Random walks pertaining to a class of deterministic weighted graphs

2009 J. Phys. A: Math. Theor. 42 275001

(http://iopscience.iop.org/1751-8121/42/27/275001)

J. Phys. A: Math. Theor. **42** (2009) 275001 (12pp) doi:10.1088/1751-8113/42/27/275001

Random walks pertaining to a class of deterministic weighted graphs

Thierry Huillet

Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, 95032, Cergy-Pontoise, France

E-mail: Thierry.Huillet@u-cergy.fr

Received 24 March 2009, in final form 20 May 2009
Published 11 June 2009
Online at stacks.iop.org/JPhysA/42/275001

Abstract
In this paper, we try to analyze and clarify the intriguing interplay between some counting problems related to specific thermalized weighted graphs and random walks consistent with such graphs.

PACS numbers: 02.50.Ey, 02.50.−r, 05.30.Pr, 05.40.−a, 02.10.Yn

1. Introduction

The purpose of this work is to underline the subtle relationship between some counting problems related to thermalized weighted graphs and random walks (RW) consistent with such graphs; see, e.g., [1] for a detailed treatise on general RWs on graphs. Let us summarize the topics developed in this paper. We start with defining finite thermalized weighted graphs. We show that the notion of a graph singularity spectrum naturally arises in the problem consisting in counting the number of paths whose transition-energy rate is asymptotically of a given order. This spectrum is classically the Legendre transform of the graph pressure function which is the logarithm of the spectral radius of its weight matrix. The corresponding Perron eigenvectors play a key role in the Perron–Frobenius theory. We next recall the Gibbs-variational principle, stating that the pressure produced by all RWs consistent with the graph structure is bounded below by the graph pressure which may itself be viewed as the pressure of some consistent canonical RW. This turns out to be a by-product of the Ruelle thermodynamic formalism. We then exhibit and interpret some important consistent RWs in the light of quasi-stationary distributions for substochastic RWs. The idea is to normalize the weight matrix of the graph by its norm to make it substochastic so that, by adding an extra absorbing coffin state, we may switch to the study of a proper RW conditioned to its absorption time. By doing so, a probabilistic interpretation of both the spectral radius and Perron eigenvectors of the graph weight matrix naturally comes out, and at least two conditionings are shown to be relevant: one is to condition locally the above RW on not hitting the absorbing state in one step at each iteration; the other is to condition it on not hitting the absorbing state in the remote future.

The latter construction is shown to be the canonical RW with smallest pressure production rate. At the end of the paper, we briefly discuss three particular cases, namely: the case where the weighted graph is reduced to its adjacency matrix, the case of a potential weighted graph and the case of a symmetric reversible weighted graph. Several general conclusions that make use of the above constructions may be drawn. One is an expression in terms of the average transition energy of the canonical RW associated with the adjacency matrix of the value at which the singularity spectrum of any weighted graph attains its maximum; another is that the entropy production rate of the locally conditioned RW is always bounded above by the logarithm of the spectral radius of its adjacency matrix, in the potential case.

2. Finite graphs with Boltzmann weights

Let $W \geqslant 0$ be some non-negative $N \times N$ weight matrix of some finite graph (i.e. with non-negative entries $W(i, j) \geqslant 0$). Let $A = [A(i, j)]$, defined by the indicator function,

$$A(i, j) = \mathbf{I}(W(i, j) > 0) \in \{0, 1\},$$

stand for the Boolean adjacency matrix associated with W. With A' denoting the transpose of A, we shall assume that $A = A'$ and that A is irreducible: in other words, the underlying topological graph is undirected and strongly connected so that for each couple of states (i, j), there is an integer m such that $A^m(i, j) > 0$. With $\beta \in \mathbb{R}$, we shall choose to represent W under the form

$$W(i, j) =: W_\beta(i, j) = A(i, j)\,\mathrm{e}^{-\beta H(i,j)}, \tag{1}$$

for some well-behaved transition energies $-\infty < H(i, j) < +\infty$ from state i to j, not all equal to the same value. The matrix, W_β, therefore appears to be the weight matrix of some thermalized weighted graph: it can be represented as the Hadamard product (say, *) of A with some positive Boltzmann kernel matrix K_β with entries $K_\beta(i, j) = \mathrm{e}^{-\beta H(i,j)}$:

$$W_\beta = A * K_\beta. \tag{2}$$

We note that the Hadamard λ-power ($\lambda > 0$) of W_β simply is $W_\beta^{*\lambda} = W_{\lambda\beta} = A * K_{\lambda\beta}$, corresponding to a rescaling of β.

Remark. Let $x_1 < x_2 < \cdots < x_N$ be N points on the line (circle). For some matrix $\mathcal{H}$, we may define

$$H(i, j) = \mathcal{H}(x_i, x_j)$$

to be the interaction energy between sites (i, j) in positions (x_i, x_j) leading to a slightly more general spatially extended model that can be treated along similar lines.

2.1. Some counting problems arising in this context

The quantity,

$$W_\beta(\mathbf{i}_n) := \prod_{m=1}^{n} W_\beta(i_{m-1}, i_m),$$

is the weight of the n-path $\mathbf{i}_n := \{i_0, i_1, \ldots, i_n\}$ which is non-null if and only if: $A(\mathbf{i}_n) := \prod_{m=1}^{n} A(i_{m-1}, i_m) \neq 0$. The total product weight of n-step paths connecting states (i_0, i_n)

therefore is given by the corresponding element of the transfer matrix:

$$W_\beta^n(i_0, i_n) = \sum_{i_1,\ldots,i_{n-1}} \prod_{m=1}^{n} W_\beta(i_{m-1}, i_m)$$
$$= \sum_{i_1,\ldots,i_{n-1}} \prod_{m=1}^{n} A(i_{m-1}, i_m)\, \mathrm{e}^{-\beta \sum_{m=1}^{n} H(i_{m-1},i_m)} = \sum_{p=1}^{N_n(i_0,i_n)} \mathrm{e}^{-\beta H_n(p)},$$

where $H_n(p)$ is the cumulative energy of the pth path connecting (i_0, i_n) and $N_n(i_0, i_n) := A^n(i_0, i_n)$ is the number of such n-paths. Summing over the endpoints of the n-paths, we obtain the full partition function of energy:

$$Z_n(\beta) := \sum_{i_0,i_n} W_\beta^n(i_0, i_n) = \sum_{i_0,\ldots,i_n} \prod_{m=1}^{n} W_\beta(i_{m-1}, i_m) = \sum_{p=1}^{N_n} \mathrm{e}^{-\beta H_n(p)}, \tag{3}$$

where $N_n := \sum_{i_0,i_n} N_n(i_0, i_n) = Z_n(0)$ is the total number of n-paths. Define

$$N_{n,\varepsilon}(\alpha) := \#\left\{ p \in [N_n] : \frac{1}{n} H_n(p) \in (\alpha - \varepsilon, \alpha + \varepsilon) \right\},$$

the number of n-paths whose transition-energy rate is asymptotically of order α. We expect

$$\lim_{\varepsilon\downarrow 0} \lim_{n\uparrow\infty} \frac{1}{n} \log N_{n,\varepsilon}(\alpha) = f(\alpha) \geqslant 0, \tag{4}$$

where $f(\alpha) = \inf_\beta(\alpha\beta - p(\beta)), \alpha \in [\alpha_-, \alpha_+]$, is the concave Legendre transform of some concave pressure function $p(\beta)$. We may call $f(\alpha)$ the *singularity spectrum of the weighted graph.*

Observing from (3) that $Z_n(\beta) = \left\| W_\beta^n \right\|$ is a matrix norm and recalling under our irreducibility assumption $\left\| W_\beta^n \right\|^{1/n} \to_{n\uparrow\infty} \rho_\beta$, the spectral radius of W_β, we get

$$-\frac{1}{n} \log Z_n(\beta) \underset{n\uparrow\infty}{\to} p(\beta) = -\log \rho_\beta. \tag{5}$$

For each β, the number ρ_β, as an eigenvalue of W_β, satisfies $\det\left(\frac{1}{\rho_\beta} W_\beta - I\right) = 0$.

For each n, the quantity,

$$\Phi_n(\beta) := \frac{Z_n(\beta)}{Z_n(0)} = \frac{1}{N_n} \sum_{p=1}^{N_n} \mathrm{e}^{-\beta H_n(p)},$$

is the Laplace–Stieltjes transform of some discrete probability measure on n-paths satisfying $-\frac{1}{n} \log \Phi_n(\beta) \to p(\beta) - p(0) = -\log \frac{\rho_\beta}{\rho_0}$. This limit therefore is the log-Laplace transform of some probability distribution which, in particular, is smooth and concave. The pressure $p(\beta)$ is classically related to the scaled free energy $\tau(\beta)$ by $p(\beta) =: \beta\tau(\beta)$. Note that, for each n, we also have

$$\Phi_n(\beta) = \frac{1}{N_n} \sum_{h_n \in \mathcal{H}_n} N_n(h_n)\, \mathrm{e}^{-\beta h_n},$$

where $\mathcal{H}_n := \mathrm{Span}(H_n(p); p \in [N_n])$, $N_n(h_n) := \#\{p \in [N_n] : H_n(p) = h_n\}$ and $N_n(h_n)/N_n$ is the probability of n-paths of energy h_n.

With $\mathbf{1} = (1, 1, \ldots, 1)'$, we define $\mathbf{w}_\beta := W_\beta \mathbf{1}$ to be the column-sum vector of W_β, with entries $w_\beta(i) = \sum_j A(i, j)\, \mathrm{e}^{-\beta H(i,j)}$. Define $w_\beta^+ = \max_i w_\beta(i)$, $w_\beta^- = \min_i w_\beta(i)$. We have $w_\beta^- \leqslant \rho_\beta \leqslant w_\beta^+$ and so

$$-\log w_\beta^+ \leqslant p(\beta) \leqslant -\log w_\beta^-.$$

Clearly, it holds that

$-\log w_\beta^+ \underset{\beta\uparrow\infty}{\sim} \beta\alpha_-$ where $\alpha_- = \min_{i,j:A(i,j)=1} H(i,j)$ and similarly,

$-\log w_\beta^- \underset{\beta\downarrow-\infty}{\sim} \beta\alpha_+$ where $\alpha_+ = \max_{i,j:A(i,j)=1} H(i,j) > \alpha_-$.

We can check that $f(\alpha_-) = f(\alpha_+) = 0$ and that the maximum of $f(\alpha)$ is attained at $\alpha = \alpha_0 = p'(0)$. We also have $f(\alpha) = f(p'(\beta)) = p'(\beta)\beta - p(\beta)$ so that $f(\alpha_0) = -p(0) = \log\rho_0 > 0$, where ρ_0 is the spectral radius of A.

Whenever $\alpha_- < 0, \alpha_+ > 0$, then there is a β_c, possibly not equal to 0, given by $p'(\beta_c) = 0$. With $\alpha_c := p'(\beta_c) = 0$, $f(\alpha_c) = -p(\beta_c) > 0$.

For all distinct pairs of nodes (i,j), let $m(i,j) = \inf(m > 1 : A^m(i,j) > 0)$. Then

$$m_* = \max_{(i,j)} m(i,j)$$

is the diameter of the adjacency graph. For each (i,j), there can be more than one path of minimal length $m(i,j)$. Let $N_{m(i,j)} \geqslant 1$ be the number of such length$-m(i,j)$ paths and let $h(i,j)$ be the energy of any path with smallest energy among these $N_{m(i,j)}$ paths. Then

$$\alpha_* = \max_{(i,j)}[h(i,j)/m(i,j)]$$

is a quantity of interest related to the energy diameter of the weighted graph. Clearly, $\alpha_- < \alpha_* < \alpha_+$ and α_* belongs to the range of the spectrum.

2.2. *Perron–Frobenius and the like*

Let $\pi'_\beta > 0$ and $\varphi_\beta > 0$ be the line and column (l_1- norm 1) Perron vectors of W_β associated with the spectral radius ρ_β of W_β:

$$\rho_\beta\pi'_\beta = \pi'_\beta W_\beta \qquad \text{and} \qquad \rho_\beta\varphi_\beta = W_\beta\varphi_\beta. \tag{6}$$

Under our hypothesis, $\rho_\beta > 0$ is the algebraically simple real dominant eigenvalue of W_β. If A is, in addition, primitive ($A^m > 0$ for some integer m), then all the other eigenvalues of W_β are strictly contained within the disc: $|\rho| < \rho_\beta$, else some could lie on the disc, $|\rho| = \rho_\beta$, because of the underlying periodicity of the problem.

We shall let $\phi_\beta := \varphi_\beta/(\pi'_\beta\varphi_\beta)$ in such a way that the Hadamard product of π_β and ϕ_β, namely the column vector $\pi_\beta * \phi_\beta$, with components $\pi_\beta * \phi_\beta(i) = \pi_\beta(i)\phi_\beta(i)$, has l_1-norm 1 (i.e. $\pi'_\beta\phi_\beta = 1$).

Remarks.

(i) When $H(i,j) = H(j,i)$, W_β is itself symmetric ($W_\beta = W'_\beta$) then $\pi_\beta = \varphi_\beta$ and $\pi_\beta * \phi_\beta = \varphi_\beta * \varphi_\beta/(\varphi'_\beta\varphi_\beta)$. Then it is useful to introduce the *probability wave* vector of l_2-norm $1 : \psi_\beta = \varphi_\beta/(\varphi'_\beta\varphi_\beta)^{1/2}$ in such a way that $\pi_\beta * \phi_\beta = \psi_\beta * \psi_\beta$.

(ii) Letting $\varepsilon_\beta := 1 - \rho_\beta/w_\beta^+ \geqslant 0$ stand for the scaled spectral gap of the graph, the equation giving the right eigenvector φ_β may be recast as

$$\varepsilon_\beta\varphi_\beta = \left(I - \frac{1}{w_\beta^+}W_\beta\right)\varphi_\beta =: -\Delta_\beta\varphi_\beta,$$

where $\Delta_\beta = \frac{1}{w_\beta^+}W_\beta - I$ is a Laplacian of the graph. Observe that $w_\beta^+ = |||W_\beta|||_\infty$ is the matrix norm induced by the l_∞-vector norm so that $\rho_\beta/w_\beta^+ \leqslant 1$ (i.e. $\varepsilon_\beta \geqslant 0$) and that $\log\rho_\beta/\log w_\beta^+ \to_{\beta\uparrow\infty} 1$.

Consistently with (6), we shall let $\pi_0' > 0$ and $\varphi_0 > 0$ stand for the line and column Perron vectors of $A = W_0$ with

$$\rho_0 \pi_0' = \pi_0' A \text{and} \rho_0 \varphi_0 = A\varphi_0,$$

associated with the spectral radius ρ_0 of $A = W_0$. We shall let $\phi_0 := \varphi_0/(\pi_0'\varphi_0)$ so that the Hadamard product $\pi_0 * \phi_0$ has l_1- norm 1. Note that, since $A = A', \pi_0 = \varphi_0$ and $\pi_0 * \phi_0 = \varphi_0 * \varphi_0/(\varphi_0'\varphi_0) =: \psi_0 * \psi_0$ where $\psi_0 = \varphi_0/(\varphi_0'\varphi_0)^{1/2}$ is the wave vector associated with A.

3. Random walks on graphs

In the study of weighted graphs, questions pertaining to counting are then relevant. Once such weighted graphs have been introduced, it is useful to consider the following particular class of random walks attached to such graphs.

Let $0 \leqslant \Pi = [\Pi(i,j)]$ denote some stochastic matrix with column sums one, $\Pi\mathbf{1} = \mathbf{1}$. Let $\mathcal{P}_{W_\beta}$ be the set of stochastic matrices which are W_β-consistent in the sense that

$$\Pi \in \mathcal{P}_{W_\beta} \leftrightarrow \{\Pi(i,j) = 0 \text{ whenever } W_\beta(i,j) = 0\}. \tag{7}$$

Let $\boldsymbol{\mu}' > 0$ be the line left Perron eigenvector of Π, satisfying: $\boldsymbol{\mu}' = \boldsymbol{\mu}'\Pi$ (the unique invariant probability measure associated with Π). Clearly, to each such Π a positive recurrent random walk can be associated.

3.1. A variational principle and first consequences

In this context, the following Gibbs-variational principle indeed holds [2], resulting from the Ruelle thermodynamic formalism [10]. It reads

$$\log \rho_\beta = \sup_{\Pi \in \mathcal{P}_{W_\beta}} \left(-\sum_i \mu(i) \sum_j \Pi(i,j) \log \Pi(i,j) + \sum_i \mu(i) \sum_j \Pi(i,j) \log W_\beta(i,j) \right)$$

$$= \sup_{\Pi \in \mathcal{P}_{W_\beta}} \left(-\sum_i \mu(i) \sum_j \Pi(i,j) \log \Pi(i,j) - \beta \sum_i \mu(i) \sum_j \Pi(i,j) H(i,j) \right), \tag{8}$$

where the supremum is attained for the unique stochastic matrix Π_* which is W_β-consistent and defined by the Doob transform:

$$\Pi_*(i,j) = \frac{1}{\rho_\beta} W_\beta(i,j) \frac{\phi_\beta(j)}{\phi_\beta(i)}. \tag{9}$$

With $D_{\phi_\beta} := \mathrm{diag}(\phi_\beta)$, this is also $\Pi_* = \frac{1}{\rho_\beta} D_{\phi_\beta}^{-1} W_\beta D_{\phi_\beta}$, in the matrix form. The corresponding invariant measure satisfying $\boldsymbol{\mu}_*' = \boldsymbol{\mu}_*'\Pi_*$ can easily be checked to be

$$\boldsymbol{\mu}_* = \boldsymbol{\pi}_\beta * \boldsymbol{\phi}_\beta, \tag{10}$$

the Hadamard product of the left and right eigenvectors of the weight matrix W_β. We shall call the RW with transition probability Π_* the *canonical* RW consistent with W_β. Using this canonical RW construction, we get

$$\log \rho_\beta = -\frac{1}{\rho_\beta} \sum_i \pi_\beta(i)\phi_\beta(i) \sum_j W_\beta(i,j) \frac{\phi_\beta(j)}{\phi_\beta(i)} \log\left(\frac{1}{\rho_\beta}\frac{\phi_\beta(j)}{\phi_\beta(i)}\right)$$

$$= -\frac{1}{\rho_\beta}\sum_i \pi_\beta(i)\sum_j W_\beta(i,j)\phi_\beta(j)[-\log\rho_\beta + \log\phi_\beta(j) - \log\phi_\beta(i)]$$

$$= \log\rho_\beta + \sum_i \pi_\beta(i)\phi_\beta(i)\log\phi_\beta(i) - \frac{1}{\rho_\beta}\sum_i \pi_\beta(i)\sum_j W_\beta(i,j)\phi_\beta(j)\log\phi_\beta(j),$$

leading to the expression

$$\rho_\beta = \frac{\sum_i \pi_\beta(i)\sum_j W_\beta(i,j)\phi_\beta(j)\log\phi_\beta(j)}{\sum_i \pi_\beta(i)\phi_\beta(i)\log\phi_\beta(i)},$$

in terms of the left and right Perron eigenvectors of W_β. As a result, we obtain

$$p(\beta) =: \beta\tau(\beta) = -\log\left[\frac{\sum_i \pi_\beta(i)\sum_j A(i,j)\,\mathrm{e}^{-\beta H(i,j)}\phi_\beta(j)\log\phi_\beta(j)}{\sum_i \pi_\beta(i)\phi_\beta(i)\log\phi_\beta(i)}\right]. \tag{11}$$

From (8), for all W_β-consistent stochastic matrix $\Pi \neq \Pi_*$:

$$\log\rho_\beta > -\sum_i \mu(i)\sum_j \Pi(i,j)\log\Pi(i,j) - \beta\sum_i \mu(i)\sum_j \Pi(i,j)H(i,j). \tag{12}$$

In the right-hand side of (12), $s := -\sum_i \mu(i)\sum_j \Pi(i,j)\log\Pi(i,j) =: \sum_i \mu(i)s(i)$ is the equilibrium Shannon entropy production rate of the ergodic Markov chain governed by Π and $u := \sum_i \mu(i)\sum_j \Pi(i,j)H(i,j) =: \sum_i \mu(i)u(i)$ its equilibrium internal transition energy. It follows from (8) that the quantity, $p(\beta) := -\log\rho_\beta$, is a universal lower bound for the equilibrium pressure production rate of all W_β-consistent walkers. Stated differently, defining the pressure of a consistent RW governed by $\Pi \in \mathcal{P}_{W_\beta}$, $\Pi \neq \Pi_*$, as

$$p_\Pi(\beta) := \beta\sum_i \mu(i)\sum_j \Pi(i,j)H(i,j) + \sum_i \mu(i)\sum_j \Pi(i,j)\log\Pi(i,j),$$

it holds that

$$\text{for all } \Pi \neq \Pi_* \in \mathcal{P}_{W_\beta}, \qquad p_\Pi(\beta) > p(\beta) = p_{\Pi_*}(\beta). \tag{13}$$

Remark. On the other hand, we also recall the Friedland–Karlin inequality [4] of a similar flavor:

$$p(\beta) = -\log\rho_\beta \geqslant \sum_i \pi_\beta(i)\phi_\beta(i)\log w_\beta(i), \tag{14}$$

where, $\mathbf{w}_\beta := W_\beta\mathbf{1}$ is the column-sum vector of W_β: $w_\beta(i) = \sum_j A(i,j)\,\mathrm{e}^{-\beta H(i,j)}$. It gives a universal lower bound of $p(\beta)$ in terms of the invariant measure $\mu_*(i) = \pi_\beta(i)\phi_\beta(i)$ associated with Π_*.

3.2. The entropy production rate of the RW governed by $\Pi \in \mathcal{P}_{W_\beta}$

We need to say a few words on the way to compute the quantity s associated with some Π. We refer to [8] for additional information. For each pair of connecting states (i_0, i_n), define

$$\Pi_\lambda^n(i_0, i_n) := \sum_{i_1,\ldots,i_{n-1}}\prod_{m=1}^n \Pi(i_{m-1}, i_m)^\lambda,$$

where $\Pi(i_{m-1}, i_m)^\lambda$ is the (i_{m-1}, i_m) entry of $\Pi^{*\lambda}$, the Hadamard λ-power of Π, $\lambda > 0$. Define the Rényi λ-entropy of all n-paths of the RW governed by Π and started using the invariant measure $\boldsymbol{\mu}$ to be

$$R_n(\lambda) := \frac{1}{1-\lambda}\log\sum_{i_0,i_n}\mu(i_0)\Pi_\lambda^n(i_0, i_n).$$

Note that with $\Pi(\mathbf{i}_n) := \prod_{m=1}^{n} \Pi(i_{m-1}, i_m)$ the probability of the n-path, $\mathbf{i}_n := \{i_0, i_1, \ldots, i_n\}$,

$$R_n(\lambda) \to {}_{\lambda\uparrow 1} S_n = -\sum_{\mathbf{i}_n} \mu(i_0)\Pi(\mathbf{i}_n)\log\Pi(\mathbf{i}_n),$$

the Shannon entropy of n-paths at equilibrium. Then, with $\alpha(\lambda) := \sum_i \mu(i) \sum_j \Pi(i,j)^\lambda$,

$$\frac{1}{n} R_n(\lambda) \underset{n\uparrow\infty}{\to} r(\lambda) := \frac{1}{1-\lambda}\log\alpha(\lambda), \tag{15}$$

where $r(\lambda)$ is the Rényi-entropy production rate of the walker. As a result,

$$r(\lambda) \to {}_{\lambda\uparrow 1} s = -\sum_i \mu(i) \sum_j \Pi(i,j)\log\Pi(i,j) = -\alpha'(1).$$

This approach is useful to compute the Shannon-entropy production rate s for specific Πs.

3.3. Random walks consistent with W_β

We now exhibit and interpret some important W_β-consistent RWs in the light of *quasi-stationary distributions* for substochastic RWs. By doing so, a probabilistic interpretation of ρ_β, $\boldsymbol{\pi}_\beta$ and $\boldsymbol{\phi}_\beta$ emerges.

Let us first normalize W_β in the following way. Consider the matrix

$$\overline{W}_\beta := \frac{W_\beta}{\|W_\beta\|}, \tag{16}$$

for some matrix-norm $\|W_\beta\|$ of W_β. For example, $\|W_\beta\| = N \max_{i,j} W_\beta(i,j)$ or $\|W_\beta\| = \sum_{i,j} W_\beta(i,j)$ or $\|W_\beta\| = w_\beta^+ = \max_i w_\beta(i)$.

The spectral radius of $\overline{W}_\beta$ now is $\overline{\rho}_\beta = \rho_\beta/\|W_\beta\| < 1$ with the same left and right strictly positive Perron eigenvectors $\boldsymbol{\pi}_\beta > 0$ and $\boldsymbol{\varphi}_\beta > 0$ as for W_β in (6). By doing so, the matrix $\overline{W}_\beta$ is *substochastic* in the sense that, with $\overline{\mathbf{w}}_\beta := \overline{W}_\beta \mathbf{1}$ being the column-sum vector of $\overline{W}_\beta$, then $\overline{w}_\beta(i) \in (0,1]$ with $\overline{w}_\beta(i) < 1$ for at least one i. To recast this problem into a stochastic problem, we may add an additional *coffin state*, say $\partial := \{0\}$ and look at the enlarged $(N+1)\times(N+1)$ stochastic matrix P:

$$P = \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{1}-\overline{\mathbf{w}}_\beta & \overline{W}_\beta \end{bmatrix}. \tag{17}$$

P now is the stochastic transition matrix of a RW, say $\{X_n\}$, having state $\{0\}$ as an additional *absorbing* state. Let τ_0 be the first hitting time of $\partial = \{0\}$ for this RW $\{X_n\}$.

Using this construction, clearly, the substochastic matrix $\overline{W}_\beta$ turns out to be the transition matrix of the process $X_n \cdot \mathbf{I}(\tau_0 > n)$ (i.e. X_n, restricted to the set $\tau_0 > n$). In other words, with $\mathbf{e}'_{i_0}$ the line vector with a single 1 in position i_0, 0 elsewhere, we have

$$\mathbb{P}_{i_0}(X_n = i_n, \tau_0 > n) = \overline{W}^n_\beta(i_0, i_n) = \mathbf{e}'_{i_0}\overline{W}^n_\beta \mathbf{e}_{i_n}, i_0, i_n \in \{1, \ldots, N\}.$$

Therefore,

$$\mathbb{P}_{i_0}(\tau_0 > n) = \mathbf{e}'_{i_0}\overline{W}^n_\beta \mathbf{1}. \tag{18}$$

We note that $\mathbb{P}_{i_0}(\tau_0 = 1) = \mathbb{P}_{i_0}(\tau_0 > 0) - \mathbb{P}_{i_0}(\tau_0 > 1) = \mathbf{e}'_{i_0}(I - \overline{W}_\beta)\mathbf{1} = 1 - \overline{w}_\beta(i_0)$, the probability mass defect of $\overline{W}_\beta$ at state i_0.

For all (i_0, i_n), we have $\lim_{n\uparrow\infty}\left[\overline{W}^n_\beta(i_0,i_n)\right]^{1/n} = \overline{\rho}_\beta$ and, only when A is primitive (irreducible and aperiodic), by the strong version of the Perron–Frobenius theorem (see [6]):

$$\lim_{n\uparrow\infty} \overline{\rho}_\beta^{-n}\, \overline{W}^n_\beta = \boldsymbol{\phi}_\beta \boldsymbol{\pi}'_\beta, \tag{19}$$

where $\boldsymbol{\pi}'_\beta > 0$ and $\phi_\beta > 0$, defined in (6), are the left (right) eigenvectors of $\overline{W}_\beta$ associated with $\overline{\rho}_\beta$, chosen, as before, so as to satisfy $\boldsymbol{\pi}'_\beta \phi_\beta = 1$. As a result of (18) and (19), when A is primitive,

$$\lim_{n\uparrow\infty} \overline{\rho}_\beta^{-n} \mathbb{P}_{i_0}(\tau_0 > n) = \phi_\beta(i_0), \tag{20}$$

meaning that τ_0 is *tail equivalent to a geometric random variable with success probability* $\overline{\rho}_\beta$. The latter formula, therefore, gives the limiting interpretation of ϕ_β in the context of the RW $\{X_n\}$. What about $\boldsymbol{\pi}_\beta$?

First, because $\boldsymbol{\pi}'_\beta$ is the left eigenprobability vector of $\overline{W}_\beta$ with the eigenvalue $\overline{\rho}_\beta$

$$\mathbb{P}_{\pi_\beta}(\tau_0 > n) := \sum_{i_0=1}^{N} \pi_\beta(i_0)\mathbb{P}_{i_0}(\tau_0 > n) = \boldsymbol{\pi}'_\beta \overline{W}_\beta^n \mathbf{1} = \overline{\rho}_\beta^n. \tag{21}$$

If the process is started with $\boldsymbol{\pi}_\beta$, the law of τ_0 is *exactly* geometrically distributed on $\{1, 2, \ldots, N\}$ with success probability $\overline{\rho}_\beta$.

Consider now the conditional probability $\mathbb{P}_{i_0}(X_n = i_n \mid \tau_0 > n)$.

Recalling $\mathbb{P}_{i_0}(X_n = i_n, \tau_0 > n) = \mathbf{e}'_{i_0} \overline{W}_\beta^n \mathbf{e}_{i_n}$, by the Bayes rule, we get

$$\mathbb{P}_{i_0}(X_n = i_n \mid \tau_0 > n) = \frac{\mathbf{e}'_{i_0} \overline{W}_\beta^n \mathbf{e}_{i_n}}{\mathbf{e}'_{i_0} \overline{W}_\beta^n \mathbf{1}} = \frac{\mathbf{e}'_{i_0} \left(\overline{\rho}_\beta^{-n} \overline{W}_\beta^n\right) \mathbf{e}_{i_n}}{\mathbf{e}'_{i_0} \left(\overline{\rho}_\beta^{-n} \overline{W}_\beta^n\right) \mathbf{1}},$$

showing that, independently of the starting point i_0,

$$\mathbb{P}_{i_0}(X_n = i_n \mid \tau_0 > n) \underset{n\uparrow\infty}{\to} \frac{\mathbf{e}'_{i_0} (\phi_\beta \boldsymbol{\pi}'_\beta) \mathbf{e}_{i_n}}{\mathbf{e}'_{i_0} (\phi_\beta \boldsymbol{\pi}'_\beta) \mathbf{1}} = \pi_\beta(i_n).$$

Such a probability measure $\boldsymbol{\pi}_\beta$ is called a Yaglom limit [11] of $\{X_n\}$.

Further, with $\mathbb{P}_{\pi_\beta}(\cdot) := \sum_{i=1}^N \pi_\beta(i_0)\mathbb{P}_{i_0}(\cdot)$, for each $n, i_n \in \{1, 2, \ldots, N\}$:

$$\mathbb{P}_{\pi_\beta}(X_n = i_n \mid \tau_0 > n) := \frac{\mathbb{P}_{\pi_\beta}(X_n = i_n, \tau_0 > n)}{\mathbb{P}_{\pi_\beta}(\tau_0 > n)} = \frac{\pi'_\beta \overline{W}_\beta^n \mathbf{e}_{i_n}}{\pi'_\beta \overline{W}_\beta^n \mathbf{1}} = \pi_\beta(i_n), \tag{22}$$

and this precisely means that $\boldsymbol{\pi}_\beta$ is the (unique) quasi-stationary distribution (QSD) of $\{X_n\}$. As is well known for Markov chains with finite state space absorbed at ∂, we observe that the Yaglom limit coincides with its QSD. *When A is primitive (strongly connected and aperiodic), equations (21), (22) and (20) provide a natural interpretation of ρ_β, $\boldsymbol{\pi}_\beta$ and ϕ_β in terms of the RW governed by P in (17) and its stopping time τ_0.*

We refer to [7] for additional informations on QSD and Yaglom limits in the context of population dynamics.

Remark. When A is irreducible but not primitive (the underlying topological graph is strongly connected but periodic), only the following weaker form of the Perron–Frobenius theorem holds true [6]:

$$\lim_{K\uparrow\infty} \frac{1}{K} \sum_{n=1}^{K} \overline{\rho}_\beta^{-n} \overline{W}_\beta^n = \phi_\beta \boldsymbol{\pi}'_\beta.$$

With $\phi_\beta^+ = \max_i \phi_\beta(i)$, $\phi_\beta^- = \min_i \phi_\beta(i)$, equation (20) has to be weakened into

$$\phi_\beta^- / \phi_\beta^+ \leqslant \overline{\rho}_\beta^{-n}\ \mathbb{P}_{i_0}(\tau_0 > n) \leqslant \phi_\beta^+ / \phi_\beta^-, \qquad \text{for all} \quad n, i_0.$$

The quantity, $\overline{\rho}_\beta^{-n} \mathbb{P}_{i_0}(\tau_0 > n)$, may oscillate and not tend to some limit; however, $-\frac{1}{n} \log \mathbb{P}_{i_0}(\tau_0 > n) \to \overline{\rho}_\beta$ still holds true.

The above construction of the RW $\{X_n\}$ allows us now to interpret two fundamental W_β-consistent RWs.

The locally conditioned random walk. With $D_{\overline{\mathbf{w}}_\beta} := \mathrm{diag}(\overline{\mathbf{w}}_\beta)$, the transition matrix of the one-step conditioned process, $(X_1 \mid \tau_0 > 1)$, is

$$\Pi(i, j) = \frac{\mathbf{e}_i' \overline{W}_\beta \, \mathbf{e}_j}{\mathbf{e}_i' \overline{W}_\beta \mathbf{1}} = \mathbf{e}_i' \big[D_{\overline{\mathbf{w}}_\beta}^{-1} \overline{W}_\beta \big] \mathbf{e}_j = \frac{\mathbf{e}_i' W_\beta \, \mathbf{e}_j}{\mathbf{e}_i' W_\beta \mathbf{1}} = \mathbf{e}_i' \big[D_{\mathbf{w}_\beta}^{-1} W_\beta \big] \mathbf{e}_j,$$

normalizing each line i by $\mathbf{e}_i' \overline{W}_\beta \mathbf{1} = \overline{w}_\beta(i)$. Clearly, $\Pi \mathbf{1} = \mathbf{1}$ and the RW with transition matrix,

$$\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta, \tag{23}$$

is W_β-consistent. Note that Π is *invariant* under the scaling $W_\beta \to \overline{W}_\beta$.

Let $\boldsymbol{\mu}$ be the invariant associated with this Π. It holds, [9], that

$$\mu(i) = \frac{(I - \Pi)_{i,i}}{\sum_i (I - \Pi)_{i,i}},$$

where $(I - \Pi)_{i,i}$ is the cofactor of the (i, i)-entry of the matrix $I - \Pi$. Then, if $\overleftarrow{\Pi}$ is the transition matrix of the reversed (backward in time) chain of Π at equilibrium, $\overleftarrow{\Pi}' = D_\mu \Pi D_\mu^{-1}$. In general, $\overleftarrow{\Pi} \neq \Pi$ and detailed balance may not hold.

Global conditioning and the canonical process. Consider now the proper Markov chain whose transition probabilities are obtained by the Doob transform,

$$\Pi_*(i, j) = \overline{\rho}_\beta^{-1} \frac{\phi_\beta(j)}{\phi_\beta(i)} \overline{W}_\beta(i, j) = \rho_\beta^{-1} \frac{\phi_\beta(j)}{\phi_\beta(i)} W_\beta(i, j), \quad i, j \in \{1, \dots, N\},$$

satisfying $\Pi_* \mathbf{1} = \mathbf{1}$. In the matrix form,

$$\Pi_* = \rho_\beta^{-1} D_{\phi_\beta}^{-1} W_\beta D_{\phi_\beta}, \tag{24}$$

and Π_* is also invariant under the scaling $W_\beta \to \overline{W}_\beta$. An important property of this RW is the following: the probability, $\Pi_*(\mathbf{i}_n) := \prod_{m=1}^n \Pi_*(i_{m-1}, i_m)$, of the n-path $\mathbf{i}_n$ is

$$\Pi_*(\mathbf{i}_n) = \rho_\beta^{-n} W_\beta(\mathbf{i}_n) \prod_{m=1}^n \frac{\phi_\beta(i_m)}{\phi_\beta(i_{m-1})} = \rho_\beta^{-n} W_\beta(\mathbf{i}_n) \frac{\phi_\beta(i_n)}{\phi_\beta(i_0)}.$$

For a bridge n-path for which $i_0 = i_n$, $\Pi_*(\mathbf{i}_n) = \rho_\beta^{-n} W_\beta(\mathbf{i}_n)$ reduces, up to a scaling constant, to the weight $W_\beta(\mathbf{i}_n)$ of the n-path $\mathbf{i}_n$.

The invariant probability distribution $\boldsymbol{\mu}_*$ on $\{1, \dots, N\}$ satisfying $\boldsymbol{\mu}_*' \Pi_* = \boldsymbol{\mu}_*'$ exists. It is given explicitly by $\boldsymbol{\mu}_* = \boldsymbol{\pi}_\beta * \phi_\beta$ and so

$$\mu_*(i) = \pi_\beta(i) \phi_\beta(i), \qquad i = 1, \dots, N. \tag{25}$$

Doob transforms have to do with conditioning a process on its lifetime. The Markov chain with one-step transition probability matrix Π_* may be shown to be that of the process whose one-step transition probability distribution is

$$\Pi_*(i, j) = \lim_{n \uparrow \infty} \mathbb{P}_i(X_1 = j \mid \tau_0 > n), \tag{26}$$

corresponding to X_n *conditioned* to never hit the coffin state $\partial = \{0\}$ in the distant future; see [7]. This process has a unique invariant measure given by $\boldsymbol{\mu}_*$ in (25).

Defining as before $\overleftarrow{\Pi}_*$ by $\overleftarrow{\Pi}'_* = D_{\mu_*}\Pi_* D_{\mu_*}^{-1}$, we can ask conditions under which detailed balance $\overleftarrow{\Pi}_* = \Pi_*$ holds. We have

$$\overleftarrow{\Pi}'_* = \rho_\beta^{-1} D_{\pi_\beta} D_{\phi_\beta} D_{\phi_\beta}^{-1} W_\beta D_{\phi_\beta} D_{\phi_\beta}^{-1} D_{\pi_\beta}^{-1} = \rho_\beta^{-1} D_{\pi_\beta} W_\beta D_{\pi_\beta}^{-1}$$

so that $\overleftarrow{\Pi}_* = \rho_\beta^{-1} D_{\pi_\beta}^{-1} W'_\beta D_{\pi_\beta}$ showing that reversibility holds when $W_\beta = W'_\beta$ since if this is the case: $\boldsymbol{\pi}_\beta = \boldsymbol{\phi}_\beta$.

No extra state. We emphasize here that there are some alternative ways to force the substochastic problem into a stochastic one. Assume $A(i,i) > 0$ for each i, in which case $A^{N-1} > 0$ and A necessarily is primitive (see lemma 8.5.5 of [6]). Consider the stochastic matrix Π which is W_β-consistent:

$$\Pi = \overline{W}_\beta + D_{\mathbf{1}-\overline{\mathbf{w}}_\beta}, \tag{27}$$

where $D_{\mathbf{1}-\overline{\mathbf{w}}_\beta} := \mathrm{diag}(\mathbf{1}-\overline{\mathbf{w}}_\beta)$, satisfying $\Pi\mathbf{1} = \mathbf{1}$. In that case, the mass defect vector $\mathbf{1}-\overline{\mathbf{w}}_\beta$ is transferred to the diagonal entries of $\overline{W}_\beta$ to make it stochastic, without appealing to an extra coffin state. Note that Π in (27) *no* longer is *invariant* under the scaling $W_\beta \to \overline{W}_\beta$ and so this normalization is norm dependent.

4. Special cases

4.1. The topological case

Assume $\beta = 0$. In this case, $W_0 = A$ and

$$\log\rho_0 > s = -\sum_i \mu(i) \sum_j \Pi(i,j)\log\Pi(i,j)$$

for all A-consistent matrix $\Pi \neq \Pi_*^0$ with $\Pi_*^0(i,j) = \frac{1}{\rho_0}A(i,j)\frac{\phi_0(j)}{\phi_0(i)}$. $\log\rho_0 > 0$ interprets as the maximal entropy production rate of all Markov chains governed by such Πs. The RW with transition matrix Π_*^0 is termed the maximal entropy RW in [3]. Its invariant measure is $\mu_*(i) = \psi_0(i)^2$. When $\Pi = D_{\mathbf{a}}^{-1}A$, with $\mathbf{a} = A\mathbf{1}$, the invariant measure is $\mu(i) = a(i)\big/\sum_i a(i)$, proportional to the node degrees. Then $s = \sum_i a(i)\log a(i)\big/\sum_i a(i)$ and $\log\rho_0 > s$ is an inequality first discussed in [3]. When disorder is present, the canonical RW associated with $W_0 = A$ was also shown therein to exhibit *localization* properties.

Consider the general inequality (12) for all W_β-consistent stochastic matrix $\Pi \neq \Pi_* =: \frac{1}{\rho_\beta}W_\beta(i,j)\frac{\phi_\beta(j)}{\phi_\beta(i)}$. Choosing for Π the above particular value, $\Pi = \Pi_*^0$, we obtain

$$\log\rho_\beta > \log\rho_0 - \beta\sum_i \psi_0(i)^2 \sum_j \Pi_*^0(i,j)H(i,j).$$

Therefore, the average transition energy,

$$\alpha_0 := \sum_i \psi_0(i)^2 \sum_j \Pi_*^0(i,j)H(i,j), \tag{28}$$

interprets as the slope at $\beta = 0$ of the graph pressure function $\beta \to p(\beta)$, namely, $\alpha_0 = p'(0)$. We have $f(\alpha_0) = -p(0) = \log\rho_0$.

4.2. The potential case

Assume $H(i, j) = U(j) - U(i)$ for some potential U attached to the nodes of the graph. In this case, the matrix K_β defining W_β is called a potential kernel. First, in this case, it follows from (28) and the equilibrium property of $(\mu_* = \psi_0 * \psi_0, \Pi_*^0)$ that

$$\alpha_0 := \sum_i \psi_0(i)^2 \sum_j \Pi_*^0(i, j)[U(j) - U(i)] = 0. \tag{29}$$

We conclude that *the singularity spectrum of all graph with potential kernel K_β attains its maximum at* $\alpha_0 = 0$.

With $\mathbf{v}_\beta$ being the column vector with entries $v_\beta(i) = \exp -\beta U(i)$, we get

$$W_\beta := A * K_\beta = D_{\mathbf{v}_\beta}^{-1} A D_{\mathbf{v}_\beta}. \tag{30}$$

We have $\mathbf{w}_\beta = W_\beta \mathbf{1} = D_{\mathbf{v}_\beta}^{-1} A \mathbf{v}_\beta$ and $D_{\mathbf{w}_\beta} = D_{\mathbf{v}_\beta}^{-1} D_{A\mathbf{v}_\beta}$. Note that W_β is diagonally similar to A so that the spectral radius of W_β is ρ_0, independently of β.

–Consider first the RW with transition matrix $\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta = D_{A\mathbf{v}_\beta}^{-1} A D_{\mathbf{v}_\beta}$. Its invariant measure is characterized by $\boldsymbol{\mu}' = \mu'\Pi$. Recalling $A = A'$, we find $\boldsymbol{\mu} \propto D_{\mathbf{v}_\beta} A \mathbf{v}_\beta$, with normalized entries weighting output degree nodes with lowest U:

$$\mu(i) = \sum_j A(i, j)\, \mathrm{e}^{-\beta(U_i + U_j)} \Big/ \sum_{i,j} A(i, j)\, \mathrm{e}^{-\beta(U_i + U_j)}. \tag{31}$$

This RW with transition matrix Π is reversible because $\overleftarrow{\Pi} = D_{\boldsymbol{\mu}}^{-1} \Pi' D_{\boldsymbol{\mu}} = \Pi$.
–Second, consider the canonical RW consistent with $W_\beta = D_{\mathbf{v}_\beta}^{-1} A D_{\mathbf{v}_\beta}$. The right eigenvector ϕ_β of $W_\beta = D_{\mathbf{v}_\beta}^{-1} A D_{\mathbf{v}_\beta}$ is $\boldsymbol{\phi}_\beta = D_{\mathbf{v}_\beta}^{-1} \phi_0$. It is associated with the eigenvalue ρ_0. Thus the canonical RW has the transition matrix Π_* is given by

$$\Pi_* = \rho_0^{-1} D_{\phi_\beta}^{-1} W_\beta D_{\phi_\beta} = \rho_0^{-1} D_{\phi_0}^{-1} A D_{\phi_0} = \Pi_*^0. \tag{32}$$

Its invariant measure is $\boldsymbol{\pi}_* = \boldsymbol{\psi}_0 * \boldsymbol{\psi}_0$. The canonical RW consistent with the potential weight matrix, $W_\beta = D_{\mathbf{v}_\beta}^{-1} A D_{\mathbf{v}_\beta}$, always coincides with the canonical RW consistent with its adjacency matrix A governed by Π_*^0.

With $\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta$ with entries

$$\Pi(i, j) = A(i, j)\, \mathrm{e}^{-\beta(U(j) - U(i))} \Big/ \sum_j A(i, j)\, \mathrm{e}^{-\beta(U(j) - U(i))},$$

and with invariant measure $\mu(i)$ displayed in (31), for all β, we get

$$\begin{aligned} \log \rho_0 &> -\sum_i \mu(i) \sum_j \Pi(i, j) \log \Pi(i, j) - \beta \sum_i \mu(i) \sum_j \Pi(i, j)(U(j) - U(i)) \\ &= -\sum_i \mu(i) \sum_j \Pi(i, j) \log \Pi(i, j) = s. \end{aligned}$$

We conclude that for potential kernels K_β, the entropy production rate of the RW with probability transition matrix $\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta = D_{\mathbf{w}_\beta}^{-1} [A * K_\beta]$ is always bounded above by $\log \rho_0$.

Remark. The W_β-consistent RW with transition matrix $\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta$ associated with the weight kernel $W_\beta = A D_{\mathbf{v}_\beta}$ was also considered in [5]. For this model, the cost of a transition from i to j only depends on the terminal state, regardless of where one starts from. Although the latter is not in the potential class, its invariant measure is also given by (31).

4.3. The symmetric case

If $H(i, j) = H(j, i)$, then $W_\beta = W'_\beta$ itself. For example, $H(i, j) = |U(j) - U(i)|$ for some potential U attached to the nodes of the graph, or $H(i, j)$ is some distance (ultrametric or not) between nodes i and j. In this case, for all β, the invariant measure $\boldsymbol{\mu}_*$ of the canonical RW governed by $\Pi_* = \frac{1}{\rho_\beta} D_{\phi_\beta}^{-1} W_\beta D_{\phi_\beta}$, is

$$\boldsymbol{\mu}_* = \boldsymbol{\psi}_\beta * \boldsymbol{\psi}_\beta, \tag{33}$$

and the corresponding RW is reversible.

When $W_\beta = W'_\beta$, the invariant measure associated with $\Pi = D_{\mathbf{w}_\beta}^{-1} W_\beta$ satisfying $\boldsymbol{\mu}' = \boldsymbol{\mu}'\Pi$ is given by

$$\mu(i) = \frac{w_\beta(i)}{\sum_i w_\beta(i)} = \frac{\sum_j A(i, j)\,\mathrm{e}^{-\beta H(i,j)}}{\sum_{i,j} A(i, j)\,\mathrm{e}^{-\beta H(i,j)}}.$$

We have $\overleftarrow{\Pi}' = D_{\boldsymbol{\mu}} \Pi D_{\boldsymbol{\mu}}^{-1} = D_{\mathbf{w}_\beta} D_{\mathbf{w}_\beta}^{-1} W_\beta D_{\mathbf{w}_\beta}^{-1} = W_\beta D_{\mathbf{w}_\beta}^{-1} = \Pi'$ so that $\overleftarrow{\Pi} = \Pi$: detailed balance also holds.

References

[1] Aldous D and Fill J 2000 *Reversible Markov Chains and Random Walks on Graphs* in preparation http://www.stat.berkeley.edu/~aldous/RWG/book.html

[2] Arnold L, Gundlach V M and Demetrius L 1994 Evolutionary formalism for products of positive random matrices *Ann. Appl. Probab.* **4** 859–901

[3] Burda Z, Duda J, Luck J M and Waclaw B 2009 Localization of maximal entropy random walk *Phys. Rev. Lett.* **102** 160602

[4] Friedland S and Karlin S 1975 Some inequalities for the spectral radius of non-negative matrices and applications *Duke Math. J.* **42** 459–90

[5] Gomez-Gardenes J and Latora V 2008 Entropy rate of diffusion processes on complex networks *Phys. Rev.* E **78** 065102

[6] Horn R A and Johnson C R 1985 *Matrix Analysis* (Cambridge: Cambridge University Press)

[7] Lambert A 2008 Population dynamics and random genealogies *Stoch. Models* **24 (Suppl. 1)** 45–163

[8] Rached Z, Alajaji F and Campbell L 2001 Rényi's divergence and entropy rates for finite alphabet Markov sources *IEEE Trans. Inf. Theory* **47** 1553–61

[9] Romanovsky V I 1970 *Discrete Markov chains* ed E Seneta (Groningen: Wolters-Noordhoff Publishing) (translated from Russian)

[10] Ruelle D 2004 Thermodynamic formalism *The Mathematical Structures of Equilibrium Statistical Mechanics* 2nd edn *(Cambridge Mathematical Library)* (Cambridge: Cambridge University Press)

[11] Yaglom A M 1947 Certain limit theorems of the theory of branching random processes *Dokl. Akad. Nauk SSSR (N. S.)* **56** 795–8