

# Ruelle-Bowen continuous-time random walk

Yongxin Chen, Tryphon T. Georgiou, Michele Pavon

**Abstract**—We define the probability structure of a continuous-time time-homogeneous Markov jump process, on a finite graph, that represents the continuous-time counterpart of the so-called Ruelle-Bowen discrete-time random walk. It constitutes the unique jump process having maximal entropy rate. Moreover, it has the property that, given the number of jumps between any two specified end-points on the graph, the probability of traversing any one of the alternative paths that are consistent with the specified number of jumps and end-points, is the same for all, and thereby depends only on the number of jumps and the end-points and not the particular path being traversed.

## I. INTRODUCTION

The motivation for this note stems from our recent work [1], [2], where the topic was the scheduling for transporting resources over a graph. Indeed, a theoretical framework was developed based on the Schrödinger bridge problem, namely, to identify a probability law on alternative paths through which (probability) mass is transported between end-point marginals. The sought probability law was taken as the closest to that of the discrete-time Ruelle-Bowen (RB) random walk [3], [5], [6], so as to disperse the flow of mass maximally over available paths and, thereby, ensure a level of robustness.

In the present note, we discuss and develop the continuous-time counterpart of the discrete-time RB-walk, as a preparation towards subsequent development of transport protocols emulating our previous aforementioned work [1], [2]. The construction builds on a standard model for jump processes as well as the discrete-time RB-probability law.

## II. MARKOV JUMP PROCESSES

We consider a time-homogeneous Markov jump process  $\{X_t \mid t \geq 0\}$  with finite state-space  $\mathcal{V} = \{1, 2, \dots, n\}$ . This is identified with the vertices (nodes) of a strongly connected directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  having no self loops. We suppose that the transition probabilities  $p_{ij}(t)$  are continuous at  $t = 0$  and that  $\lim_{t \rightarrow 0} p_{i,j}(t) = \delta_{ij}$ , where  $\delta_{ij} = 1$  when  $i = j$  and 0 otherwise. Then, as is well known, the transition probabilities are of the form

$$[p_{ij}(t)]_{i,j=1}^n = \exp(Qt),$$

Y. Chen Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011; email: yongchen@iastate.edu

T. T. Georgiou is with the Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA; email: tryphon@uci.edu

M. Pavon is with the Dipartimento di Matematica “Tullio Levi Civita”, Università di Padova, 35121 Padova, Italy; email: pavon@math.unipd.it

where the infinitesimal generator

$$Q = [q_{ij}]_{i,j=1}^n$$

of the process satisfies

$$\begin{aligned} -q_{ii} &=: q_i > 0, \\ q_{ij} &\geq 0 \text{ for } i \neq j, \\ q_i &= \sum_{j \neq i} q_{ij}. \end{aligned}$$

That is,  $Q\mathbf{1} = \mathbf{0}$ , where  $\mathbf{1}, \mathbf{0}$  denote column vectors with ones and zeros, respectively. We note that the chain is ergodic,  $p_{ij}(t) > 0$  for all  $i, j$  and  $t > 0$ , which follows from the fact that the graph is strongly connected.

It is quite standard to realize the jump process [4, Chapter 2] via the bivariate process

$$\{(Z_k, T_k) \mid k = 0, 1, \dots\}$$

on the product space  $\mathcal{S} = \mathcal{V} \times (0, \infty)$ , where  $Z_k$  is a discrete-time Markov chain with transition probabilities

$$\pi_{ij} = \mathbb{P}\{Z_{k+1} = j \mid Z_k = i\} = \frac{q_{ij}}{q_i},$$

taking place at a random time  $t_{k+1} = \sum_{i=0}^k T_i, t_0 = 0$ , where  $T_k$  are random times exponentially distributed with density  $q_{z_k} \exp(-q_{z_k} t)$  representing time-intervals between successive transitions between elements of  $\mathcal{V}$ . Thence,  $X_t$  is the right-continuous process

$$X_t = Z_k \text{ for } t_k \leq t < t_{k+1}.$$

Note that, in general, the random time  $T_k$  is a function of  $Z_k$ . The bivariate realization of  $X_t$  models a random walker on the network, taking steps and transitioning from one node to the next, according to this probabilistic model.

## III. DISCRETE-TIME MARKOV CHAINS AND THE RUELLE BOWEN RANDOM WALK

We briefly outline the construction of the discrete-time RB random walk [3] on  $\mathcal{G}$ , as it will be needed in the continuous-time construction that follows.

Once again we begin with a strongly connected directed graph<sup>1</sup>  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , having nodes  $\mathcal{V} = \{1, 2, \dots, n\}$ , whose adjacency matrix is denoted by  $A = [a_{ij}]_{i,j=1}^n$ . We consider a stationary, i.e., time-homogeneous, discrete-time Markov chain  $\{Z_k \mid k = 0, 1, \dots\}$  taking values on  $\mathcal{V}$ , and we let  $P = [p_{ij}]_{i,j=1}^n$  denote the matrix of transition probabilities

<sup>1</sup>In discrete-time, self loops are allowed.

and  $\pi = [\pi_i]_{i=1}^n$  the corresponding stationary probability distribution.<sup>2</sup> The entropy rate of the process is

$$H(P) = - \sum_{i,j} \pi_i p_{ij} \log p_{ij}. \quad (1)$$

The discrete-time RB random walk is the Markov chain that is consistent with the topology of the graph and has maximal entropy rate. It is well known that this is unique [5], [3] and the construction proceeds as follows. Its transition probability matrix  $P^{\text{RB}} = [p_{ij}]_{i,j=1}^n$  has entries

$$p_{ij}^{\text{RB}} = \frac{1}{\lambda_A} \frac{\varphi_j}{\varphi_i},$$

where  $\varphi$  is the right Frobenius-Perron eigenvector of  $A$  and  $\lambda_A$  the corresponding Frobenius-Perron eigenvalue, i.e.,  $A\varphi = \lambda_A\varphi$  and  $\lambda_A > 0$  and maximal. The stationary distribution  $\pi_i = \hat{\varphi}_i\varphi_i$  where  $\hat{\varphi}$  is the left Frobenius-Perron eigenvector of  $A$ , normalized so that  $\sum_i \pi_i = 1$ . It is easy to verify that the entropy rate of the corresponding discrete-time Markov chain is in fact  $\log \lambda_A$ , and that the probability of transitioning between states  $i$  and  $j$  in  $N$  steps following any of the several possible alternative  $N$ -step paths is  $\hat{\varphi}_i\varphi_j/\lambda_A^N$ . This is a rather remarkable property, that the probability is independent of which of the alternative paths was taken! The chances that the random walker takes any of those is the same, and the entropy rate is in fact the topological entropy of the graph. Thus, the purpose of this note is to see that there is a completely analogous continuous-time random walk with exactly the same properties.

#### IV. CONTINUOUS-TIME RUELLE BOWEN RANDOM WALK

We will now resume with the construction of a continuous-time analogue of the Ruelle Bowen random walk, which is the main contribution of this work. As in the discrete-time case, the process maximizes the entropy rate over all such continuous-time random walks. As noted, interestingly, it also equalizes the probability of traversing any path, for any specified number of transitions in its flight between any two specific nodes, in complete analogy with the standard RB random walk [3].

Consider the time-homogenous Markov jump process  $X_t$  as before, with generator  $Q$  and invariant measure  $\pi$ , and the right-continuous discretized-in-time process

$$X_t^\Delta := X_{k\Delta}, \text{ for } k\Delta \leq t < (k+1)\Delta.$$

The entropy rate of  $X_t^\Delta$  is the same as that of the discrete-time Markov chain with transition probability matrix  $\Pi = \exp(Q\Delta)$ . For  $\Delta$  small, the entropy rate of  $X_{k\Delta}$ , and hence, of  $X_t^\Delta$  as they carry the same amount of information, is

$$H(X_{k\Delta}) \simeq -\Delta(1 - \log \Delta) \sum_i \pi_i q_{ii} - \Delta \sum_{j \neq i} \pi_i q_{ij} \log q_{ij}.$$

Clearly, the differential entropy rate  $H(X_{k\Delta})/\Delta$  grows unbounded as the time-scale resolution  $\Delta$  goes to zero.

<sup>2</sup>Vectors, such as  $\pi$ , will be thought as column vectors throughout.

A meaningful way to define a suitable rate of information for a jump process is separate the information content of the transition between nodes with that of the actual timing. Thus, we may specify at the outset the value of the average “retention rate”  $-\sum_i \pi_i q_{ii}$  at nodes. Equivalently, we may define instead as *differential entropy rate* for the jump process  $X_t$

$$h_\eta(Q) := -\eta \sum_i \pi_i q_{ii} - \sum_{i \neq j} \pi_i q_{ij} \log q_{ij}, \quad (2)$$

for any fixed choice of a value  $\eta > 0$ . The parameter  $\eta$  effectively serves as a Lagrange multiplier in the optimization that follows, that dictates a corresponding value for  $-\sum_i \pi_i q_{ii}$ . Whatever the interpretation, our problem is now to seek the maximum of  $h_\eta(Q)$  over choices of the infinitesimal generator matrix  $Q$  and in accordance with the geometry of the given graph  $\mathcal{G}$ .

Since the graph is strongly connected, a sufficiently high power of the adjacency matrix  $A$  has all of its entries strictly positive. By Frobenius-Perron theory,  $A$  has a unique eigenvalue  $\lambda_A$  having maximal modulus which is also real and positive. The corresponding right and left eigenvectors  $\varphi$ ,  $\hat{\varphi}$  can be taken to have positive entries, and also normalized so that

$$\langle \varphi, \hat{\varphi} \rangle = \sum_{i=1}^n \varphi_i \hat{\varphi}_i = 1.$$

As before,

$$\pi^{\text{RB}} := \varphi \hat{\varphi}, \quad (3)$$

but now we also define

$$Q^{\text{RB}} = \text{diag}(\varphi)^{-1} A \text{diag}(\varphi) - \lambda_A I. \quad (4)$$

Clearly,  $Q^{\text{RB}}$  is a generator having  $\pi^{\text{RB}}$  as invariant measure, since it has non-negative off-diagonal entries, positive diagonal entries,

$$\begin{aligned} Q^{\text{RB}} \mathbf{1} &= \text{diag}(\varphi)^{-1} A \varphi - \lambda_A \mathbf{1} \\ &= \text{diag}(\varphi)^{-1} \lambda_A \varphi - \lambda_A \mathbf{1} = \mathbf{0}, \end{aligned}$$

and

$$\begin{aligned} (Q^{\text{RB}})' \pi^{\text{RB}} &= \text{diag}(\varphi) A' \hat{\varphi} - \lambda_A \pi^{\text{RB}} \\ &= \lambda_A \text{diag}(\varphi) \hat{\varphi} - \lambda_A \pi^{\text{RB}} = \mathbf{0}, \end{aligned}$$

where “/” denotes transposition and  $\pi^{\text{RB}}$  is thought of as a column vector.

*Theorem 1:* The differential entropy rate  $h_1(Q)$  has a unique maximum at  $Q^{\text{RB}}$  given above, over all infinitesimal generators that are consistent with the topology of a graph with adjacency matrix  $A$ . Moreover,

$$h_1(Q^{\text{RB}}) = \lambda_A.$$

**Proof:** Define  $r_{ij} = \pi_i Q_{ij}$  (i.e., for  $j \neq i$  these represent “flow rates” from node  $i$  to  $j$ ), then maximizing  $h(Q)$  reduces to maximizing

$$f(r, \pi) = - \sum_i r_{ii} - \sum_{i \neq j} r_{ij} a_{ij} \log \frac{r_{ij}}{\pi_i a_{ij}}$$

over  $r, \pi$  subject to the constraints

$$\sum_j r_{ij} = 0, \quad \text{and} \quad \sum_i r_{ij} = 0.$$

To see this note that  $R = [r_{ij}]_{i,j=1}^n$  equals  $\text{diag}(\pi) \times Q$ , and the above constraints are precisely the requirements that  $Q\mathbf{1} = \mathbf{0}$  and  $Q'\pi = \mathbf{0}$ .

Introducing Lagrangian multipliers  $\alpha, \beta$  for these two constraints gives

$$\begin{aligned} \mathcal{L}(r, \pi, \alpha, \beta) = & - \sum_i r_{ii} - \sum_{i \neq j} r_{ij} a_{ij} \log \frac{r_{ij}}{\pi_i a_{ij}} \\ & + \sum_i \alpha_i \sum_j r_{ij} + \sum_j \beta_j \sum_i r_{ij}. \end{aligned}$$

By standard duality theory we have that the maximum of  $h_1(Q)$  is bounded above by

$$\min_{\alpha, \beta} \max_{r, \pi} \mathcal{L}(r, \pi, \alpha, \beta).$$

Now choose a specific pair of  $\alpha^*$  and  $\beta^*$  as

$$\beta_j^* = \log \varphi_j, \quad \text{and} \quad \alpha_i^* = 1 - \beta_i^*,$$

then

$$\min_{\alpha, \beta} \max_{r, \pi} \mathcal{L}(r, \pi, \alpha, \beta) \leq \max_{r, \pi} \mathcal{L}(r, \pi, \alpha^*, \beta^*).$$

For any fixed  $\pi$ , we have that

$$\mathcal{L}(r, \pi, \alpha^*, \beta^*) = \sum_{i \neq j} (-r_{ij} a_{ij} \log \frac{r_{ij}}{\pi_i a_{ij}} + (\alpha_i^* + \beta_j^*) r_{ij})$$

is a concave function of  $r$ . Considering first order optimality conditions we conclude that the maximizer is

$$r_{ij}^* = \begin{cases} \pi_i e^{\alpha_i^* + \beta_j^* - 1} & \text{if } a_{ij} = 1, \\ 0 & \text{if } a_{ij} = 0. \end{cases}$$

corresponding to the maximal value

$$\begin{aligned} & \sum_{i \neq j} (-r_{ij}^* a_{ij} \log \frac{r_{ij}^*}{\pi_i a_{ij}} + (\alpha_i^* + \beta_j^*) r_{ij}^*) \\ = & \sum_{i \neq j} -\pi_i e^{\alpha_i^* + \beta_j^* - 1} a_{ij} \log \frac{\pi_i e^{\alpha_i^* + \beta_j^* - 1}}{\pi_i a_{ij} e^{\alpha_i^* + \beta_j^*}} \\ = & \sum_{i \neq j} \pi_i e^{\alpha_i^* + \beta_j^* - 1} a_{ij} \\ = & \sum_{i \neq j} \pi_i \frac{\varphi_j}{\varphi_i} a_{ij} = \lambda_A. \end{aligned}$$

Therefore  $h_1(Q)$  is bounded above by  $\lambda_A$ . On the other hand, with  $Q = Q^{\text{RB}}$ ,

$$\begin{aligned} h_1(Q^{\text{RB}}) &= - \sum_i \pi_i q_{ii}^{\text{RB}} - \sum_{i \neq j} \pi_i q_{ij}^{\text{RB}} \log q_{ij}^{\text{RB}} \\ &= \lambda_A - \sum_{i \neq j} \pi_i \frac{\varphi_j}{\varphi_i} a_{ij} (\log \varphi_j - \log \varphi_i) \\ &= \lambda_A - \lambda_A \sum_j \hat{\varphi}_j \varphi_j \log \varphi_j + \lambda_A \sum_i \pi_i \log \varphi_i \\ &= \lambda_A, \end{aligned}$$

and achieves the upper bound. This completes the proof.  $\square$

The argument carries through for other values of  $\eta$ , in which case the solution is a scaled multiple of (4), corresponding to a different value for the average rate  $-\sum_i \pi_i q_{ii}$  for staying at the current node.

*Corollary 1:* The differential entropy rate  $h_\eta(Q)$  has a unique maximum at  $e^{\eta-1} Q^{\text{RB}}$  with  $Q^{\text{RB}}$  given in (4), over all infinitesimal generators that are consistent with the topology of a graph with adjacency matrix  $A$ . Moreover,

$$h_\eta(e^{\eta-1} Q^{\text{RB}}) = e^{\eta-1} \lambda_A.$$

The continuous-time time-homogeneous Markov jump process with infinitesimal generator  $Q^{\text{RB}}$  in (4) we refer to as the continuous-time Ruelle-Bowen random walk. We now complete the final claim in this note that the continuous-time RB random walk shares the property, of its discrete-time counterpart, of equalizing probability across alternative paths. The proof is almost verbatim with only subtlety the need to condition on the number of steps taken in a flight between two nodes, since we are now dealing with a continuous-time jump process. But this can be easily done using the realization of the process as outlined in Section II.

*Theorem 2:* Let  $X_t^{\text{RB}}$  be a continuous-time RB walk on a graph  $\mathcal{G}$ . The probability of transitioning from node  $i$  to node  $j$ , over any specified window of time, say  $[0, t_f]$ , via a sequence of precisely  $N$  nodes, is independent of the particular sequence of nodes.

**Proof:** Consider the corresponding bivariate process  $(Z_k^{\text{RB}}, T_k^{\text{RB}})$ . The key is to observe that the random time  $T_k$  are identically distributed with the same exponent,  $-\lambda_A$ , independent of  $Z_k$ . In fact,  $Z_k$  and  $T_k$  are independent. Thus,

$$\begin{aligned} \mathbb{P}(T_0 = t_1, \sum_0^1 T_k = t_2, \dots, \sum_0^{N-1} T_k = t_N, \\ t_N \leq t_f < t_{N+1}, \\ X_0 = i, \dots, X_{t_f} = j) \end{aligned}$$

is equal to the product

$$\mathbb{P}(t_N \leq t_f < t_{N+1}) \times \mathbb{P}(Z_0 = i, \dots, Z_N = j).$$

The former is independent of the particular path and the latter, using the facts detailed in Section III, is

$$\frac{\hat{\varphi}_i \varphi_j}{\lambda_A^N}.$$

Thus, the discrete-time analysis applies verbatim and the probability is independent of the path and only a function of beginning and ending nodes, and of the number of steps taken.  $\square$

The relation between the infinitesimal generator  $Q^{\text{RB}}$  of the continuous-time RB walk and the transition probability matrix  $P$  of the discrete-time RB walk, on the same graph, is

$$Q^{\text{RB}} = \lambda_A (P^{\text{RB}} - I).$$

## REFERENCES

- [1] Y. Chen, T.T. Georgiou, M. Pavon and A. Tannenbaum, Robust transport over networks. *IEEE Trans. Aut. Control*, 62(9):4675–4682, 2017.
- [2] Y. Chen, T.T. Georgiou, M. Pavon and A. Tannenbaum, Efficient-robust routing for single commodity network flows, *IEEE Trans. Aut. Control*, to appear, arXiv: 1701.07625v2
- [3] Jean-Charles Delvenne and Anne-Sophie Libert. Centrality measures and thermodynamic formalism for complex networks. *Physical Review E*, 83(4):046117, 2011.
- [4] James R. Norris, Markov chains. Cambridge university press, 1998.
- [5] William Parry. Intrinsic markov chains. *Transactions of the American Mathematical Society*, 112(1):55–66, 1964.
- [6] Ruelle, David. Thermodynamic formalism: the mathematical structure of equilibrium statistical mechanics. Cambridge University Press, 2004.