GRAPH DYNAMICAL SYSTEMS

Some combinatorial problems related to Markov chains

AG15, NIMS, August 3, 2015

Yaokun Wu

Shanghai Jiao Tong University
Shanghai, **China**

TABLE OF CONTENTS

TOPOLOGICAL MARKOV CHAINS

Let K be a k-element set. The set of probability distributions on K is the $(k-1)$-dimensional probability simplex

$$\Delta^K = \{x \in \mathbb{R}^K = \mathbb{R}_{k\times 1}, x \geq 0, \sum_{i\in K} x_i = 1\}.$$

A **Markov chain of order t on K**, also known as a Markov chain with memory t or a t-step Markov shift on K, is a sequence of points $x(0), x(1), \ldots$ in Δ^K where each point $x(1+t+h)$ is linearly determined by its earlier t points $x(1+h), \ldots, x(t+h)$, written as $x(1+t+h) = Mx(1+h)\cdots x(t+h)$, where M is a t-linear map (hypermatrix/tensor of order $t+1$) and

$$(Mx(1+h)\cdots x(t+h))_i = \sum_{i_1,\ldots,i_t\in K} M_{ii_t\ldots i_1} x(1+h)_{i_1} \cdots x(t+h)_{i_t}.$$

FROM PROBABILISTIC TO TOPOLOGICAL

- A Markov chain of order t on K is determined by the initial t points from Δ^K as well as the t-linear map M which sends $\underbrace{\Delta^K \times \cdots \times \Delta^K}_{t}$ into Δ^K. If M varies in time, we have an inhomogeneous chain; If M is constant, it is a homogeneous chain.
- We replace Δ^K by 2^K (mapping a probability vector to its support) and choose coefficients from the Boolean semiring instead of from nonnegative reals, thus arriving at a nonparametric version of a Markov chain, called a **topological Markov chain**. The probability transition tensor now becomes a **Boolean tensor**. If the chain is homogeneous, the topological chain is known as a **t-step subshift of finite type** in symbolic dynamics [2, 3].

ONE-STEP TOPOLOGICAL MARKOV CHAIN

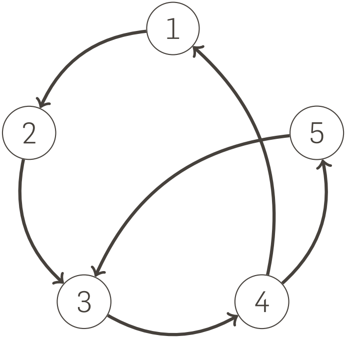


Abbildung: A digraph f representing a nonparametric 1-step Markov chain.

ONE-STEP TOPOLOGICAL MARKOV CHAIN

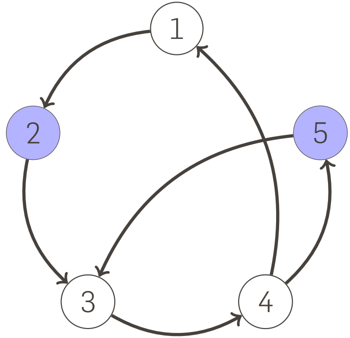


Abbildung: A digraph f and its evolving vertex subset

$\{2, 5\}$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

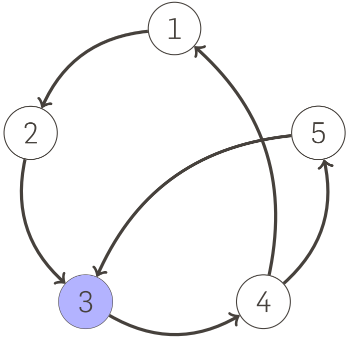


Abbildung: A digraph f and its evolving vertex subset

$\{2,5\} \to \{3\}$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

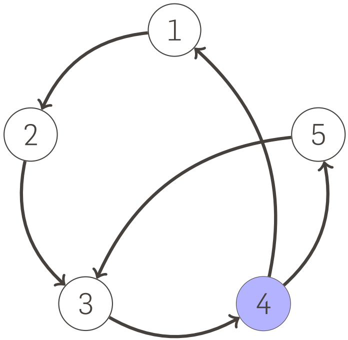


Abbildung: A digraph f and its evolving vertex subset

$\{2,5\} \to \{3\} \to \{4\}$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

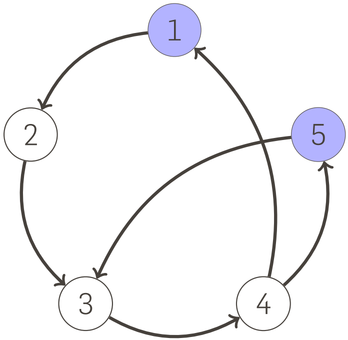


Abbildung: A digraph f and its evolving vertex subset

$\{2,5\} \to \{3\} \to \{4\} \to \{1,5\}$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

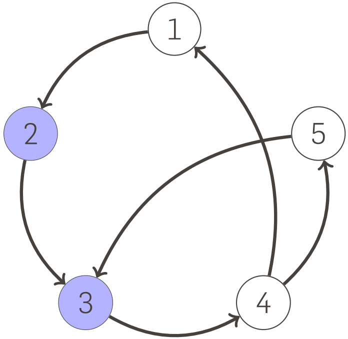


Abbildung: A digraph f and its evolving vertex subset

$\{2,5\} \to \{3\} \to \{4\} \to \{1,5\} \to \{2,3\}$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

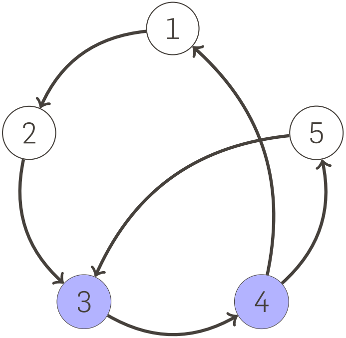


Abbildung: A digraph f and its evolving vertex subset

$$\{2,5\} \to \{3\} \to \{4\} \to \{1,5\} \to \{2,3\} \to \{3,4\}$$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

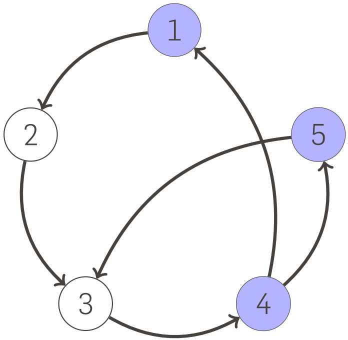


Abbildung: A digraph f and its evolving vertex subset

$$\{2,5\} \to \{3\} \to \{4\} \to \{1,5\} \to \{2,3\} \to \{3,4\} \to \{1,4,5\}$$

ONE-STEP TOPOLOGICAL MARKOV CHAIN

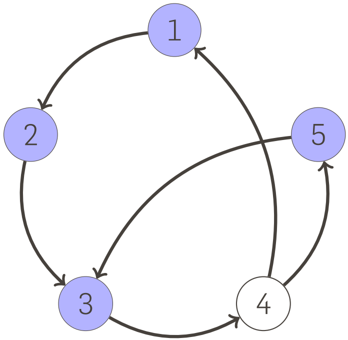


Abbildung: A digraph f and its evolving vertex subset

$$\{2,5\} \to \{3\} \to \{4\} \to \{1,5\} \to \{2,3\} \to \{3,4\} \to \{1,4,5\} \to \{1,2,3,5\}$$

PHASE SPACE $\mathcal{PS}_F$ OF THE DIGRAPH F

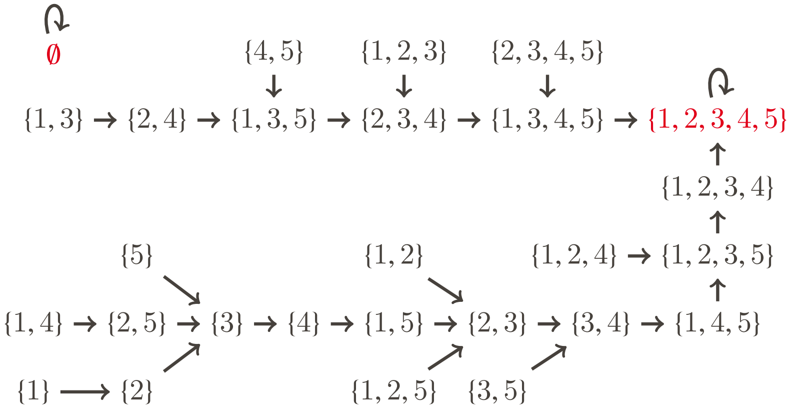


Abbildung: Besides ∅, all vertices will eventually visit $V(f)$.

PHASE SPACE $\mathcal{PS}_F$ OF THE DIGRAPH F

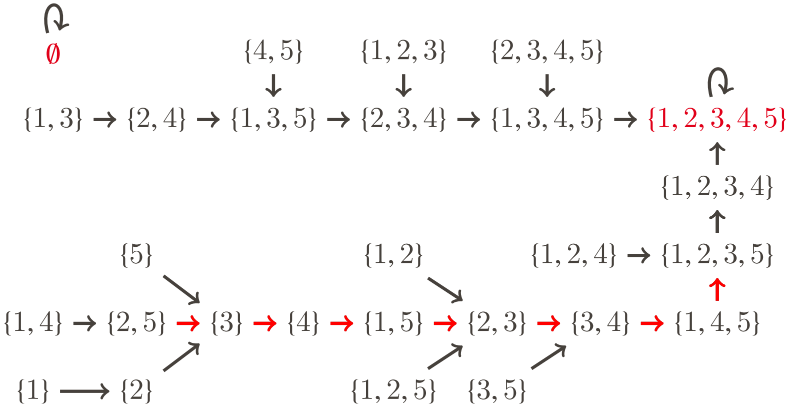


Abbildung: In seven steps $\{2, 5\}$ moves to a set which properly contains itself.

The digraph f stands for the local connection mechanism and its phase space $\mathcal{PS}_f$ displays the global evolving picture.

The theory of dynamical systems aims to relate a system's global behaviour to its local behaviour and the forces that shape it.

PRIMITIVITY: FROM SINGLETON SET TO WHOLE VERTEX SET

A Boolean linear map f on K is called **primitive** if every vertex from $2^K \setminus \{\emptyset\}$ will reach K in $\mathcal{PS}_f$. If f is primitive, the length of a longest path in $\mathcal{PS}_f$ is called the **primitive exponent** of f and is denoted $g(f)$.

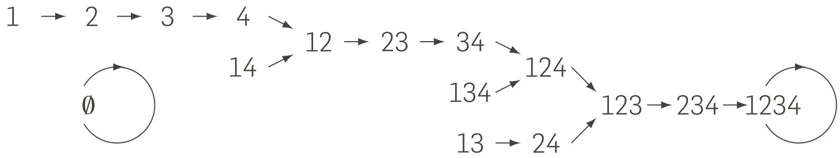
Every digraph f is identified with a corresponding Boolean linear map.

WIELANDT-TYPE MATRICES

Take a positive integer $k \geq 2$ and choose $i \in \{1, \ldots, k-1\}$ satisfying $\gcd(i, k) = 1$. A **Wielandt-type matrix/digraph** $W_{k;i}$ is the matrix/digraph with vertex set $\mathbb{Z}/k\mathbb{Z}$ and arc set $\{i \to i+1 : \ i = 1, \ldots, k\} \cup \{k \to 1+i\}$.

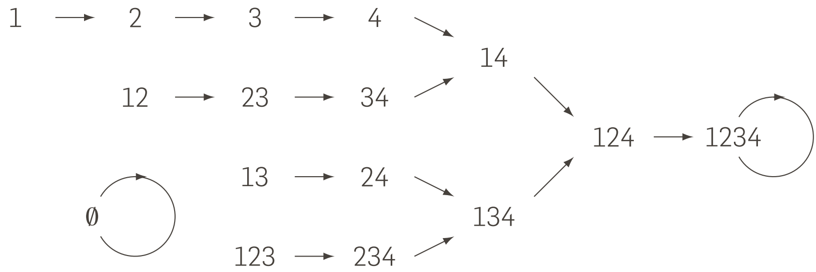
$$W_{4;1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, W_{4;3} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, W_{5;4} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$\mathcal{PS}_{W_{4;1}}$



$g(W_{4;1}) = 10 = (4-1)^2 + 1.$

$\mathcal{PS}_{W_{4;3}}$



In $6 = 2 \times 4 - 2$ steps every vertex other than $\emptyset$ reaches 1234.

Helmut Wielandt (1959)

Let f be a primitive digraph with k vertices. Then the primitive exponent of f is at most $(k-1)^2+1$ and the bound is attained if and only if the digraph f is isomorphic to $W_{k;1}$.

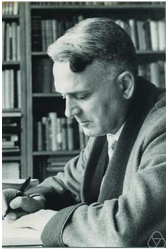


Abbildung: Helmut Wielandt, 19 December 1910 – 14 February 2001

MORE GENERAL REACHABILITY RESULTS

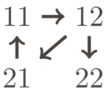
W., Zeying Xu, Yinfeng Zhu

Let k be a positive integer and let f be a digraph on $[k]$. Let $Y = f^h(X)$ where $X \subseteq [k]$ and $h \in \mathbb{N}$.

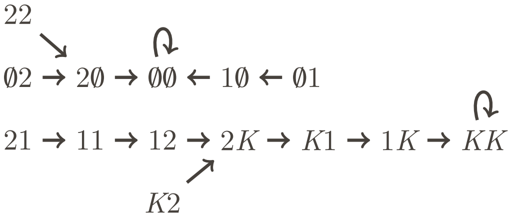
(A) If Y falls into a strongly connected component of f, then there exists $h' \in \mathbb{N}$ such that $h' \leq |Y| L_f$ and $Y = f^{h'}(X)$, where L_f is the length of a longest path in f.

(B) If f itself is irreducible, then there exists $h' \in \mathbb{N}$ such that $h' \leq |Y| \mathrm{D}_f$ and $Y = f^{h'}(X)$, where D_f is the diameter of f.

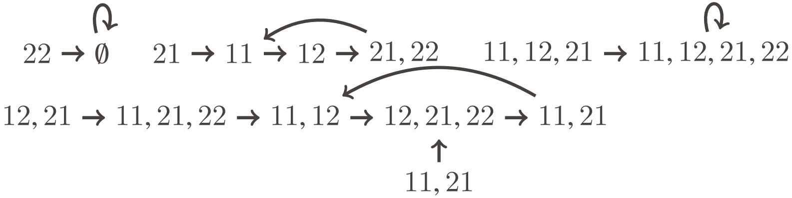
WAVE-PARTICLE DUALITY FOR HIGHER-ORDER MARKOV CHAIN



A 2-step MC on $K = [2]$

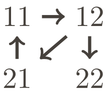


Phase space of the 2-step MC (wave view)

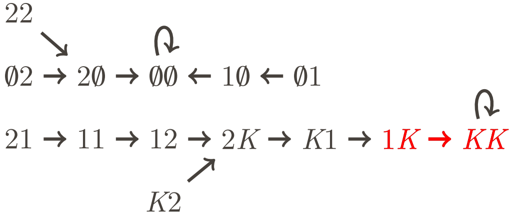


Phase space of the encoded 1-step MC on 2^K (particle view)

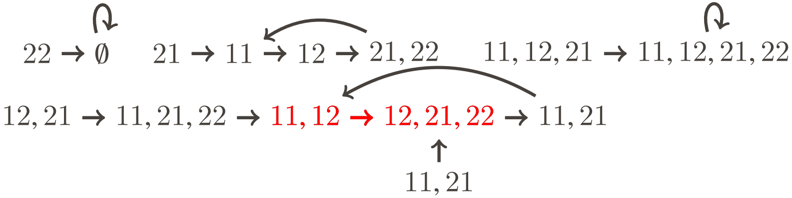
WAVE-PARTICLE DUALITY FOR HIGHER-ORDER MARKOV CHAIN



A 2-step MC on $K = [2]$



Phase space of the 2-step MC (wave view)



Phase space of the encoded 1-step MC on 2^K (particle view)

PRIMITIVE TENSOR (DIRECTED HYPERGRAPH)

A Boolean t-linear map f on K is called **primitive** if every vertex from $(2^K \setminus \{\emptyset\})^t$ will reach $(K, \ldots, K)$ in $\mathcal{PS}_f$. If f is primitive, the length of a longest path in $\mathcal{PS}_f$ is called the **primitive exponent** of f.

Every spanning subdigraph f of the $(t-1)$th line digraph of a complete digraph (a De Bruijn digraph) is identified with a Boolean t-linear map.

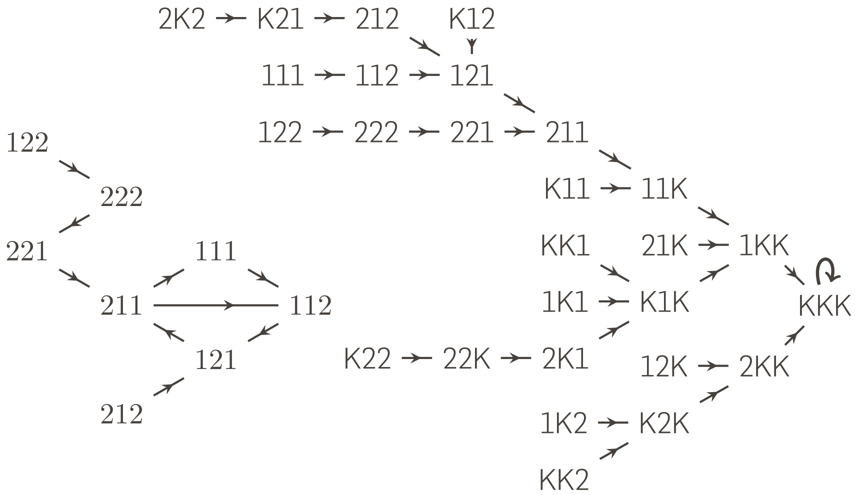


Abbildung: A primitive 3-linear map and part of its phase space.

Conjecture (W., Zeying Xu, Yinfeng Zhu)

Take an integer $k \geq 3$. The maximum primitive exponent for which a primitive Boolean 2-linear map on $[k]$ can achieve is $5k^2 - 8k + 2$.

The Perron-Frobenius Theory for nonnegative matrices basically give all information of a 1-step Markov chain.

For a higher-order Markov chain, to establish corresponding higher-order Perron-Frobenius Theory has been a hot topic in the study of tensors.

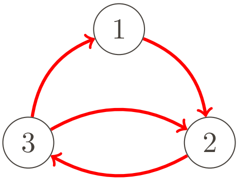
The process of going from linear to multilinear/polynomial seems to suggest many nontrivial problems.

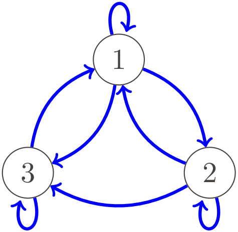
PRIMITIVE INHOMOGENEOUS CHAIN

The Boolean linear dynamical system $(2^K, \mathcal{F})$ is **primitive** provided every walk of length $2^k - 2$ in $\mathcal{PS}_\mathcal{F}$ starting from a vertex other than $\emptyset$ will reach K. If $\mathcal{F}$ is primitive, the length of a longest path in $\mathcal{PS}_\mathcal{F}$ is called the **primitive exponent** of $\mathcal{F}$ and is denoted by $g(\mathcal{F})$.

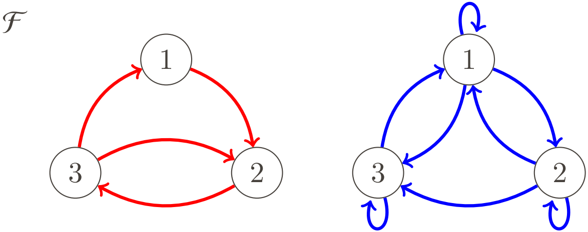
PRIMITIVE DIGRAPH SET

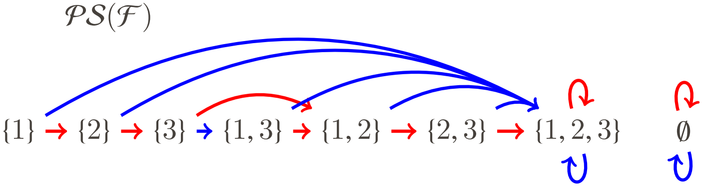
$\mathcal{F}$





PRIMITIVE DIGRAPH SET





In $2^3 - 2 = 6$ steps every nonempty set is arriving at $\{1, 2, 3\}$.

A PROBLEM OF PROTASOV

Let $g_{k,t}$ be the maximum possible value of $g(\mathcal{A})$ where $\mathcal{A}$ is a primitive set of Boolean linear maps on $[k]$ of size t. Protasov suggested to estimate $g_{k,t}$.

V. Yu. Protasov, *Semigroups of non-negative matrices*, Communications of the Moscow Mathematical Society 65 (2010) 1186–1188.

Besides Wielandt's bound of $g_{k,1} = (k-1)^2 + 1$ and the trivial bound of $g_{k,t} \le 2^k - 2$, very little is known about $g_{k,t}$.

$$\begin{cases} g_{2,2} = 2 = 1 \times 2 = 2^2 - 2 \\ g_{3,2} = 6 = 2 \times 3 = 2^3 - 2 \\ g_{4,2} = 12 = 3 \times 4 = 2^4 - 2^2 \\ g_{5,2} \ge 23 \quad (24 = 2^5 - 2^3 < 30 = 2^5 - 2) \\ g_{6,2} \ge 39 \quad (48 = 2^6 - 2^4) \end{cases}$$

GENERAL REACHABILITY

Conjecture (W., Zeying Xu, Yinfeng Zhu)

Let $\mathcal{F}$ be a primitive set of Boolean linear maps on $[k]$. For any $A, B \in 2^{[k]}$, it holds $\mathrm{dist}_{\mathcal{PS}_{\mathcal{F}}}(A, B) \leq |B|k^{|\mathcal{F}|}$ as long as A can reach B in $\mathcal{PS}_{\mathcal{F}}$.

Cohen-Sellers suggested to estimate the parameter γ_k, where $\gamma_k = \min\{t : \; g_{k,t} = 2^k - 2\}$.

Cohen-Sellers, 1982

For every $k \geq 2$, it holds $\gamma_k \leq 2^k - 2$.

Joel E. Cohen, Peter H. Sellers, *Sets of nonnegative matrices with positive inhomogenous products*, Linear Algebra and its Applications 47 (1982) 185–192.

W., Yinfeng Zhu

For every $k \geq 2$, it holds $\gamma_k \leq k$.

Given a set $\mathcal{A}$ of Boolean linear maps on $[k]$, can we find a set $\mathcal{F}$ of Boolean linear maps on $[k]$ such that $g(\mathcal{F} \cup \mathcal{A}) = 2^k - 2$?

We let $\gamma(\mathcal{A})$ be the minimum size of such $\mathcal{F}$ if it exists and let $\gamma(\mathcal{A}) = \infty$ otherwise.

W., Yinfeng Zhu

Take $k \geq 2$. For a set $\mathcal{A}$ of Boolean linear maps on $[k]$, it holds $\gamma(\mathcal{A}) \leq 2^k - 2$ if and only if the only possible cycles in $\mathcal{PS}_{\mathcal{A}}$ are loops at $\emptyset$ and $[k]$.

Xinmao Wang, W., Ziqing Xiang

It holds $\gamma(W_{k,k-1}) \leq \binom{k-2}{\lfloor (k-2)/2 \rfloor}$ for all integers $k \geq 2$.

GRAPH-INDEXED MARKOV CHAINS

GRAPH-INDEXED MARKOV CHAIN

If we think of time not as infinite path (integers) but as a general graph, we go from usual Markov chains to graph-indexed Markov chains.

Take a connected graph G. A function f from $V(G)$ to integers is **Lipschitz** if

$$\min f = 0$$

and

$$0 \leq |f(u) - f(v)| \leq \mathrm{dist}_G(u, v)$$

for all $u, v \in V(G)$. The **range** of such a map f is $\max f$. The average range of all Lipschitz functions on G is called the **average range of G**.

GRAPH-INDEXED MARKOV CHAIN: CONTD.

Conjecture (Loebl-Nešetřil-Reed, 2003)

The maximum average range among all connected graphs with n vertices is achieved by the n-vertex path.

W., Zeying Xu, Yinfeng Zhu

Among all trees with the same number of vertices, the path is the unique tree whose average range takes the maximum value.

THANKS!

Zhangye, China

→ **ykwu@sjtu.edu.cn**

→ **http://math.sjtu.edu.cn/faculty/ykwu/home.php**

S.L. Kalpazidou,
»Cycle Representations of Markov Processes«
Second Edition, Springer, 2006.

B.P. Kitchens,
»Symbolic Dynamics: One-sided, Two-sided and Countable State Markov Shifts«
Springer, 1998.

D. Lind and B. Marcus,
»An Introduction to Symbolic Dynamics and Coding«
Cambridge University Press, 1995.

B. Sturmfels,
»Geometry of higher-order Markov chains«
Journal of Algebraic Statistics, 3, pp. 1–10, 2012.