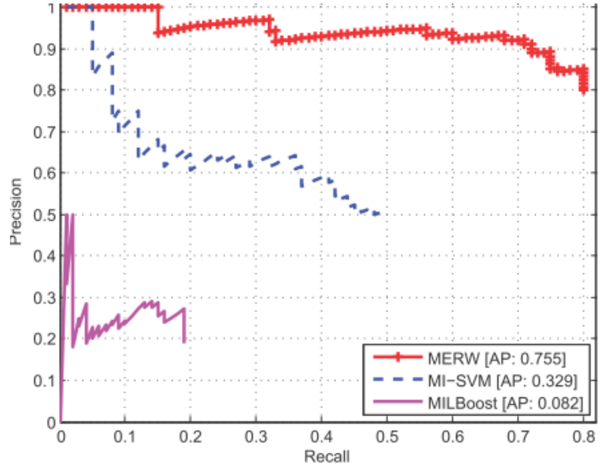


[**WEAKLY SUPERVISED OBJECT LOCALIZATION VIA MAXIMAL ENTROPY RANDOM WALK**](),

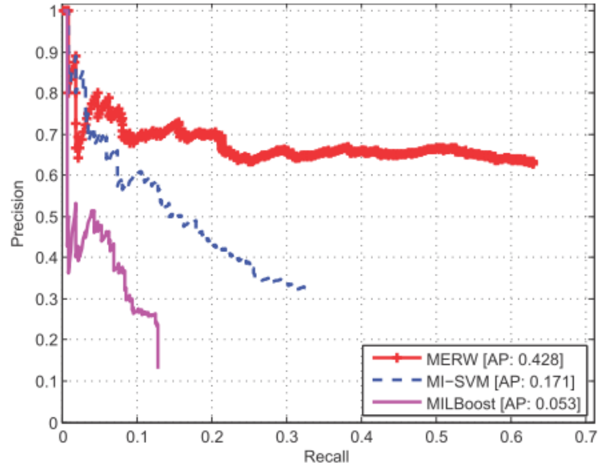
Liantao Wang, Ji Zhao, Xuelei Hu, Jianfeng Lu, IEEE ICIP 2014

Divide the picture into regions and use SVM to evaluate weights of features (w_i) for different objects (e.g. car, dog)

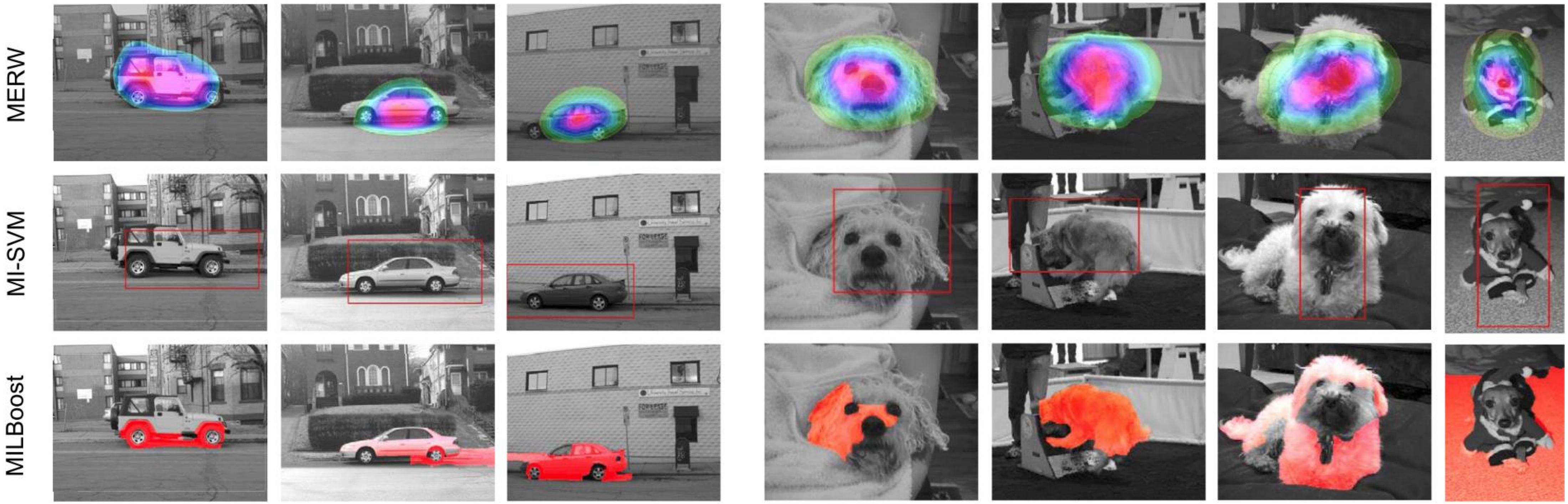
$$a_{ij} = \begin{cases} e^{\gamma(w_i + w_j)}, & \text{if} \quad z_j \in \mathcal{N}_k(z_i) \\ 0, & \text{otherwise} \end{cases}$$



(a) PittCar



(b) 'dog'



Centrality (graph theory, http://en.wikipedia.org/wiki/Centrality): **indicators which identify the most important vertices within a graph.**

Examples (for the same graph):

A) Degree centrality
($e.g. C(v) \propto \deg(v) \quad - \;$ GRW),

B) Closeness centrality
($e.g. C(v) \propto \sum_{w\neq v} 1/d(v,w)$),

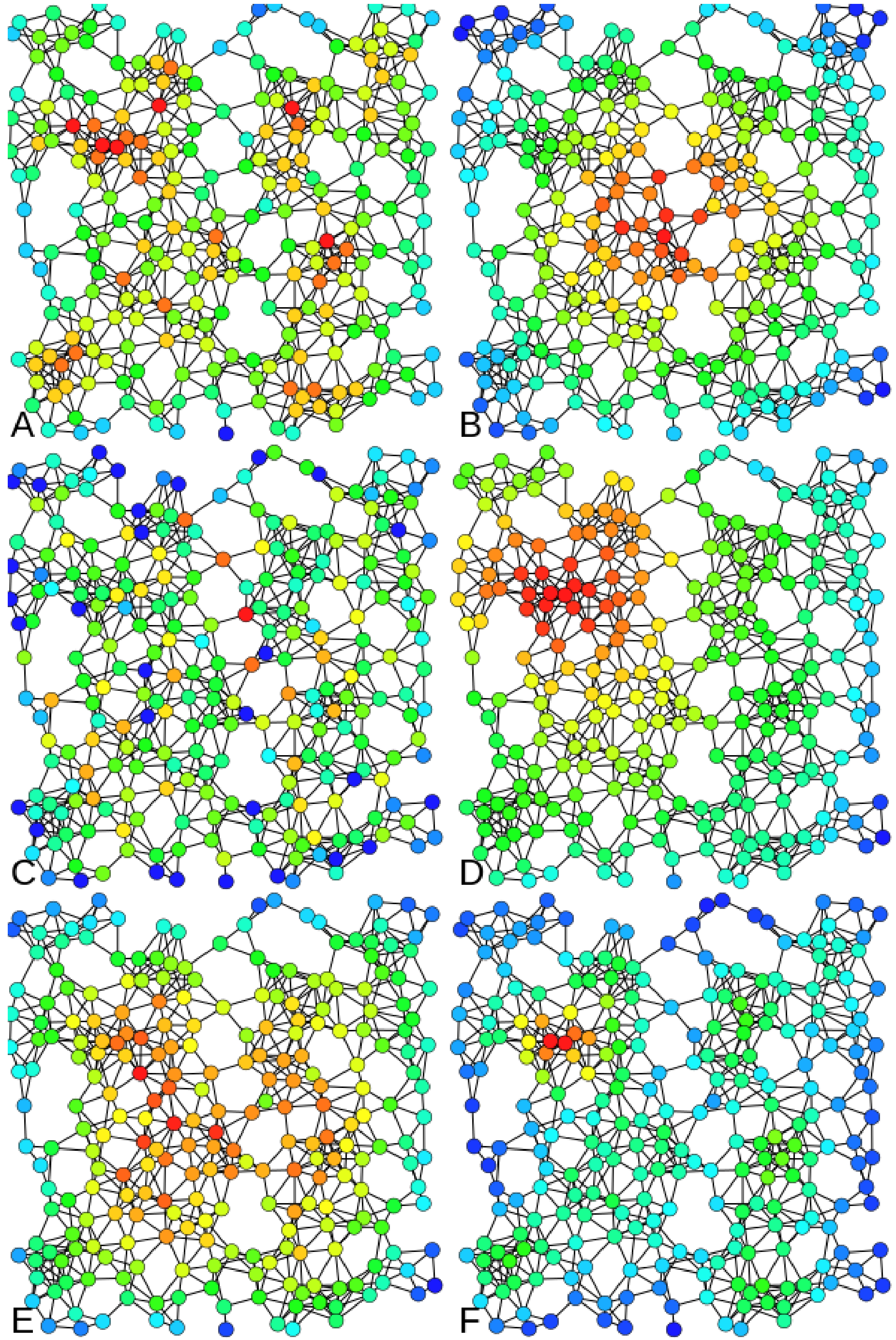
C) Betweenness centrality
(how many shortest paths go through v)

D) **Eigenvector centrality (MERW-like)**,

E) Katz centrality (e.g. PageRank),

F) Alpha centrality.

Drawing 2D diagrams for graphs:
positions from two high eigenvectors
(of M or Laplacian: $L = \mathrm{diag}(\deg(i)) - M$)



Delvenne, J.-C. & Libert, A.-S. *Centrality measures and thermodynamic formalism for complex networks*, Phys. Rev. E 83, 046117 (2011).

(e.g. Google) PageRank (GRW) → Entropy Rank (MERW)

(α = Pr(going to a random page), $E = e^{-U_0}$ weight out of the graph edges)

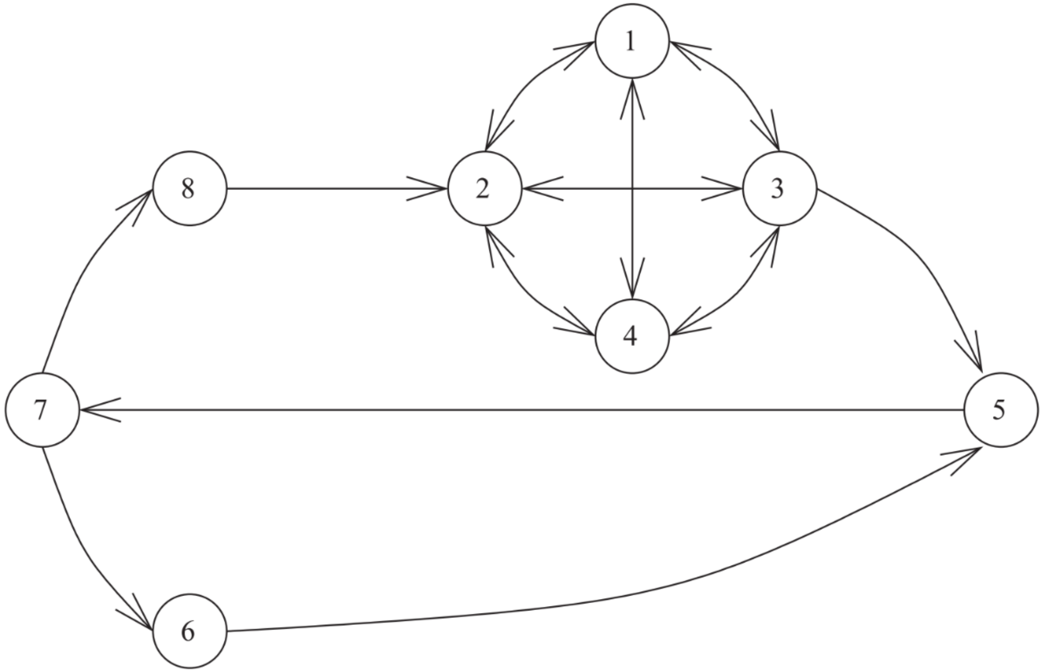
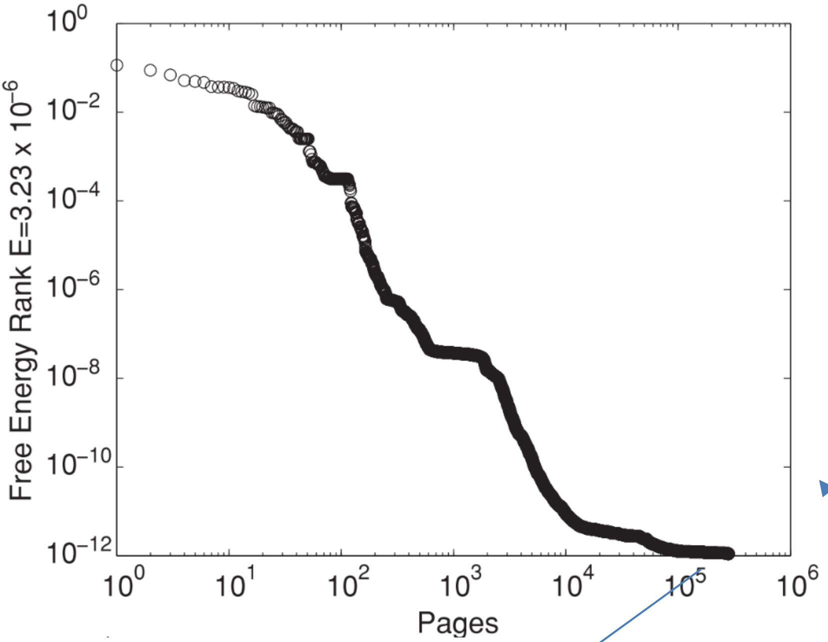


TABLE I. PageRank, free-energy rank, and entropy rank for the network of Fig. 1.

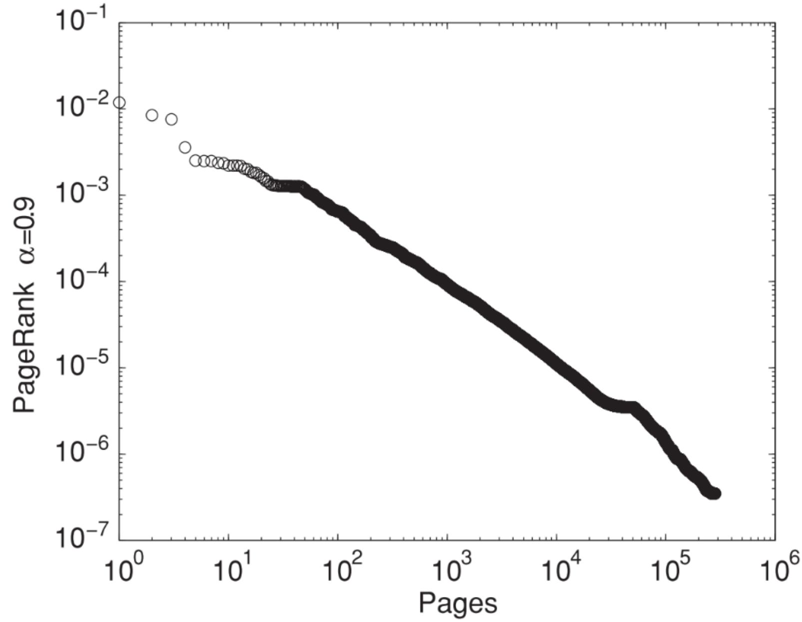
Vertex	PageRank ($\alpha = 1$)	PageRank ($\alpha = 0.9$)	Entropy rank	Free-energy rank ($E = 0.03$)
1	0.1705	0.1549	0.2464	0.2400
2	0.2045	0.1965	0.2487	0.2458
3	0.1818	0.1644	0.2487	0.2460
4	0.1705	0.1549	0.2464	0.2400
5	0.0909	0.1035	0.0032	0.0099
6	0.0455	0.0601	0.0001	0.0019
7	0.0909	0.1057	0.0032	0.0076
8	0.0455	0.0601	0.0031	0.0087

- vertex 8 becomes more interesting than 6 (pointing to “good pages”),
- cliques are swelling (localization) – problem with “link farms” ...

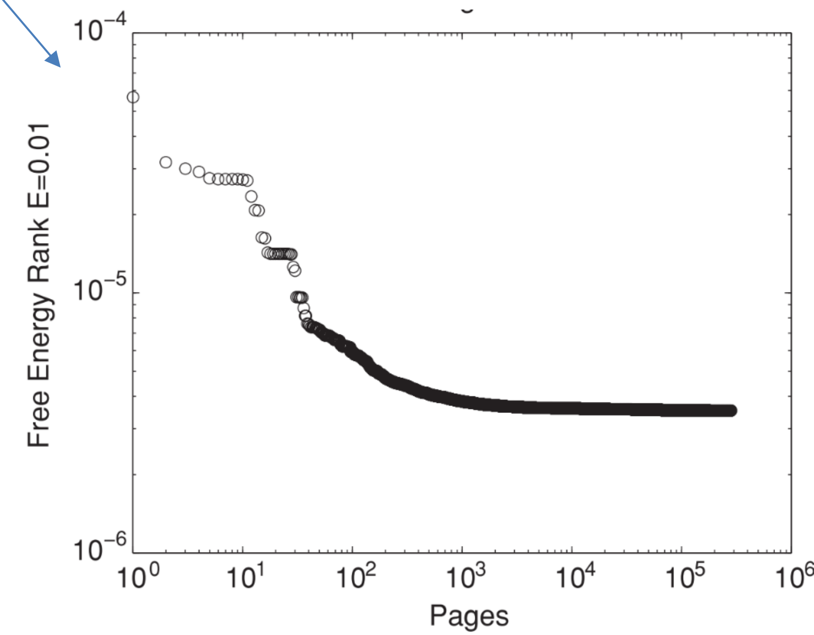
Experiments on "289 000 – node piece of the Stanford web (http://www.kamvar.org/)"



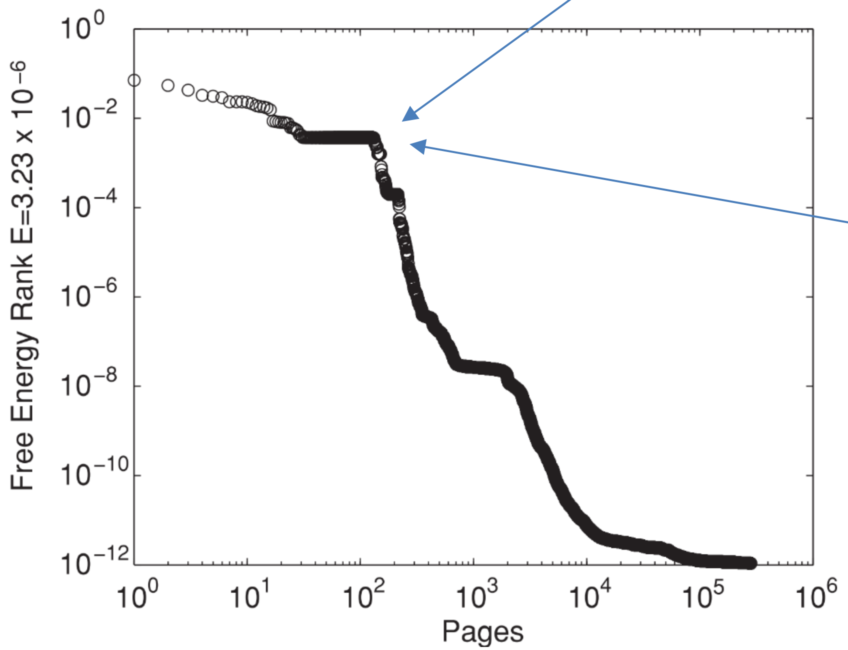
PageRank



High E FER
(good for finding
best pages)



low H FER



low H FER
vertex with added
100 vert. clique
("farm link")
200 000th → 627th
(plateau → clique ?)

Mean first-passage time (MFPT) (e.g. for community detection)

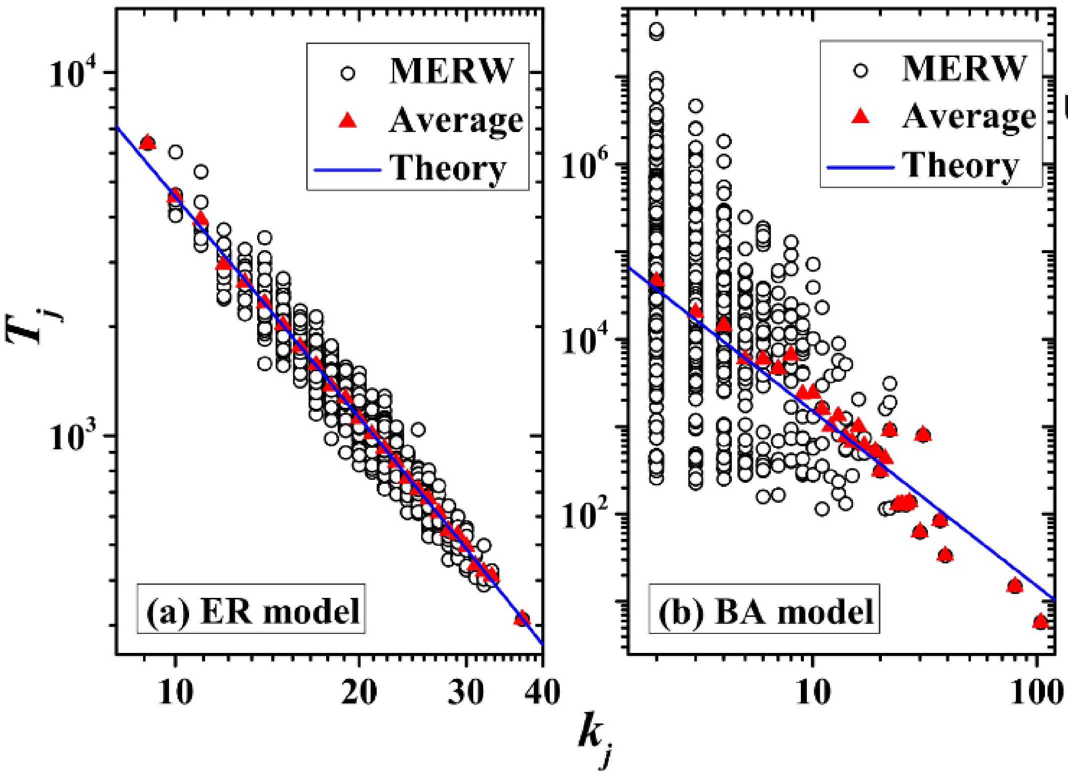
M_{ij} – expected minimal time to reach vertex j starting from i.

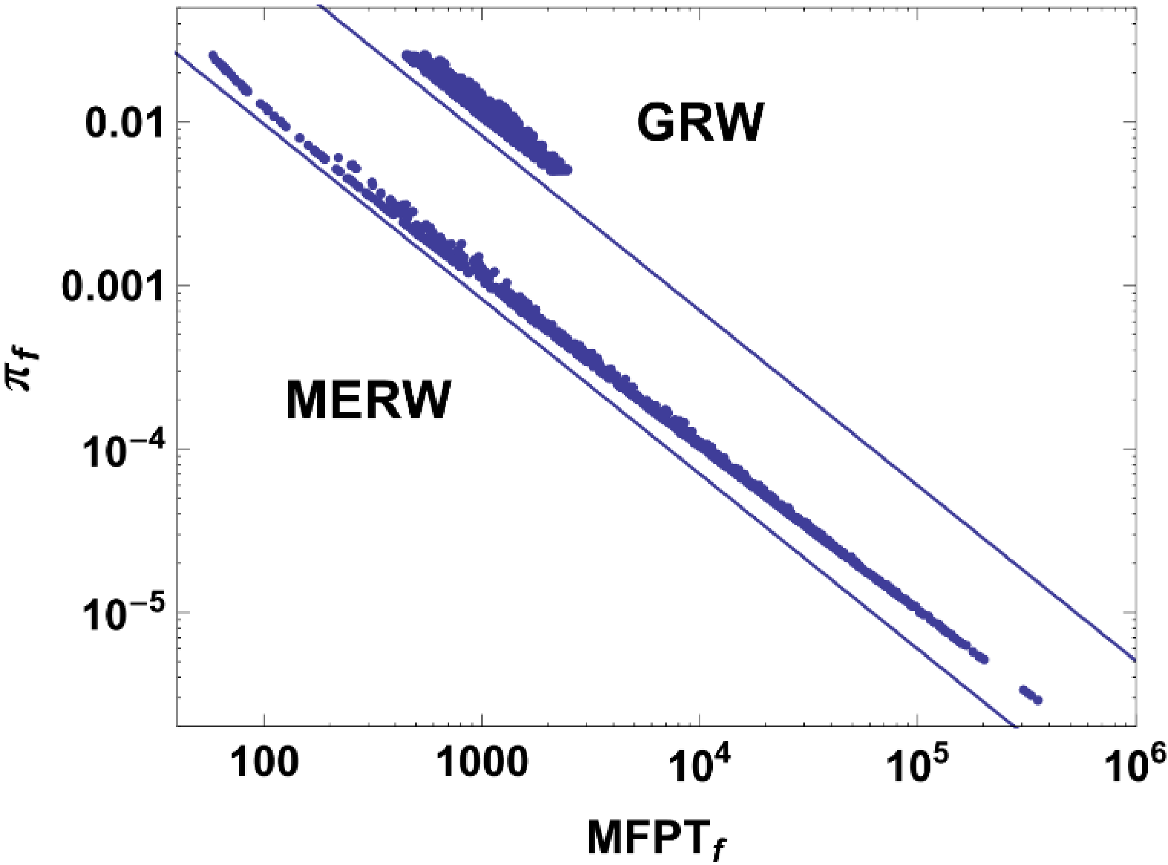
Y. Lin, Z. Zhang, *Mean first-passage time for maximal-entropy random walks in complex networks* (Nature, 2014)

Erdős–Rényi (ER): $\Pr(\to v_j) = \text{const}$

Barabási–Albert (BA): $\Pr(\to v_j) \propto k_j$

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





1000 vertices

J. Ochab, *Maximal-entropy random walk unifies centrality measures* (Phys. Rev. E, 2012)

SimRank: measure how similar two vertices are

G. Jeh and J. Widom. *Simrank: a measure of structural-context similarity* (KDD 2002)

$s(a,b) = \frac{C}{|N(a)||N(b)|} \sum_{x\in N(a)} \sum_{y\in N(b)} s(x,y)$ $\quad$ $(1\ if\ a=b,\ 0\ if\ I(a)\cap I(b)=\emptyset)$

can be expressed by Expected$-f$ Meeting Distance (EMD) of two walkers (a,b)

$s'(a,b) = \sum_{t:(a,b)\rightsquigarrow(x,x)} P[t]\, f\big(l(t)\big)$ $\quad$ for $f(z)=z$ $\quad$ or $\quad$ $f(z)=C^z$

$P[t]$ - GRW probability of path t

Link prediction – which new interactions (links) are likely to occur?

Predicting evolution, suggesting connections, finding weak/fake links

The more similar they are, the more likely they will connect

Li, R. H., Yu, J. X. & Liu, J. *Link prediction: the power of maximal entropy random walk* (ACM conference, 2011):

Replace GRW with MERW in $P[t]$, getting $S(a,b) = \frac{C\psi_a\psi_b}{\lambda^2} \sum_{x\in N(a)} \sum_{y\in N(b)} \frac{S(x,y)}{\psi_x\psi_y}$

MERW – more distinctive, scale-free (does not depend on discretization)

27 link prediction methods (AUC: the higher the better), “ME” – maximal entropy

SM	ER	BA	SW	USAir	C.ele	Yeast	Power	NetSci	GrQc	HepPh	HepTh
CTT	0.710	0.750	0.791	0.847	0.784	0.709	0.713	0.917	0.520	0.523	0.525
CTME	0.720	0.746	0.745	0.855	0.798	0.501	0.501	0.866	0.556	0.645	0.534
CK	0.805	0.883	0.804	0.856	0.809	0.715	0.501	0.799	0.513	0.501	0.513
MECK	0.940	0.981	0.845	0.936	0.856	0.757	0.501	0.975	0.517	0.501	0.503
NCK	0.502	0.501	0.501	0.708	0.706	0.501	0.501	0.501	0.503	0.508	0.501
NMECK	0.903	0.983	0.982	0.931	0.969	0.710	0.501	0.971	0.623	0.750	0.675
DK	0.835	0.813	0.983	0.836	0.838	0.829	0.764	0.965	0.501	0.605	0.593
MEDK	0.999	0.983	0.998	0.991	0.971	0.749	0.812	0.963	0.739	0.735	0.746
NDK	0.786	0.711	0.956	0.920	0.778	0.731	0.857	0.908	0.531	0.530	0.530
NMEDK	**0.999**	**0.983**	**0.998**	**0.997**	**0.978**	**0.970**	0.857	**0.996**	**0.739**	0.755	**0.758**
RK	0.851	0.907	0.973	0.898	0.887	0.803	0.864	0.624	0.632	0.608	0.561
MERK	0.999	0.983	0.998	0.981	0.949	0.812	0.812	0.963	0.618	0.745	0.735
NRK	0.504	0.501	0.501	0.719	0.501	0.703	0.806	0.501	0.501	0.508	0.504
NMERK	0.999	0.983	0.998	0.983	0.975	0.968	0.857	0.986	0.739	0.755	0.756
MENK	0.999	0.983	0.998	0.936	0.975	0.799	0.812	0.963	0.618	0.730	0.746
NNK	0.503	0.501	0.501	0.819	0.501	0.705	0.806	0.501	0.501	0.508	0.504
NMENK	0.999	0.983	0.998	0.983	0.965	0.965	0.857	0.996	0.739	0.755	0.752
PD	0.926	0.974	0.953	0.971	0.866	0.887	0.857	0.722	0.666	0.618	0.628
MEPD	0.999	0.976	0.998	0.993	0.964	0.968	0.857	0.913	0.739	0.755	0.758
PDM	0.805	0.764	0.957	0.972	0.798	0.886	0.857	0.874	0.616	0.660	0.530
MEPDM	0.999	0.983	0.998	0.990	0.976	0.970	0.857	0.996	0.739	0.755	0.758
SR	–	–	–	0.905	0.860	–	–	0.955	–	–	–
MESR	–	–	–	0.960	0.876	–	–	0.963	–	–	–
CN	0.884	0.782	0.501	0.386	0.971	0.752	0.802	0.961	0.617	0.623	0.635
AA	0.886	0.781	0.501	0.409	0.975	0.793	0.806	0.969	0.623	0.630	0.638
HPLP+	0.983	0.971	0.978	0.979	0.974	0.965	**0.886**	0.984	0.725	0.753	0.732
SRW	0.991	0.977	0.989	0.983	0.972	0.967	0.863	0.983	0.731	**0.760**	0.754

Kernel between G and G': $k(G,G') = q_\times^T \cdot \left(\sum_{k\geq 0} \mu(k)\, W_\times^k\right) \cdot p_\times$ e.g. $(1-\lambda W_\times)^{-1}$ or $e^{\lambda W_\times}$

NMEDK – normalized maximal entropy heat diffusion kernel, **NMERK** - ...Laplacian kernel

$S_{ij}^{GRW} = \frac{M_{ij}}{\deg(i)}$ $\quad \pi_i^{GRW} \propto \deg(i)$ $\qquad (S^{MERW})_{ij}^t = \frac{(M)^t{}_{ij}}{\lambda^k}\frac{\psi_j}{\psi_j}$ $\quad \pi_i^{MERW} \propto \psi_i^2$

GRW Laplacian $(M_{ii}=0)$: $\Delta_{ij} = -\mathrm{L}_{ij} = M_{ij} - \deg(i)\cdot\delta_{ij}$ $\quad \left(w^T \mathrm{L} w = \sum_{\{i,j\}\in E}\left(w_i - w_j\right)^2\right)$

In analogy to discretized continuous Laplacian: $(\partial_{xx} w)(x) \approx w(x-1) - 2w(x) + w(x+1)$

Or relaxation of capacitor network: $\frac{d}{dt} Z_i(t) = \sum_{j:\, i\sim j}\left(Z_j(t) - Z_i(t)\right)$.

General Laplacian $\left(\textbf{"continuity equation"}: \forall_i \sum_j L_{ij} = 0,\ M_{ij} = M_{ji} \Rightarrow \Pr(i,j) = \Pr(j,i)\right)$**:**

(const ·) $\quad \Delta_{ij} = \left(\Pi(S-\mathbf{1})\right)_{ij} = \Pr(i,j) - \Pr(i)\cdot\delta_{ij}$ $\qquad$ **MERW:** $\Delta_{ij} = M_{ij}\frac{\psi_i\psi_j}{\lambda} - \psi_i^2\cdot\delta_{ij}$

Normalized MERW Laplacian: $\left(\Delta_{\mathrm{sym}}\right)_{ij} = \frac{M_{ij}}{\lambda} - \delta_{ij}$

Heat equation and **kernel**: $\frac{d}{dt}K_t = \Delta K_t$ $\qquad K_t = \exp(t\Delta) = \lim_{n\to\infty}\left(1+\frac{t\Delta}{n}\right)^n = \sum_k \frac{(t\Delta)^k}{k!}$

MEPDM – maximal entropy inverse p-distance with matrix exponentiation

Inverse P-distance: $P(i,j) = \sum_{t_{ij}: i \rightsquigarrow j} P[t]\cdot\alpha^{l(t_{ij})}$ $\quad$ (or $\ \alpha^l/l!$)

for MERW : $l(t_{ij}) = l(t'_{ij}) \Rightarrow P[t_{ij}] = P[t'_{ij}]$ $\quad$ so $\quad P(i,j) = \frac{\psi_j}{\psi_i}\sum_{l\geq 1}\left(\frac{\alpha}{\lambda}\right)^l (A^l)_{ij}$

Hitting/commute time (MFPT): $h(i,j) = [i\neq j]\left(1 + \sum_k S_{ik} h(k,j)\right)$ $\qquad c(i,j) = h(i,j) + h(j,i)$

MERW – the most random among random walks

uniform distribution among paths, not edges (GRW)

- As the choice of statistical parameters of an **information channel**

MERW allows to maximize channel capacity under some constraints (language?)

- **As random walk/diffusion** (scale-free)

GRW: the walker indeed performs successive random decisions

MERW: only represents our (lack of) knowledge about a complex dynamics

- **For metrics to analyze complex network**

GRW sees only degrees of vertices, poorly distinguish nodes

MERW allows to evaluate importance in the space of possible paths

- social/evolutionary entropy (Lloyd Demetrius):

"thinking" in terms of paths (reason→result chains) of possibilities?

GRW → MERW

in many cases improves performance or agreement