On the fastest finite Markov processes

Vivek Borkar, Laurent Miclo

On the fastest finite Markov processes*

Vivek Borkar† and Laurent Miclo‡

†Department of Electrical Engineering
Indian Institute of Technology Bombay, India

‡Institut de Mathématiques de Toulouse, UMR 5219
Université de Toulouse and CNRS, France

Abstract

Consider a finite irreducible Markov chain with invariant probability π. Define its inverse communication speed as the expectation to go from x to y, when x, y are sampled independently according to π. In the discrete time setting and when π is the uniform distribution υ, Litvak and Ejov [10] have shown that the permutation matrices associated to Hamiltonian cycles are the fastest Markov chains. Here we prove (A) that the above optimality is with respect to all processes compatible with a fixed graph of permitted transitions (assuming that it does contain a Hamiltonian cycle), not only the Markov chains, and, (B) that this result admits a natural extension in both discrete and continuous time when π is close to υ: the fastest Markov chains/processes are those moving successively on the points of a Hamiltonian cycle, with transition probabilities/jump rates dictated by π. Nevertheless, the claim is no longer true when π is significantly different from υ.

Keywords: Fastest Markov chains/processes, communication speed, spectra of Markov operators, uniform distribution, Hamiltonian cycles, dynamic programming, differentiation of Markov operators.

MSC2010: primary: 60J22, secondary: 60J27, 60J10, 15A42, 05C45, 49L20, 90C40.

*Work supported by the Indo-French Centre for Applied Mathematics and by the ANR-12-BS01-0019.

1 Introduction

Given a finite oriented (strongly) connected graph $G = (V, E)$ and a positive probability measure π on V, it is natural to wonder what is the fastest Markov chain leaving π invariant and whose permitted transitions are included in E. This depends on the way the speed is measured. In this paper the goal is to minimize the expectation $\mathfrak{F}$ of the time needed to go from x to y, when x and y are independently sampled according to π. Litvak and Ejov [10] have shown that if π is the **uniform distribution** υ and if G contains a Hamiltonian cycle, then the fastest Markov chains are exactly those following deterministically the succession of the states given by a Hamiltonian cycle when one exists (the corresponding quantity $\mathfrak{F}$ does not depend on the choice of the admissible Hamiltonian cycle). Our objectives in this paper are: (A) to extend this result to the continuous time framework (under an appropriate renormalization of the jump rates), (B) to establish the above optimality over a larger class of processes, and to begin an investigation of the situation where π is not the uniform distribution by showing, (C) that when G contains a Hamiltonian cycle and that π is close to υ, the fastest Markov chains/processes are still those appropriately associated to Hamiltonian cycles, and (D) that this is no longer true when π is 'far away' from υ.

The plan of the paper is as follows. The above results (A) and (B) are proved in the next section via a dynamic programming approach, which also provides an alternative proof of the discrete time result of Litvak and Ejov [10]. In Section 3, we decompose the generators leaving π invariant into convex sums of generators associated to (not necessarily Hamiltonian) cycles and we differentiate the expectations of hitting times with respect to the generators. This is the basic tool for the proof of (C) (see Theorem 5 in Section 3) in Section 4 , through small perturbations of the uniform probability measure. At the other extreme, large perturbations lead to the proof of (D) (cf. Theorem 6 in Section 3) at the end of the same section. Section 5 contains some observations about the links between continuous time and discrete time. In the appendix, we compute the fastest normalized birth and death generators leaving invariant any fixed positive probability measure π on $\{0, 1, 2\}$. The underlying graph is the segment graph of length 2, i.e. the simplest example not containing a Hamiltonian cycle.

2 The dynamic programming approach

2.1 Introduction

The aim of this section is to show that the Hamiltonian cycles, when one exists, are the fastest in the sense we have defined among *all* processes compatible with the given graph, not just the Markov chains. The proof uses dynamic programming. We first recall the eigentime identity in the next subsection and then establish the desired result for resp. discrete and continuous time in the subsections that follow.

2.2 The eigentime identity

We shall use the notation $\mathcal{L}(X)$ to denote the law of a random variable X and $|A|$ for the cardinality of a finite set A. Consider a discrete time Markov chain $(X_n)_{n\in\mathbb{Z}_+}$ on a finite state space V with transition matrix $P = (p(i,j))_{i,j\in V}$. We assume it to be **irreducible**, i.e., for any $i, j \in V$, there exists a path $i_0 = i, i_1, \cdots, i_{n-1}, i_n = j$ such that $p(i_k, i_{k+1}) > 0$, for $k \in [\![0, n-1]\!] \coloneqq \{0, 1, ..., n-1\}$. Let $\pi \coloneqq (\pi(i))_{i\in V}$ denote its unique **stationary distribution**, which is its left eigenvector corresponding to the Perron-Frobenius eigenvalue $\theta_1 = 1$. In particular, if $\mathcal{L}(X_0) = \pi$, then for any $n \in \mathbb{Z}_+$, we have $\mathcal{L}(X_n) = \pi$. This justifies the term 'stationary', its uniqueness being

a well-known consequence of the irreducibility hypothesis. Denote by $\theta_2, \cdots, \theta_{|V|}$ the remaining eigenvalues of P. Also define the hitting times

$$\forall\ i \in V, \qquad \tau_i \ := \ \min\{n \in \mathbb{Z}_+ \ : \ X_n = i\}. \tag{1}$$

The **eigentime identity** states that

$$\forall\ i \in V, \qquad \sum_j \pi(j) E_i[\tau_j] \ = \ \sum_{m=2}^{|V|} \frac{1}{1-\theta_m}. \tag{2}$$

Here each eigenvalue is counted as many times as its (algebraic) multiplicity. For reversible chains, this is Proposition 3.13, p. 75, of Aldous and Fill [1]. It was extended to the general case in Cui and Mao [2], see also [11] for a simple proof and further extensions. The left hand side gives the mean hitting time of a target state picked randomly with distribution π. A minor modification is to consider instead the stopping times

$$\forall\ i \in V, \qquad T_i \ := \ \min\{n \in \mathbb{N} \ : \ X_n = i\},$$

where $\mathbb{N} := \mathbb{Z}_+ \backslash \{0\}$, the set of positive integers. Since $\pi(i) E_i[T_i] = 1$, we can replace (2) by

$$\forall\ i \in V, \qquad \sum_j \pi(j) E_i[T_j] \ = \ 1 + \sum_{m=2}^{|V|} \frac{1}{1-\theta_m}. \tag{3}$$

This variant is essentially contained in Theorem 2.4 of Hunter [6], which also gives the pre–history of the problem going back to Kemeny and Snell [9]. An immediate corollary is the 'symmetrized' version

$$\sum_{i,j} \pi(i)\pi(j) E_i[\tau_j] \ = \ \sum_{m=2}^{|V|} \frac{1}{1-\theta_m}, \tag{4}$$

$$\sum_{i,j} \pi(i)\pi(j) E_i[T_j] \ = \ 1 + \sum_{m=2}^{|V|} \frac{1}{1-\theta_m}. \tag{5}$$

Let Π be the rank one matrix whose rows are all equal to π. From Theorem 2.4 of Hunter [6], we have:

(5) equals $\mathrm{tr}(I - P + \Pi)^{-1}$, where I is the identity matrix. $(*)$

A **Hamiltonian cycle** A of V is an ordering $(a_0, a_1, ..., a_{N-1})$ of the elements of V, where $N := |V|$. We will make the convention that $a_N = a_0$, as the indices should be seen as elements of $\mathbb{Z}_N := \mathbb{Z}/(N\mathbb{Z})$. More precisely, the cycles $(a_0, a_1, ..., a_{N-1})$ and $(a_k, a_{k+1}, ..., a_{k+N-1})$ should be identified as the same cycle, for all $k \in \mathbb{Z}_N$. This will be implicit in the sequel, even though for notational convenience, we will represent a cycle A as $(a_0, a_1, ..., a_{N-1})$. Consider an irreducible directed graph $G = (V, E)$ where V, E denote respectively its node and edge sets. In the discrete time setting, such graphs will always be assumed to contain all the self-loops, i.e. $(i, i) \in E$ for any $i \in V$. The Hamiltonian cycle $A := (a_0, a_1, ..., a_{N-1})$ is said to be **admissible** for G if $(a_k, a_{k+1}) \in E$ for all $k \in \mathbb{Z}_n$. It means that G_A is a subgraph of G, where G_A is the oriented graph on V whose edges are the (a_k, a_{k+1}), for $k \in \mathbb{Z}_N$. The set of all Hamiltonian cycles (respectively, admissible for G) is denoted $\mathcal{H}$ (resp., $\mathcal{H}(G)$) and the graph G is said to be **Hamiltonian** if $\mathcal{H}(G) \neq \varnothing$. Obviously, we have $\mathcal{H} = \mathcal{H}(K_V)$, where K_V is the complete oriented graph on V.

Consider the optimization problem of minimizing (2)/(4), or equivalently, (3)/(5), where π is equal to υ, over all irreducible P compatible with the given graph G (in the sense that for any

$i \neq j \in V$, $P(i,j) > 0 \Rightarrow (i,j) \in E$). Since these quantities will be infinite for reducible P, we might as well consider the problem of minimizing it over all stochastic matrices P compatible with G. Say that P is Hamiltonian if there exists a Hamiltonian cycle $(a_1, \cdots, a_{|V|})$ such that $p(a_k, a_{k+1}) = 1 = p(a_{|V|}, a_1)$ for $1 \leqslant k < |V|$. That is, the transitions deterministically trace a Hamiltonian cycle. Recall that a Hamiltonian cycle need not exist in general and the problem of determining whether one does is NP-hard (see, e.g., Garey and Johnson [4]). By Proposition 2.1 of Litvak and Ejov [10] in combination with $(*)$ above, we have:

Theorem 1 *When $\pi = \upsilon$, either of the quantities (2), (3) is minimized by a Hamiltonian P if there exists one.*

In the next subsection, we give an alternative proof, inspired by the Held-Karp algorithm for scheduling problems [5], which gives a strengthening of this result and has interesting implications for random search. Specifically, we improve on the cited result of Litvak and Ejov [10], insofar as the cost is shown to be minimized over *all* G-compatible random processes and not only the Markov chains, by tracing the Hamiltonian cycle, when one exists, deterministically.

2.3 A dynamic programming solution

As in Held and Karp [5], a natural state space for the dynamic program is

$$V^* := \{(i, A) : i \in V,\ A \subset V \backslash \{i\}\}.$$

With each $(i, A) \in V^*$, we associate an action space $U_i :=$ the set of probability vectors on the set

$$V_i := \{j \in V : (i,j) \in E\} \subset V$$

of successors of i in G. Note that this does not depend on A. Suppose V_i is enumerated as $(j_1, \cdots, j_{m_i})$. Given a 'control' $q = (q(j_1), \cdots, q(j_{m_i}))$, the transition probability

$$\hat{p}((j,B)|(i,A),q) \tag{6}$$

of going from $(i, A) \in V^*$ to $(j, B) \in V^*$ under control q is zero if either $j \notin V_i$ or $B \neq A \backslash \{j\}$. Otherwise it equals $q(j)$. Consider an V^*-valued controlled Markov chain $(X_n, Z_n)_{n \in \mathbb{Z}_+}$ governed by a control process $(q_n)_{n \in \mathbb{Z}_+}$ with $q_n \in U_{X_n}$, for all $n \in \mathbb{Z}_+$, evolving according to the above controlled transition probability function. That is, for any $n \in \mathbb{Z}_+$,

$$\begin{aligned} P((X_{n+1}, Z_{n+1}) = (j,B)|(X_m, Z_m), q_m, m \leqslant n) \\ = \quad P((X_{n+1}, Z_{n+1}) = (j,B)|(X_n, Z_n), q_n) \\ = \quad q_n(j)\delta_{B, A\backslash\{j\}}, \end{aligned}$$

where $\delta_{\cdot,\cdot}$ denotes the Kronecker delta. Since we are allowed to choose any past dependent transition probability compatible with G, this covers *all* V-valued random processes that are compatible with G, i.e., that make transitions only along the edges in E.

Our objective is to minimize, for a prescribed initial state i_0[1] the quantity

$$E\left[\sum_j \tau_j \;\middle|\; X_0 = i_0, Z_0 = V\backslash\{i_0\}\right], \tag{7}$$

[1] more generally, for a prescribed initial distribution

which is proportional to (2) when $\pi = \upsilon$, the uniform distribution (i.e., when P is doubly stochastic). Note, however, that we do not require $(X_n)_{n\in\mathbb{Z}_+}$ to be even Markov. Let

$$\zeta := \min\{n \geqslant 0 : Z_n = \varnothing\}.$$

Then (7) can be equivalently written as

$$E\left[\sum_{m=0}^{\zeta} |Z_n| \;\Big|\; X_0 = i_0, Z_0 = V\backslash\{i_0\}\right]. \tag{8}$$

This allows us to apply the dynamic programming principle to the 'value function' or 'cost to go function'

$$V(i, A) := \inf E\left[\sum_{m=0}^{\zeta} |Z_n| \;\Big|\; X_0 = i, Z_0 = A\right],$$

where the infimum is over all admissible controls. Standard arguments yield the dynamic programming equation

$$V(i, A) = \min_{q\in U_i}\left(|A| + \sum_{j\in V_i} q(j)V(j, A\backslash\{j\})\right),\ A \neq \varnothing, \tag{9}$$

$$V(\cdot, \varnothing) \equiv 0. \tag{10}$$

Furthermore, the optimal control in state (i, A) is any minimizer of the right hand side of (9). Since the expression being minimized is affine in q, this minimum will be attained at a Dirac measure, implying that the optimal choice in state (i, A) is to deterministically move to a certain $j \in V_i$. In other words, the optimal trajectory is deterministic and perforce visits each node at least once, otherwise the cost would be infinite. Since at most one new node can be visited each time, the total cost is at least $\sum_{i=1}^{|V|-1} i = \frac{|V|(|V|-1)}{2}$, which equals the cost for tracing a Hamiltonian cycle if one exists.

A parallel treatment can be given for the cost

$$E\left[\sum_j T_j \mid X_0 = i_0, Z_0 = V\right], \tag{11}$$

which can be equivalently written as

$$E\left[\sum_{m=0}^{\zeta} |Z_n| \mid X_0 = i_0, Z_0 = V\right]. \tag{12}$$

The minimum cost for this, again attained by tracing a Hamiltonian cycle deterministically, will be $\frac{s(s+1)}{2}$.

We have proved:

Theorem 2 *Minimum of either the cost (7) or the cost (11) over all V-valued random processes compatible with G is attained by tracing a Hamiltonian cycle when one exists.*

This has interesting implications to some random search schemes. For example, consider the problem of searching for an N bit binary password given a device or 'oracle' that can verify whether

a password is correct or not. Random search schemes for this problem have been proposed, involving Markov chains on the discrete N-cube $\{0,1\}^N$, where any two strings differing in one position are deemed to be neighbors. This undirected graph can be rendered directed by replacing each undirected edge by two directed edges. A simple induction argument shows that it has a Hamiltonian cycle. Then the foregoing leads to the conclusion that no random search scheme can do better on average than simply listing the N-strings and checking them one by one.

2.4 Continuous time problem

We now consider the continuous time counterparts of the foregoing. Recall that a **Markov generator** on V can be represented by a matrix $L \coloneqq (L(x,y))_{x,y\in V}$ whose off-diagonal entries are non-negative and whose row sums all vanish. Corresponding Markov processes, defined through the corresponding martingale problems, will be denoted $X \coloneqq (X_t)_{t\geqslant 0}$. The law of X then only depends on the **initial distribution**, namely on the law $\mathcal{L}(X_0)$ of X_0. The Markov generator L is said to be **compatible** with G, if we have

$$\forall\ x \neq y \in V, \qquad L(x,y) > 0 \quad \Rightarrow \quad (x,y) \in E.$$

The probability measure π, viewed as a row vector, is said to be **invariant** for the generator L, if $\pi L = 0$. Its probabilistic interpretation is that if initially $\mathcal{L}(X_0) = \pi$, then for any $t \geqslant 0$, $\mathcal{L}(X_t) = \pi$, similarly to the discrete time case. The generator L is said to be **irreducible** if for any $x, y \in V$, there exists a path $x_0 = x, x_1, ..., x_l = y$, with $l \in \mathbb{Z}_+$ the length of the path, such that $L(x_k, x_{k+1}) > 0$ for all $k \in [\![0, l-1]\!]$. In our finite setting, a Markov generator L always admits an invariant probability measure, the irreducibility of L ensures that it is unique. The irreducible Markov generator L is said to be **normalized**, if

$$\sum_{x\in V} L(x)\pi(x) \quad = \quad 1, \tag{13}$$

where μ is the invariant measure of L and where $L(x) \coloneqq -L(x,x) = \sum_{y\neq x} L(x,y)$ for any $x \in V$. It means that at its equilibrium μ (i.e. for the stationary X starting with $\mathcal{L}(X_0) = \mu$), the jump rate of X is 1. Denote by $\mathcal{L}(G,\pi)$ the convex set of irreducible normalized Markov generators L compatible with G and admitting π for invariant probability. To simplify notation, we will also write $\mathcal{L}(\pi) \coloneqq \mathcal{L}(K_V, \pi)$, when G is the complete graph K_V on V. For $y \in V$, let τ_y be the **hitting time** of y:

$$\tau_y \quad \coloneqq \quad \inf\{t \geqslant 0\ :\ X_t = y\}.$$

We are particularly interested in the functional

$$F\ :\ \mathcal{L}(G,\pi) \ni L \quad \mapsto \quad \sum_{x,y\in V} \pi(x)\pi(y)\mathbb{E}_x^L[\tau_y], \tag{14}$$

where the subscript x (respectively the superscript L) in the expectation indicates that X is starting from x (resp. is generated by L). The probabilistic interpretation of $F(L)$ is the mean time to go from x to y for the Markov process generated by L, when x and y are sampled independently according to its invariant probability π.

Remark 3 The smaller the $F(L)$, the faster the underlying Markov process goes between the elements of V. It does not necessarily imply that the faster the time-marginal distributions go to equilibrium in large time (especially in the discrete time analogue). It is more related to the asymptotic behavior of the variance associated with the convergence of the empirical measures. This point of view will not be investigated here.

□

The quantity $F(L)$ also admits a nice spectral formulation: for any $L \in \mathcal{L}(G,\pi)$, let $\Lambda(L)$ be the spectrum of $-L$, removing the eigenvalue 0. To take into account the possible multiplicities of the eigenvalues, $\Lambda(L)$ should be seen a **multiset** (i.e. a eigenvalue of $-L$ of multiplicity m, appears m times in $\Lambda(L)$). By irreducibility of L, $\Lambda(L)$ is a priori a sub(multi)set of $\mathbb{C}_+ := \{z \in \mathbb{C} : \Re(z) > 0\}$ that is invariant under conjugation. The **eigentime relation** asserts

$$F(L) \;=\; \sum_{\lambda \in \Lambda(L)} \frac{1}{\lambda}. \tag{15}$$

The references Cui and Mao [2] and [11] given in the discrete time setting also deal with the continuous time case. The quantity $F(L)$ can also be written in terms of return times. Define for any $y \in V$,

$$\begin{aligned} \sigma &:= \min\{t > 0 : X_t \neq X_0\}, \\ T_y &:= \min\{t \geqslant \sigma : X_t = y\}. \end{aligned}$$

By irreducibility of L, we have the following eigentime identities (see Cui and Mao [2]), for any $y \in V$:

$$\sum_{x \in V} \pi(x) E_x[\tau_y] \;=\; \sum_{\lambda \in \Lambda(L)} \frac{1}{\lambda} \;=\; \sum_{x,z \in V} \pi(x)\pi(z) E_x[\tau_z], \tag{16}$$

$$\sum_{x} \pi(x) E_x[T_y] \;=\; 1 + \sum_{\lambda \in \Lambda(L)} \frac{1}{\lambda} \;=\; \sum_{x,z \in V} \pi(x)\pi(z) E_x[T_z], \tag{17}$$

Similarly to Subsection 2.3, our goal is to find the minimizers of F on $\mathcal{L}(G,\pi)$, or at least to deduce some information about them, since they correspond to the fastest normalized Markov processes compatible with G with invariant distribution π. There is no loss of generality in imposing that L is irreducible, because the functional F is infinite for non-irreducible Markov generators admitting π as invariant measure.

Consider next a continuous time V-valued *controlled* Markov chain, denoted $(X_t)_{t \geqslant 0}$ again by abuse of notation, controlled by a control process $(Z_t)_{t \geqslant 0}$. The latter takes values in $\mathcal{U}_{X_t}$, where $\mathcal{U}_i := [0,\infty)^{|V_i|}$, identified with the instantaneous transition rate of X_t. That is, as δ goes to 0_+,

$$\begin{aligned} P(X_{t+\delta} = j | X_s, Z_s, s \leqslant t, X_t = i) &= P(X_{t+\delta} = j | X_t = i, Z_t) \\ &= \begin{cases} Z_t(X_t, j)\delta + o(\delta) & , \text{if } j \in V_{X_t}, \\ -\sum_{j \in V_{X_t}} Z_t(X_t, j)\delta + o(\delta) & , \text{if } j = X_t, \\ 0 & , \text{if } j \notin V_{X_t} \cup \{X_t\}, \end{cases} \end{aligned}$$

where we write $Z_t = (Z_t(X_t, j_1), \cdots, Z_t(X_t, j_{m_{X_t}}))$ for a suitable enumeration $(j_1, \cdots, j_{m_{X_t}})$ of V_{X_t}.

For the remaining part of this subsection, we consider the case where $\pi = \upsilon$, the uniform measure on V. The renormalization condition (13) can be written in the form

$$\sum_{i \neq j} Z_t(i,j) \;=\; |V|. \tag{18}$$

If for any $t \geqslant 0$, Z_t is a function of X_t alone, say $Z_t = r(X_t, \cdot) \in \mathcal{U}_{X_t}$, then $(X_t)_{t \geqslant 0}$ is a time-homogeneous Markov process with rate matrix $R = (r(i,j))_{i,j \in V}$, where we set $r(i,j) = 0$ for $j \notin V_i$. Consider the problem of minimizing (16). As before, we augment the state process to the V^*-valued process $(\hat{X}_t)_{t \geqslant 0} = (X_t, A_t)_{t \geqslant 0}$, with the understanding that A_t can change only when

X_t does and a transition of X_t from i to j leads to a transition of A_t to $A_t \backslash \{j\}$. Consider the control problem of minimizing the cost

$$E\left[\int_0^\zeta |A_t| dt \;\middle|\; X_0 = i_0, A_0 = V\backslash\{i_0\}\right], \tag{19}$$

for

$$\zeta := \{t \geqslant 0 : A_t = \varnothing\},$$

which is equivalent to (16), subject to the normalization constraint (18). The constraint (18) couples decisions across different states, so dynamic programming arguments cannot be directly applied. Therefore we modify the formulation for the time being, this modification will be dropped later. The modification is as follows. Let $(a_i)_{i\in V}$ be scalars in $(0, |V|)$ such that $\sum_i a_i = |V|$. For state i, we restrict the rates to be from the set

$$\tilde{U}_i := \{r(i,j) : r(i,j) = 0 \;\forall\; j \notin V_i,\; \sum_{j\in V_i} r(i,j) = a_i\}.$$

Consider the value function

$$V(i, A) := \inf E\left[\int_0^\zeta |A_t| dt \;\middle|\; X_0 = i, A_0 = A\right],$$

where the infimum is over all admissible controls. The dynamic programming equation then is

$$\min_{r(i,\cdot)\in\tilde{U}_i} \left(|A| + \sum_{j\in V_i, A} r(i,j)(V(j, A\backslash\{j\}) - V(i, A))\right) = 0,\; V(\cdot, \varnothing) \equiv 0. \tag{20}$$

Once again it is clear that the quantity being minimized is affine in the variables it is being minimized over and hence the optimum is attained for a deterministic choice of $r(i,\cdot)$ in the sense that $r(i,j)$ can be non-zero for at most one $j \in V_i$. Thus the optimal path traces the nodes of G in a deterministic manner, visiting each of them at least once. This is true for any choice of $\{a_i\}$ and therefore true in general for the constraint (18). Unlike the discrete time case, this does not, however, mean that the trajectory is deterministic, because the sojourn time in each node is still random. It is clear that the cost for any such trajectory will be

$$E\left[\sum_{i\in V} a_i^{-1} N_i\right],$$

where N_i is the number of times the trajectory passed through i. For a given choice of $(a_i)_{i\in V}$, this is clearly minimized if $N_i = 1$ for all $i \in V$, which can be achieved by tracing a Hamiltonian cycle if one exists. Optimizing next over the choice of $(a_i)_{i\in V}$ subject to (18), namely $\sum_{i\in V} a_i = |V|$, a simple induction argument shows that the choice $a_i = 1$, for all $i \in V$, is optimal.

As in the discrete case, a similar treatment is possible for the cost (17) or its equivalent

$$E\left[\int_0^\zeta |A_t| dt \;\middle|\; X_0 = i_0, A_0 = V\right], \tag{21}$$

with the constraint (18).

We have proved:

Theorem 4 *Minimum of either the cost (19) or the cost (21) over all V-valued random processes compatible with G is attained by tracing a Hamiltonian cycle when one exists.*

3 Perturbation of Markov generators

3.1 Introduction

We can expect the optimality of Hamiltonian cycle to persist under small perturbations of the Markov chains considered above. For specific classes of perturbations, such results were established in Ejov et al [3]. Here we establish a vastly more general result, first for the continuous time framework (which turns out to be more natural in some sense for the kind of techniques we employ) and then for the discrete case.

When $A \coloneqq (a_0, a_1, ..., a_{N-1}) \in \mathcal{H}$ and a positive probability measure π on V are fixed, the set $\mathcal{L}(G_A, \pi)$ is reduced to a singleton, its element will be denoted L_A. It is indeed given by

$$\forall\ x, y \in V, \qquad L_A(x,y) \quad = \quad \begin{cases} \frac{1}{N\pi(x)} & \text{, if } x = a_k \text{ and } y = a_{k+1} \text{ for some } k \in \mathbb{Z}_n \\ -\frac{1}{N\pi(x)} & \text{, if } x = y \\ 0 & \text{, otherwise} \end{cases}$$

Theorem 4 may seem a little deceptive: the fastest normalized Markov processes X leaving invariant υ, the uniform probability measure on V, follow a prescribed cyclic ordering of the states of V, their randomness comes only from their waiting times, distributed according to the exponential law of intensity 1. Such a derandomization of the successive points visited by X is also valid for probability measures π close to υ:

Theorem 5 *Assume that the graph G is Hamiltonian. Then there exists a neighborhood $\mathcal{N}$ of υ in the set $\mathcal{P}_+(V)$ of positive probability measures on V (endowed with the topology inherited from that of $(0,1]^V$) such that for any $\pi \in \mathcal{N}$, the set of minimizers of F on $\mathcal{L}(G,\pi)$ is exactly $\{L_A\ :\ A \in \mathcal{H}(G)\}$.*

Nevertheless, this result cannot be extended to all positive probability measures π, at least for the graphs which are not a Hamiltonian cycle, a situation where $\mathcal{L}(G,\pi)$ is not reduced to a singleton, in particular, this requires $N \geqslant 3$.

Theorem 6 *Assume that G is not a Hamiltonian cycle. Then there exist positive probability measures π on V such that none of the elements of $\{L_A\ :\ A \in \mathcal{H}(G)\}$ is a minimizer of F on $\mathcal{L}(G,\pi)$.*

Thus for some (G,π), the minimizers of F on $\mathcal{L}(G,\pi)$ are (spatially) **hesitating** Markov processes: at some vertex, the next visited point is not chosen deterministically. For a given Hamiltonian graph G which is not reduced to a Hamiltonian cycle, it would be interesting to describe the probability measures π leading to a transition between non-hesitating and hesitating minimizers. This issue remains open at present.

3.2 Differentiation on $\mathcal{L}(\pi)$

This section introduces some elements of differential calculus on $\mathcal{L}(\pi)$, which will be helpful in the proof of Theorem 5. Here we will be working mainly with the complete graph K_V.

We begin by presenting a more analytical expression for the functional F. For $y \in V$, consider the function

$$f_y\ :\ V \ni x \quad \mapsto \quad \frac{\mathbb{1}_{\{y\}}(x)}{\pi(y)} - 1.$$

Note that $\pi[f_y] = 0$, so for any $L \in \mathcal{L}(\pi)$, by irreducibility, there exists a unique function φ_y^L on V satisfying the Poisson equation

$$\left\{ \begin{array}{rcl} L[\varphi_y^L] & = & f_y, \\ \varphi_y^L(y) & = & 0. \end{array} \right. \tag{22}$$

The following relation with the functional F is well-known:

Lemma 7 *For any $L \in \mathcal{L}(\pi)$ and any $x, y \in V$, we have*

$$\varphi_y^L(x) \;=\; \mathbb{E}_x^L[\tau_y],$$

so that

$$F(L) \;=\; \sum_{y \in V} \pi(y)\pi[\varphi_y^L].$$

To simplify notation, from now on, we will remove the L in the exponent of $\mathbb{E}_x^L$ and φ_y^L, when the underlying generator L is clear from the context.

Proof

Let us recall a simple argument, which will be used again in the sequel. Through the martingale problem characterization of X, we have that for any given function φ on V, the process $(M_t)_{t \geqslant 0}$ defined by

$$\forall\ t \geqslant 0, \qquad M_t \;\coloneqq\; \varphi(X_t) - \varphi(X_0) - \int_0^t L[\varphi](X_s)\, ds$$

is a martingale. In particular, for any stopping time τ, the process $(M_{\tau \wedge t})_{t \geqslant 0}$ is also a martingale. Thus, starting from $x \in V$, we get,

$$\begin{array}{rcl} \mathbb{E}_x[M_{\tau_y \wedge t}] & = & 0 \\ & = & \mathbb{E}_x\left[\varphi(X_{\tau \wedge t}) - \varphi(X_0) - \displaystyle\int_0^{\tau \wedge t} L[\varphi](X_s)\, ds\right] \\ & = & \mathbb{E}_x\left[\varphi(X_{\tau \wedge t})\right] - \varphi(x) - \mathbb{E}_x\left[\displaystyle\int_0^{\tau \wedge t} L[\varphi](X_s)\, ds\right]. \end{array}$$

Since τ is a.s. finite and $\varphi(X_{\tau \wedge t}), t \geqslant 0$, uniformly integrable, we obtain, by letting t go to infinity

$$\mathbb{E}_x\left[\varphi(X_\tau)\right] - \varphi(x) - \mathbb{E}_x\left[\int_0^{\tau} L[\varphi](X_s)\, ds\right] \;=\; 0.$$

For any $y \in V$, consider $\varphi \coloneqq \varphi_y$ and $\tau \coloneqq \tau_y$. From (22) and from the fact that $f_y(z) = -1$ for any $z \in V \backslash \{y\}$, we deduce

$$\varphi_y(x) \;=\; \mathbb{E}_x[\tau_y].$$

The last identity of the lemma comes from

$$\begin{array}{rcl} \displaystyle\sum_{x \in V} \pi(x)\mathbb{E}_x[\tau_y] & = & \displaystyle\sum_{x \in V} \pi(x)\varphi_y(x) \\ & = & \pi[\varphi_y]. \end{array}$$

■

Since we are looking for minimizers of F on $\mathcal{L}(\pi)$, it is natural to differentiate this functional. Let $\bar{\mathcal{L}}(\pi)$ be the convex set of normalized Markov generators L admitting π for invariant probability.

The difference with $\mathcal{L}(\pi)$ is that the elements of $\bar{\mathcal{L}}(\pi)$ are not required to be irreducible. For $L \in \mathcal{L}(\pi)$, $\widetilde{L} \in \bar{\mathcal{L}}(\pi)$ and $\epsilon \in [0,1)$, let $L_\epsilon \coloneqq (1-\epsilon)L + \epsilon\widetilde{L} \in \mathcal{L}(\pi)$. Define

$$D_{\widetilde{L}}F(L) \quad \coloneqq \quad \lim_{\epsilon \to 0_+} \frac{F(L_\epsilon) - F(L)}{\epsilon}.$$

In the proof of the following result, it will be shown that this limit exists.

Lemma 8 *With the above notation, we have*

$$\begin{aligned} D_{\widetilde{L}}F(L) &= \sum_{y \in V} \pi(y)(\pi[\varphi_y] - \pi[\psi_y]) \\ &= F(L) - \sum_{y \in V} \pi(y)\pi[\psi_y]. \end{aligned}$$

where ψ_y is the unique solution of another Poisson equation

$$\left\{ \begin{array}{rcl} L[\psi_y] &=& \widetilde{L}[\varphi_y], \\ \psi_y(y) &=& 0 \end{array} \right. . \tag{23}$$

Proof

Let $\mathcal{F}_\pi$ stand for the space of functions f on V whose mean with respect to π vanishes. By restriction to $\mathcal{F}_\pi$, $L \in \mathcal{L}(\pi)$ can be seen as an invertible endomorphism of $\mathcal{F}_\pi$, denote by $L^{-1}_{|\mathcal{F}_\pi}$ its inverse. Similarly, for $\epsilon \in [0,1)$, let $L^{-1}_{\epsilon,|\mathcal{F}_\pi}$ be the inverse of L_ϵ on $\mathcal{F}_\pi$. The mapping $[0,1) \ni \epsilon \mapsto L_\epsilon$ being analytical, the same is true for $[0,1) \ni \epsilon \mapsto L^{-1}_{\epsilon,|\mathcal{F}_\pi}$. Since we have

$$\forall\ \epsilon \in [0,1),\ \forall\ y \in V, \qquad \varphi_y^{L_\epsilon} \quad = \quad L^{-1}_{\epsilon,|\mathcal{F}_\pi}[f_y] - L^{-1}_{\epsilon,|\mathcal{F}_\pi}[f_y](y),$$

we deduce that the mapping

$$[0,1) \ni \epsilon \quad \mapsto \quad \varphi_y^{L_\epsilon}$$

is analytical. The same is true for $[0,1) \ni \epsilon \mapsto F(L_\epsilon)$, due to the equality

$$\forall\ \epsilon \in [0,1), \qquad F(L_\epsilon) \quad = \quad \sum_{y \in V} \pi(y)\pi[\varphi_y^{L_\epsilon}].$$

In particular its derivative $D_{\widetilde{L}}F(L)$ exists and is equal to $\sum_{y \in V} \pi(y)\pi[\varphi'_y]$, where φ'_y is the derivative of $\varphi_y^{L_\epsilon}$ at $\epsilon = 0$. Differentiating the relation $L_\epsilon[\varphi_y^{L_\epsilon}] = f_y$, we get

$$(\widetilde{L} - L)[\varphi_y] + L[\varphi'_y] \quad = \quad 0.$$

Furthermore, we have that $\varphi'_y(y) = \partial_\epsilon \varphi_y^{L_\epsilon}(y)|_{\epsilon=0} = 0$, so that $\varphi_y - \varphi'_y$ satisfies the equation (23) and must be equal to ψ_y. The claim then follows from the equality $\varphi'_y = \varphi_y - \psi_y$, for all $y \in V$.

■

In the above proof, we have seen that $[0,1) \ni \epsilon \mapsto F(L_\epsilon)$ is analytic, so we can differentiate it a second time at $\epsilon = 0$. Denote $D^2_{\widetilde{L}}F(L) = \partial^2_\epsilon F(L_\epsilon)|_{\epsilon=0}$.

Lemma 9 *For $L \in \mathcal{L}(\pi)$, $\widetilde{L} \in \bar{\mathcal{L}}(\pi)$, we have*

$$\begin{aligned} D^2_{\widetilde{L}}F(L) &= \sum_{y \in V} \pi(y)(2\pi[\varphi_y] - 4\pi[\psi_y] + 2\pi[\psi'_y]) \\ &= 4D_{\widetilde{L}}F(L) - 2F(L) + 2\sum_{y \in V} \pi(y)\pi[\psi'_y], \end{aligned}$$

where ψ'_y is the unique solution of

$$\left\{ \begin{array}{rcl} L[\psi'_y] &=& \widetilde{L}[\psi_y], \\ \psi'_y(y) &=& 0 \end{array} \right. . \tag{24}$$

Proof

For any $y \in V$, denote φ''_y the second derivative of $\varphi_y^{L_\epsilon}$ at $\epsilon = 0$. By differentiating twice the relation $L_\epsilon[\varphi_y^{L_\epsilon}] = f_y$ at $\epsilon = 0$, we get

$$(\partial_\epsilon^2 L_\epsilon)[\varphi_y] + 2(\partial_\epsilon L_\epsilon)[\varphi'_y] + L[\varphi''_y] \;=\; 0.$$

namely, since $\partial_\epsilon^2 L_\epsilon = 0$,

$$\begin{aligned} L[\varphi''_y] &= 2(L - \widetilde{L})[\varphi'_y] \\ &= 2(L - \widetilde{L})[\varphi_y - \psi_y] \\ &= 2L[\varphi_y - \psi_y] - 2\widetilde{L}[\varphi_y - \psi_y] \\ &= 2L[\varphi_y - \psi_y] - 2\widetilde{L}[\varphi_y] + 2\widetilde{L}[\psi_y] \\ &= 2L[\varphi_y - \psi_y] - 2L[\psi_y] + 2\widetilde{L}[\psi_y] \\ &= L[2\varphi_y - 4\psi_y] + 2\widetilde{L}[\psi_y]. \end{aligned}$$

It follows that $\varphi''_y/2 - \varphi_y + 2\psi_y$ satisfies the first condition of equation (24). It also vanishes at y, since $\varphi''_y(y) = 0 = \varphi_y(y) = \psi_y(y)$. Thus we get that $\varphi''_y = 2\varphi_y - 4\psi_y + 2\psi'_y$. The announced result is now a consequence of the equality

$$D^2_{\widetilde{L}} F(L) \;=\; \sum_{y \in V} \pi(y)\pi[\varphi''_y].$$

■

It will be convenient to use these differentiations with respect to particular generators $\widetilde{L} \in \bar{\mathcal{L}}(\pi)$. A cycle A in V is a finite sequence $(a_0, a_1, ..., a_{n-1})$ of distinct elements of V, with $n \in \mathbb{N}\backslash\{1\}$ (up to the identification with $(a_k, a_{k+1}, ..., a_{k+n-1})$, for all $k \in \mathbb{Z}_n$). As with Hamiltonian cycles (corresponding to $n = N$), we will make the convention that $a_n = a_0$, as the indices should be seen as elements of $\mathbb{Z}_n$. The set of all cycles is denoted by $\mathcal{A}$. For any $A \in \mathcal{A}$, there is a unique element $L \in \bar{\mathcal{L}}(\pi)$ such that

$$\forall\; x, y \in V, \qquad L(x,y) > 0 \;\;\Leftrightarrow\;\; \exists\; l \in \mathbb{Z}_n \;:\; x = a_l \text{ and } y = a_{l+1}.$$

It is indeed the generator, denoted L_A in the sequel, given by

$$\forall\; x, y \in V, \qquad L_A(x,y) \;=\; \begin{cases} \frac{1}{n\pi(x)} & \text{, if } x = a_l \text{ and } y = a_{l+1} \text{ for some } l \in \mathbb{Z}_n, \\ -\frac{1}{n\pi(x)} & \text{, if } x = y, \\ 0 & \text{, otherwise.} \end{cases} \tag{25}$$

Lemma 10 *Let $A = (a_l)_{l \in \mathbb{Z}_n} \in \mathcal{A}$ be given and for $y \in V$, consider the function ψ_y defined by* (23) *with $\widetilde{L} = L_A$. Then we have*

$$\forall\; x \in V, \qquad \psi_y(x) \;=\; \frac{1}{n} \sum_{l \in \mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l))(\varphi_{a_l}(x) - \varphi_{a_l}(y)). \tag{26}$$

Furthermore, we get that

$$\begin{aligned} \sum_{y \in V} \pi(y)\pi[\psi_y] &= -\frac{1}{n} \sum_{l \in \mathbb{Z}_n} \sum_{y \in V} \pi(y)(\varphi_y(a_{l+1}) - \varphi_y(a_l))\varphi_{a_l}(y) \\ &= \frac{1}{n} \left(\sum_{l \in \mathbb{Z}_n} \frac{1}{2} \mathbb{E}_{a_{l+1}}[\tau_{a_l}^2] - \mathbb{E}_\pi[\tau_{a_l}] \mathbb{E}_{a_{l+1}}[\tau_{a_l}] \right), \end{aligned}$$

where $\mathbb{E}_\pi$ stands for the expectation relative to the initial distribution π for X.

Proof

For any function φ on V, we have

$$L_A[\varphi] \quad = \quad \sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n\pi(a_l)} \mathbb{1}_{\{a_l\}}.$$

Let ψ be a function such that $L[\psi] = L_A[\varphi]$. Using the martingale problem as in the proof of Lemma 8, we get for any $x, y \in V$,

$$\begin{aligned}
\psi(y)-\psi(x) &= \mathbb{E}_x\left[\int_0^{\tau_y} L[\psi](X_s)\,ds\right] \\
&= \mathbb{E}_x\left[\int_0^{\tau_y} L_A[\varphi](X_s)\,ds\right] \\
&= \sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n} \mathbb{E}_x\left[\int_0^{\tau_y} \frac{\mathbb{1}_{\{a_l\}}}{\pi(a_l)}(X_s)\,ds\right] \\
&= \sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n} \mathbb{E}_x\left[\int_0^{\tau_y} 1+f_{a_l}(X_s)\,ds\right] \\
&= \sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n} \left(\varphi_y(x)+\mathbb{E}_x\left[\int_0^{\tau_y} f_{a_l}(X_s)\,ds\right]\right).
\end{aligned}$$

Taking into account that $L[\varphi_{a_l}] = f_{a_l}$, we deduce that

$$\mathbb{E}_x\left[\int_0^{\tau_y} f_{a_l}(X_s)\,ds\right] \quad = \quad \varphi_{a_l}(y)-\varphi_{a_l}(x),$$

so that

$$\psi(y)-\psi(x) \quad = \quad \sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n} \left(\varphi_y(x)+\varphi_{a_l}(y)-\varphi_{a_l}(x)\right).$$

Note that

$$\begin{aligned}
\sum_{l\in\mathbb{Z}_n} \frac{\varphi(a_{l+1})-\varphi(a_l)}{n}\varphi_y(x) &= \frac{\varphi_y(x)}{n}\sum_{l\in\mathbb{Z}_n} \varphi(a_{l+1})-\varphi(a_l) \\
&= 0.
\end{aligned}$$

Thus

$$\psi(y)-\psi(x) \quad = \quad \frac{1}{n}\sum_{l\in\mathbb{Z}_n} \left(\varphi(a_{l+1})-\varphi(a_l)\right)\left(\varphi_{a_l}(y)-\varphi_{a_l}(x)\right).$$

Considering for $y \in V$ the functions $\varphi = \varphi_y$ and $\psi = \psi_y$ and recalling that $\psi_y(y) = 0$, gives the first relation of the lemma. Integrating this relation with respect to π in x, we get

$$\pi[\psi_y] \quad = \quad \frac{1}{n}\sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1})-\varphi_y(a_l))(\pi[\varphi_{a_l}]-\varphi_{a_l}(y)).$$

A well-known result (recall (2) or see e.g. the book of Aldous and Fill [1]) asserts that the quantity $\sum_{y\in V}\pi(y)\varphi_y(x)$ does not depend on $x \in V$. It follows that

$$\sum_{y\in V} \pi(y)(\varphi_y(a_{l+1})-\varphi_y(a_l))\pi[\varphi_{a_l}] \quad = \quad 0 \tag{27}$$

and hence

$$\sum_{y\in V}\pi(y)\pi[\psi_y] \;=\; -\frac{1}{n}\sum_{y\in V}\pi(y)\sum_{l\in\mathbb{Z}_n}(\varphi_y(a_{l+1})-\varphi_y(a_l))\varphi_{a_l}(y),$$

which is the second equality of the lemma. For any $l\in\mathbb{Z}_n$, let ϕ_{a_l} be the function defined by:

$$\forall\ x\in V,\qquad \phi_{a_l}(x) \;=\; \sum_{y\in V}\pi(y)\varphi_{a_l}(y)(\varphi_y(x)-\varphi_y(a_l)). \tag{28}$$

We have $\sum_{y\in V}\pi(y)\pi[\psi_y]=-\frac{1}{n}\sum_{l\in\mathbb{Z}_n}\phi_{a_l}(a_{l+1})$. To compute ϕ_{a_l}, note that $\phi_{a_l}(a_l)=0$ and that

$$\begin{aligned}
L[\phi_{a_l}] &= \sum_{y\in V}\pi(y)\varphi_{a_l}(y)L[\varphi_y]\\
&= \sum_{y\in V}\pi(y)\varphi_{a_l}(y)f_y\\
&= \sum_{y\in V}\pi(y)\varphi_{a_l}(y)\left(\frac{\mathbb{1}_{\{y\}}}{\pi(y)}-1\right)\\
&= \varphi_{a_l}-\pi[\varphi_{a_l}].
\end{aligned}$$

This observation leads us to resort once again to the martingale problem, to get for any $x\in V$,

$$\begin{aligned}
\phi_{a_l}(a_l) &= \phi_{a_l}(x)+\mathbb{E}_x\left[\int_0^{\tau_{a_l}}\varphi_{a_l}(X_s)-\pi[\varphi_{a_l}]\,ds\right]\\
&= \phi_{a_l}(x)+\mathbb{E}_x\left[\int_0^{\tau_{a_l}}\varphi_{a_l}(X_s)\,ds\right]-\pi[\varphi_{a_l}]\mathbb{E}_x[\tau_{a_l}]\\
&= \phi_{a_l}(x)-\pi[\varphi_{a_l}]\varphi_{a_l}(x)+\mathbb{E}_x\left[\int_0^{\tau_{a_l}}\varphi_{a_l}(X_s)\,ds\right]\\
&= \phi_{a_l}(x)-\pi[\varphi_{a_l}]\varphi_{a_l}(x)+\frac{1}{2}\mathbb{E}_x[\tau_{a_l}^2]
\end{aligned}$$

according to Lemma 11 below. Recalling that $\phi_{a_l}(a_l)=0$, we get

$$\phi_{a_l}(x) \;=\; \pi[\varphi_{a_l}]\varphi_{a_l}(x)-\frac{1}{2}\mathbb{E}_x[\tau_{a_l}^2] \tag{29}$$

and this leads immediately to the last equality of the lemma.

■

In the previous proof, we needed the following result.

Lemma 11 *For any $x,y\in V$, we have*

$$\mathbb{E}_x\left[\int_0^{\tau_y}\varphi_y(X_s)\,ds\right] \;=\; \frac{1}{2}\mathbb{E}_x[\tau_y^2].$$

Proof

Coming back to the probabilistic interpretation of φ_y, we get

$$\begin{aligned}
\mathbb{E}_x\left[\int_0^{\tau_y}\varphi_y(X_s)\,ds\right] &= \int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s\leqslant\tau_y\}}\mathbb{E}_{X_s}[\tau_y]\right]\,ds\\
&= \int_0^{+\infty}\int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s\leqslant\tau_y\}}\mathbb{E}_{X_s}[\mathbb{1}_{\{t\leqslant\tau_y\}}]\right]\,ds\,dt\\
&= \int_0^{+\infty}\int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s\leqslant\tau_y\}}\mathbb{E}[\mathbb{1}_{\{t\leqslant\tau_y\circ\theta_s\}}|\sigma(X_u\,:\,u\in[0,s])]\right]\,ds\,dt\\
&= \int_0^{+\infty}\int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s\leqslant\tau_y\}}\mathbb{1}_{\{t\leqslant\tau_y\circ\theta_s\}}\right]\,ds\,dt\\
&= \int_0^{+\infty}\int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s+t\leqslant\tau_y\}}\right]\,ds\,dt,
\end{aligned}$$

where we used the Markov property and where θ_s is the shift by time s of the trajectories of X. Using the Fubini theorem, we get

$$\begin{aligned}\int_0^{+\infty}\int_0^{+\infty}\mathbb{E}_x\left[\mathbb{1}_{\{s+t\leqslant\tau_y\}}\right]ds\,dt &= \mathbb{E}_x\left[\int_0^{+\infty}\int_0^{+\infty}\mathbb{1}_{\{s+t\leqslant\tau_y\}}\,ds\,dt\right]\\ &= \frac{1}{2}\mathbb{E}_x\left[\int_0^{+\infty}\int_0^{+\infty}\mathbb{1}_{\{s\leqslant\tau_y\}}\mathbb{1}_{\{t\leqslant\tau_y\}}\,ds\,dt\right]\\ &= \frac{1}{2}\mathbb{E}_x\left[\left(\int_0^{\tau_y}ds\right)^2\right]\\ &= \frac{1}{2}\mathbb{E}_x[\tau_y^2]\end{aligned}$$

■

Before treating the second derivative in a similar way, let us present two remarks about the quantities entering Lemma 10. We believe they will be relevant for further study of the minimizers of the mapping F on $\mathcal{L}(\pi)$.

Define the following quantities, associated with a given $L\in\mathcal{L}(\pi)$:

$$\begin{aligned}\forall\ x,y\in V,\qquad h_L(x,y) &:= \frac{1}{2}\mathbb{E}_y[\tau_x^2]-\mathbb{E}_\pi[\tau_x]\mathbb{E}_y[\tau_x]\\ \forall\ A=(a_1,...,a_n)\in\mathcal{A},\qquad H_A(L) &= \frac{1}{n}\sum_{l\in\mathbb{Z}_n}h_L(a_l,a_{l+1})\end{aligned}\tag{30}$$

Lemma 10 can be rewritten under the form

$$\forall\ A=(a_1,...,a_n)\in\mathcal{A},\qquad D_AF(L) = F(L)-H_A(L)\tag{31}$$

where $D_AF(L)$ is a short hand for $D_{L_A}F(L)$.

Let us say that a cycle $A=(a_1,...,a_n)\in\mathcal{A}$ is below the generator L, if

$$\forall\ l\in\mathbb{Z}_n,\qquad L(a_l,a_{l+1})>0$$

and denote by $\mathcal{A}(L)$ the set of cycles below L. Then we have:

Lemma 12 *Assume that $L\in\mathcal{L}(\pi)$ is a minimizer of F on $\mathcal{L}(\pi)$. Then,*

$$\begin{aligned}\forall\ A\in\mathcal{A}(L),\qquad H_A(L) &= F(L),\\ \forall\ A\in\mathcal{A}\backslash\mathcal{A}(L),\qquad H_A(L) &\leqslant F(L).\end{aligned}$$

In particular, we get

$$F(L) = \max_{A\in\mathcal{A}}H_A(L).$$

Proof

Consider a minimizer $L\in\mathcal{L}(\pi)$ of F on $\mathcal{L}(\pi)$ and $A\in\mathcal{A}(L)$. Then for $\epsilon\in\mathbb{R}$ small enough, $(1-\epsilon)L+\epsilon L_A$ remains a Markov generator and belongs to $\mathcal{L}(\pi)$. Differentiating $F(L_\epsilon)$ at $\epsilon=0$, we thus get that $D_AF(L)=0$, which implies $H_A(L)=F(L)$. For $A\in\mathcal{A}\backslash\mathcal{A}(L)$, the operator $(1-\epsilon)L+\epsilon L_A$ is not Markovian for $\epsilon<0$. So $D_AF(L)$ only corresponds to the right derivative of $F(L_\epsilon)$ at $\epsilon=0_+$. The minimizing assumption on L implies that $D_AF(L)\geqslant 0$, namely $H_A(L)\leqslant F(L)$. The last identity of the lemma is an immediate consequence of the previous observations and of the fact that there exists at least one cycle below L, by irreducibility.

■

Next we mention a spectral relation satisfied by the quantities $(h_L(x,y))_{x,y\in V}$, reminiscent of (15). Indeed, it is proved in a similar way, as will become clear from the following proof where the arguments for (15) will be recalled.

Lemma 13 *For any $L \in \mathcal{L}(\pi)$, he have*

$$\sum_{x,y\in V} \pi(x)\pi(y)h_L(x,y) \quad = \quad \sum_{\lambda\in\Lambda(L)} \frac{1}{\lambda^2}. \tag{32}$$

Proof

As in the proof of Lemma 7, let $\mathcal{F}_\pi$ stand for the space of functions f on V whose mean with respect to π vanishes and denote by Π the orthogonal projection from $\mathbb{L}^2(\pi)$ to $\mathcal{F}_\pi$:

$$\forall\ f \in \mathbb{L}^2(\pi), \qquad \Pi[f] \quad = \quad f - \pi[f].$$

Let $(g_y)_{y\in V}$ be an orthonormal basis of $\mathbb{L}^2(\pi)$ and R be any endomorphism of $\mathcal{F}_\pi$. We have seen in Lemma 6 of [11] that

$$\mathrm{tr}(R) \quad = \quad \sum_{y\in V} \pi[\Pi[g_y]R[\Pi[g_y]]].$$

In [11], we considered the orthonormal basis given by

$$\forall\ y \in V, \qquad g_y \quad \coloneqq \quad \frac{\mathbb{1}_{\{y\}}}{\sqrt{\pi(y)}}$$

and the operator $L^{-1}_{|\mathcal{F}_\pi}$ defined in Lemma 7, in order to conclude (15), taking into account the fact that $\Pi[g_y] = \sqrt{\pi(y)}f_y$, for all $y \in V$, and that $\mathrm{tr}(L^{-1}_{|\mathcal{F}_\pi}) = \sum_{\lambda\in\Lambda(L)} \frac{1}{\lambda}$.

To prove (32), we use $R = (L^{-1}_{|\mathcal{F}_\pi})^2$. Remark that for any $y \in V$,

$$\begin{aligned}(L^{-1}_{|\mathcal{F}_\pi})^2[f_y] &= L^{-1}_{|\mathcal{F}_\pi}[\varphi_y - \pi[\varphi_y]] \\ &= \phi_y - \pi[\phi_y]\end{aligned}$$

where ϕ_y is the unique solution of

$$\left\{\begin{array}{rcl} L[\phi_y] &=& \varphi_y - \pi[\varphi_y] \\ \phi_y(y) &=& 0. \end{array}\right. \tag{33}$$

(This notation agrees with that introduced in (28)). Thus we get

$$\begin{aligned}\sum_{\lambda\in\Lambda(L)} \frac{1}{\lambda^2} &= \mathrm{tr}((L^{-1}_{|\mathcal{F}_\pi})^2) \\ &= \sum_{y\in V} \pi[\Pi[g_y](L^{-1}_{|\mathcal{F}_\pi})^2[\Pi[g_y]]] \\ &= \sum_{y\in V} \pi(y)\pi[f_y(L^{-1}_{|\mathcal{F}_\pi})^2[f_y]] \\ &= \sum_{y\in V} \pi(y)\pi[f_y(\phi_y - \pi[\phi_y])] \\ &= \sum_{y\in V} \pi(y)\pi[f_y\phi_y] \\ &= \sum_{y\in V} \pi(y)\phi_y(y) - \pi(y)\pi[\phi_y] \\ &= -\sum_{y\in V} \pi(y)\pi[\phi_y] \\ &= -\sum_{x,y\in V} \pi[x]\pi[y]\phi_y(x).\end{aligned}$$

In the proof of Lemma 10 (see (29)), it was shown that

$$\forall\ x, y \in V, \qquad \phi_x(y) \ = \ -h_L(x,y),$$

which leads immediately to (32).

■

Lemma 10 can be extended to the second derivative presented in Lemma 9, by computing similarly the function ψ'_y defined by (24) with $\widetilde{L} = L_A$, for fixed $A = (a_l)_{l\in\mathbb{Z}_n} \in \mathcal{A}$ and $y \in V$. For our purposes, it is convenient to consider a generalization of this situation. Given another cycle $A' = (a'_l)_{l\in\mathbb{Z}_{n'}} \in \mathcal{A}$, consider the equation in the function Ψ_y:

$$\left\{ \begin{array}{rcl} L[\Psi_y] & = & L_{A'}[\psi_y], \\ \Psi_y(y) & = & 0, \end{array} \right. \tag{34}$$

where ψ_y is still associated to L, A and y as in Lemma 10. Of course, when $A' = A$, we recover $\Psi_y = \psi'_y$.

Lemma 14 *For $A = (a_l)_{l\in\mathbb{Z}_n} \in \mathcal{A}$, $A' = (a'_l)_{l\in\mathbb{Z}_{n'}} \in \mathcal{A}$ and $y \in V$ given as above, consider the function Ψ_y defined by (34). Then we have, for any $x \in V$,*

$$\Psi_y(x) \ = \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n, k\in\mathbb{Z}_{n'}} (\varphi_y(a_{l+1}) - \varphi_y(a_l))(\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k))(\varphi_{a'_k}(x) - \varphi_{a'_k}(y)). \tag{35}$$

Furthermore, we get that

$$\begin{aligned} &\sum_{y\in V} \pi(y)\pi[\Psi_y] \\ &= \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n, k\in\mathbb{Z}_{n'}} (h_L(a'_k, a_{l+1}) - h_L(a'_k, a_l))(\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k)) \qquad (36) \\ &= \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n, k\in\mathbb{Z}_{n'}} \left(\frac{1}{2}(\mathbb{E}_{a_{l+1}}[\tau^2_{a'_k}] - \mathbb{E}_{a_l}[\tau^2_{a'_k}]) - \mathbb{E}_\pi[\tau_{a'_k}](\mathbb{E}_{a_{l+1}}[\tau_{a'_k}] - \mathbb{E}_{a_l}[\tau_{a'_k}]). \right) \left(\mathbb{E}_{a'_{k+1}}[\tau_{a_l}] - \mathbb{E}_{a'_k}[\tau_{a_l}] \right) \end{aligned}$$

Proof

From Lemma 10, we have

$$\begin{aligned} L_{A'}[\psi_y] \ &= \ \frac{1}{n} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) L_{A'}[\varphi_{a_l}] \\ &= \ \frac{1}{n} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) \sum_{k\in\mathbb{Z}_{n'}} \frac{\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k)}{n'\pi(a'_k)} \mathbb{1}_{\{a'_k\}} \\ &= \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) \sum_{k\in\mathbb{Z}_{n'}} (\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k)) \frac{\mathbb{1}_{\{a'_k\}}}{\pi(a'_k)} \\ &= \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) \sum_{k\in\mathbb{Z}_{n'}} (\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k)) f_{a'_k}, \end{aligned}$$

where we used that for any $l \in \mathbb{Z}_n$,

$$\sum_{k\in\mathbb{Z}_{n'}} \varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k) \ = \ 0.$$

Thus, denoting

$$\xi_y \ \coloneqq \ \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) \sum_{k\in\mathbb{Z}_{n'}} (\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k))(\varphi_{a'_k} - \varphi_{a'_k}(y)),$$

we get that $L[\xi_y] = L_{A'}[\psi_y]$ and $\xi_y(y) = 0$. It follows that $\Psi_y = \xi_y$, as announced.

We deduce that

$$\pi[\psi'_y] \;=\; \frac{1}{nn'} \sum_{l\in\mathbb{Z}_n} (\varphi_y(a_{l+1}) - \varphi_y(a_l)) \sum_{k\in\mathbb{Z}_{n'}} (\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k))(\pi[\varphi_{a'_k}] - \varphi_{a'_k}(y))$$

and

$$\begin{aligned}\sum_{y\in V} \pi(y)\pi[\psi'_y] &= \frac{1}{nn'} \sum_{y\in V} \pi(y) \sum_{l\in\mathbb{Z}_n, k\in\mathbb{Z}_{n'}} (\varphi_y(a_{l+1}) - \varphi_y(a_l))(\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k))(\pi[\varphi_{a'_k}] - \varphi_{a'_k}(y)) \\ &= -\frac{1}{nn'} \sum_{y\in V} \sum_{l\in\mathbb{Z}_n, k\in\mathbb{Z}_{n'}} \pi(y)(\varphi_y(a_{l+1}) - \varphi_y(a_l))(\varphi_{a_l}(a'_{k+1}) - \varphi_{a_l}(a'_k))\varphi_{a'_k}(y), \qquad (37)\end{aligned}$$

where we used again (recall (27)) that

$$\sum_{y\in V} \pi(y)(\varphi_y(a_{l+1}) - \varphi_y(a_l)) \;=\; 0.$$

Remember also (cf. (28)) that

$$\begin{aligned}\forall\; x \in V, \qquad \sum_{y\in V} \pi(y)(\varphi_y(x) - \varphi_y(a'_k))\varphi_{a'_k}(y) &= \phi_{a'_k}(x) \\ &= -h_L(a'_k, x).\end{aligned}$$

Thus substituting in (37)

$$\varphi_y(a_{l+1}) - \varphi_y(a_l) \;=\; \varphi_y(a_{l+1}) - \varphi_y(a'_k) - (\varphi_y(a_l) - \varphi_y(a'_k))$$

we deduce (36). The last equality of the lemma is obtained by expressing h_L and φ_x, for $x \in V$, in terms of expectation of hitting times.

■

Denote by $H_{A',A}(L)$ the expression given by (36). Considering the case $A' = A$, Lemma 9 leads to

$$D_{A,A}F(L) \;=\; 2F(L) + 4H_A(L) + 2H_{A,A}(L)$$

where $D_{A,A}F(L)$ is a shorthand for $D^2_{L_A}F(L)$. But the importance of Lemma 14, is because, if we define for any $A, A' \in \mathcal{A}$, $D_{A',A}F(L) \coloneqq D_{A'}(D_AF(L))$, then we get

$$D_{A',A}F(L) \;=\; 2(F(L) + H_A(L) + H_{A'}(L) + H_{A',A}(L)).$$

The previous expressions for the differentiations up to order 2 with respect to Markov generators associated to cycles can be extended to general Markov generators from $\bar{\mathcal{L}}(\pi)$. To go in this direction, we need to recall a simple result:

Lemma 15 *The extremal points of the convex set $\bar{\mathcal{L}}(\pi)$ are exactly the generators L_A for $A \in \mathcal{A}$.*

As a consequence, any $L \in \bar{\mathcal{L}}(\pi)$ can be decomposed into a barycentric sum

$$L \;=\; \sum_{A\in\mathcal{A}} p(A)L_A,$$

where p is a probability measure on $\mathcal{A}$. For an extensive discussion of such decompositions, see the book of Kalpazidou [7]. Note that the above decomposition is not unique in general, because $\bar{\mathcal{L}}(\pi)$ is not a simplex for $N \geqslant 3$. For instance, the generator

$$L \;\coloneqq\; \frac{1}{2}\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

of the simple random walk on $\mathbb{Z}_3$ can be written in the form $L = \frac{1}{2}L_{(0,1,2)} + \frac{1}{2}L_{(0,2,1)}$ and $L = \frac{1}{3}L_{(0,1)} + \frac{1}{3}L_{(1,2)} + \frac{1}{3}L_{(2,0)}$.

Nevertheless, given $\widetilde{L}, \widehat{L} \in \bar{\mathcal{L}}(\pi)$, decompose them into

$$\begin{aligned} \widetilde{L} &= \sum_{A \in \mathcal{A}} \widetilde{p}(A) L_A, \\ \widehat{L} &= \sum_{A \in \mathcal{A}} \widehat{p}(A) L_A, \end{aligned}$$

where $\widetilde{p}, \widehat{p}$ are probability measures on $\mathcal{A}$. Then we get for any $L \in \mathcal{L}(\pi)$.

$$\begin{aligned} D_{\widetilde{L}} F(L) &= \sum_{A \in \mathcal{A}} \widetilde{p}(A) D_A F(L), \\ D_{\widehat{L}} D_{\widetilde{L}} F(L) &= \sum_{A, A' \in \mathcal{A}} \widetilde{p}(A) \widehat{p}(A') D_{A,A'} F(L). \end{aligned}$$

It follows that we can write

$$\begin{aligned} D_{\widetilde{L}} F(L) &= F(L) - H_{\widetilde{L}}(L), \\ D_{\widehat{L}} D_{\widetilde{L}} F(L) &= 2(F(L) + H_{\widetilde{L}}(L) + H_{\widehat{L}}(L) + H_{\widetilde{L},\widehat{L}}(L)), \end{aligned}$$

where

$$\begin{aligned} H_{\widetilde{L}}(L) &= \sum_{x \neq y} \pi(x) L(x,y) h_L(x,y) \\ H_{\widehat{L},\widetilde{L}}(L) &= \sum_{x \neq y,\, x' \neq y'} \pi(x') \widehat{L}(x',y') \pi(x) \widetilde{L}(x,y) (h(x',y) - h(x',x))(\varphi_x(y') - \varphi_x(x')) \end{aligned}$$

(definitions which conform to (30) and (36) when $\widetilde{L} = L_A$ and $\widehat{L} = L_{A'}$).

In view of (26) and (35), the following quantity seems to play an important role in bounding the derivatives:

$$\begin{aligned} M(L) &\coloneqq \max_{y,x,x' \in V} \left| \varphi_y(x) - \varphi_y(x') \right| \\ &= \max_{y,x \in V} \varphi_y(x). \end{aligned}$$

Proposition 16 *We have for any $L \in \mathcal{L}(\pi)$ and $\widetilde{L}$, $\widehat{L} \in \bar{\mathcal{L}}(\pi)$,*

$$\begin{aligned} F(L) &\leqslant M(L), \\ \left| D_{\widetilde{L}} F(L) \right| &\leqslant M(L) + M(L)^2, \\ \left| D_{\widehat{L}} D_{\widetilde{L}} F(L) \right| &\leqslant 2(M(L) + M(L)^2 + M(L)^3). \end{aligned}$$

Proof

The first bound is obvious. For the second, note that (26) can be extended to the solution of (23) for general $\widetilde{L} \in \bar{\mathcal{L}}(\pi)$: we get

$$\forall\ y, x \in V, \qquad \psi_y(x) = \sum_{z \neq z' \in V} \pi(z) \widetilde{L}(z,z') (\varphi_y(z') - \varphi_y(z))(\varphi_z(x) - \varphi_z(y)).$$

Taking into account the renormalization of $\widetilde{L}$, it follows that for any $y \in V$, we have for the supremum norm:

$$\|\psi_y\|_\infty \leqslant M(L)^2$$

For the third bound of the lemma, note that (35) can also be extended to Ψ_y for given $y \in V$, which is the solution of

$$\left\{ \begin{array}{rcl} L[\Psi_y] & = & \widehat{L}[\psi_y] \\ \Psi_y(y) & = & 0, \end{array} \right.$$

where ψ_y is the solution of (23). It follows that for any $x, y \in V$,

$$\Psi_y(x) \;=\; \sum_{u \neq v,\, u' \neq v'} \pi(u') \widehat{L}(u', v') \pi(u) \widetilde{L}(u, v) (\varphi_y(v) - \varphi_y(u))(\varphi_u(v') - \varphi_u(u'))(\varphi_{u'}(x) - \varphi_{u'}(y)).$$

Thus

$$\|\Psi_y\|_\infty \;\leqslant\; M(L)^3.$$

■

A natural question is how to upper bound $M(L)$. A first answer is to use the operator norm $\|\!| \cdot |\!\|_{\infty \to \infty}$ from $\mathbb{L}^\infty(\pi)$ to $\mathbb{L}^\infty(\pi)$ with the operator $L^{-1}_{|\mathcal{F}_\pi}$ introduced in Lemma 7:

$$\begin{array}{rcl} M(L) & = & \max\limits_{y \in V} \|\varphi_y\|_\infty \\ & \leqslant & \max\limits_{y \in V} \|\varphi_y - \pi[\varphi_y]\|_\infty + \max\limits_{y \in V} \pi[\varphi_y] \\ & \leqslant & \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{\infty \to \infty} \max\limits_{y \in V} \|f_y\|_\infty + \max\limits_{y \in V} (1 - \pi(y)) M(L) \\ & \leqslant & \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{\infty \to \infty} \dfrac{1}{\pi_\wedge} + (1 - \pi_\wedge) M(L) \end{array}$$

where $\pi_\wedge \coloneqq \min_{x \in V} \pi(x)$. It follows that

$$M(L) \;\leqslant\; \frac{\|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{\infty \to \infty}}{\pi_\wedge^2}.$$

But the norm $\|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{\infty \to \infty}$ does not seem so easy to evaluate. One can instead resort to the operator norm from $\mathbb{L}^2(\pi)$ to $\mathbb{L}^2(\pi)$ as follows. Denoting I the identity operator on $\mathcal{F}_\pi$, we have as above

$$\begin{array}{rcl} M(L) & \leqslant & \|\!| I |\!\|_{2 \to \infty} \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2} \max\limits_{y \in V} \|f_y\|_2 + (1 - \pi_\wedge) M(L) \\ & \leqslant & \dfrac{1}{\sqrt{\pi_\wedge}} \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2} \max\limits_{y \in V} \sqrt{\dfrac{1}{\pi(y)} + 1 - \pi(y)} + (1 - \pi_\wedge) M(L) \\ & \leqslant & \dfrac{1}{\sqrt{\pi_\wedge}} \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2} \sqrt{1 + \dfrac{1}{\pi_\wedge}} + (1 - \pi_\wedge) M(L) \\ & \leqslant & \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2} \dfrac{\sqrt{2}}{\pi_\wedge} + (1 - \pi_\wedge) M(L). \end{array}$$

As a consequence, we get

$$M(L) \;\leqslant\; \frac{\sqrt{2} \|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2}}{\pi_\wedge^2}.$$

This expression is advantageous when L is reversible with respect to π, since in this situation, $\|\!| L^{-1}_{|\mathcal{F}_\pi} |\!\|_{2 \to 2} = 1/\lambda$, where λ is the spectral gap of L, namely the smallest element of $\Lambda(L)$ (which is then in $(0, +\infty)$). Nevertheless, since we are interested in $F(L)$, note there is a simple comparison:

$$M(L) \;\leqslant\; \frac{F(L)}{\pi_\wedge^2}. \tag{38}$$

We now concentrate on the case $\pi = \upsilon$, the uniform measure and $L = L_A$, with A a Hamiltonian cycle. The following result will be crucial in the proof of Theorem 5.

Proposition 17 *For any $A \in \mathcal{H}$ and $\widetilde{A} \in \mathcal{A}\backslash\{A\}$, we have on $\mathcal{L}(\upsilon)$,*

$$D_{\widetilde{A}} F(L_A) \;\geqslant\; \frac{N-1}{2N}.$$

Proof

There is no loss of generality in assuming that $V = \mathbb{Z}_N$ and that $A = (0, 1, 2, ..., N-1)$. To simplify the notation, let us write $L = L_A$. By invariance of L and υ through the rotations $\mathbb{Z}_N \ni x \mapsto x + y \in \mathbb{Z}_N$ for any fixed $y \in \mathbb{Z}_N$, it follows that the quantity $\mathbb{E}_\upsilon[\tau_x]$ does not depend on the choice of $x \in \mathbb{Z}_N$. It is then necessarily equal to $F(L)$. Furthermore, since under L, the Markov process waits an exponential time before adding 1 to the current state, we get that for any $x, y \in \mathbb{Z}_N$, $\mathbb{E}_x[\tau_y] = \rho(x, y)$, where

$$\forall\; x, y \in \mathbb{Z}_N, \qquad \rho(x, y) \;\coloneqq\; \min\{n \in \mathbb{Z}_+ \,:\, y = x + n\}.$$

It follows easily that $F(L) = (N-1)/2$ (for an alternative proof, see Corollary 20 in the next section). Thus we get that

$$\begin{aligned} \forall\; x, y \in \mathbb{Z}_N, \qquad h_L(x, y) &= \frac{1}{2}\left(\mathbb{E}_y[\tau_x^2] - (N-1)\mathbb{E}_y[\tau_x]\right) \\ &= \frac{1}{2}\left(\mathbb{E}_y[\tau_x^2] - (N-1)\rho(y, x)\right). \end{aligned}$$

Since under $\mathbb{P}_y$, τ_x is a sum of $\rho(y, x)$ independent exponential random variables of parameter 1, we compute that

$$\mathbb{E}_y[\tau_x^2] \;=\; \mathbb{E}_y[\tau_x]^2 + \rho(y, x)$$

(because for any exponential random variable $\mathcal{E}$, we have $\mathbb{E}[\mathcal{E}^2] = 2\mathbb{E}[\mathcal{E}]^2$). Thus we get that for any $x, y \in \mathbb{Z}_N$, $h_L(x, y) = h_N(\rho(y, x))$, where

$$h_N \,:\, [0, N-1] \ni r \;\mapsto\; \frac{1}{2}\left(r^2 - (N-2)r\right).$$

This function h_N is decreasing on $[0, (N-2)/2]$, increasing on $[(N-2)/2, N-1]$ and we have $h_N(0) = 0 < (N-1)/2 = h_N(N-1)$.

Thus from the definition (30), we get that

$$\begin{aligned} \forall\; \widetilde{A} \in \mathcal{A}, \qquad H_{\widetilde{A}}(L) &\leqslant h_N(N-1) \\ &= H_A(L). \end{aligned}$$

More precisely, with $\widetilde{A} \coloneqq (a_0, a_1, ..., a_n) \in \mathcal{A}$, we get, except if for any $l \in \mathbb{Z}_n$, $h_N(\rho(a_l, a_{l+1})) = h_N(N-1)$,

$$\begin{aligned} H_{\widetilde{A}}(L) &\leqslant \frac{n-1}{n} h_N(N-1) + \frac{1}{n}\max\{h_N(0), h_N(N-2)\} \\ &= \frac{n-1}{n} H_A(L) \\ &\leqslant \frac{N-1}{N} H_A(L), \end{aligned}$$

where in the equality, we used that $h_N(0) = h_N(N-2) = 0$ and that $h_N(N-1) = H_A(L)$, according to (30). But if for any $l \in \mathbb{Z}_n$, we have $h_N(\rho(a_l, a_{l+1})) = h_N(N-1)$, it means that

$a_{l+1} = a_l + 1$. Since this must be true for all $l \in \mathbb{Z}_n$, it follows that $n = N$ and that $\widetilde{A}$ must be of the form $(k, k+1, ..., k+N-1)$, for some $k \in \mathbb{Z}_N$, namely, it is Hamiltonian.

From (31), we obtain,

$$\begin{aligned} \forall\ \widetilde{A} \in \mathcal{A}\backslash\{A\}, \qquad D_{\widetilde{A}}F(L) &= F(L) - H_{\widetilde{A}}(L) \\ &\geqslant F(L) - \frac{N-1}{N} H_{\widetilde{A}}(L) \\ &= D_A F(L) + \frac{1}{N} H_A(L) \\ &= \frac{N-1}{2N} \end{aligned}$$

due to $D_A F(L) = 0$, because L is not modified by modifying it in the direction of the cycle A. ■

Above we worked with the complete graph K_V and the associated set of Markov generators $\mathcal{L}(\pi)$. But all the previous considerations can be extended to the case of $\mathcal{L}(G, \pi)$, where the graph G is as in the introduction. The only difference is that $\mathcal{A}$ has to be replaced by $\mathcal{A}(G)$, the set of cycles using only edges from E. For instance, Lemma 15 has to be replaced by

Lemma 18 *The extremal points of the convex set $\bar{\mathcal{L}}(G, \pi)$ (the set of normalized Markov generators L, compatible with G and admitting π for invariant probability) are exactly the generators L_A for $A \in \mathcal{A}(G)$.*

3.3 Perturbations of the uniform probability measure

Our main goal here is to show Theorems 5 and 6. Their proofs are respectively based on small and large perturbations of the uniform probability measure υ.

First we check that all Hamiltonian cycles have the same speed in $\mathcal{L}(\pi)$, as was announced in the introduction in the discrete time setting and for the uniform distribution υ, but this is true more generally.

Lemma 19 *Let $A = (a_0, ..., a_{N-1}) \in \mathcal{H}$ be a Hamiltonian cycle, we have*

$$F(L_A) = \frac{N}{2} \sum_{x \neq y} \pi(x)\pi(y).$$

In particular this quantity does not depend on the choice of the Hamiltonian cycle A.

Proof

The generator L_A can be represented by the matrix

$$\begin{pmatrix} -\frac{1}{N\pi(a_0)} & \frac{1}{N\pi(a_0)} & 0 & \cdots & \cdots & 0 \\ 0 & -\frac{1}{N\pi(a_1)} & \frac{1}{N\pi(a_1)} & 0 & \cdots & 0 \\ & & \cdots & \cdots & & \\ 0 & \cdots & \cdots & 0 & -\frac{1}{N\pi(a_{N-2})} & \frac{1}{N\pi(a_{N-2})} \\ \frac{1}{N\pi(a_{N-1})} & 0 & \cdots & \cdots & 0 & -\frac{1}{N\pi(a_{N-1})} \end{pmatrix}.$$

It follows that the polynomial in X given by

$$P(X) \;\coloneqq\; \det \begin{pmatrix} X-\frac{1}{N\pi(a_0)} & \frac{1}{N\pi(a_0)} & 0 & \cdots & \cdots & 0 \\ 0 & X-\frac{1}{N\pi(a_1)} & \frac{1}{N\pi(a_1)} & 0 & \cdots & 0 \\ & & \cdots & \cdots & & \\ 0 & \cdots & \cdots & 0 & X-\frac{1}{N\pi(a_{N-2})} & \frac{1}{N\pi(a_{N-2})} \\ \frac{1}{N\pi(a_{N-1})} & 0 & \cdots & \cdots & 0 & X-\frac{1}{N\pi(a_{N-1})} \end{pmatrix},$$

is equal to $X\prod_{\lambda\in\Lambda(L_A)}(X-\lambda)$. Expanding the latter expression into $X(\alpha_0+\alpha_1 X+\cdots+\alpha_{N-1}X^{N-1})$, we get that

$$\sum_{\lambda\in\Lambda(L_A)}\frac{1}{\lambda} \;=\; -\frac{\alpha_1}{\alpha_0}$$

This is indeed a consequence of

$$\begin{aligned} \alpha_0 &= (-1)^{N-1}\prod_{m\in[\![N-1]\!]}\theta_m \\ \alpha_1 &= (-1)^{N-2}\sum_{k\in[\![N-1]\!]}\prod_{m\in[\![N-1]\!]\setminus\{k\}}\theta_m \end{aligned}$$

where $\Lambda(L_A)$ is parametrized as the multiset consisting of the θ_m, for $m\in[\![N-1]\!]\coloneqq\{1,2,...,N-1\}$.

On another hand, we compute directly from the definition of $P(X)$, by expanding the determinant, that

$$P(X) \;=\; \prod_{l\in\mathbb{Z}_N}\left(X-\frac{1}{N\pi(a_l)}\right)-\prod_{l\in\mathbb{Z}_N}\left(-\frac{1}{N\pi(a_l)}\right)$$

It follows that

$$\begin{aligned} \alpha_0 &= \sum_{k\in\mathbb{Z}_N}\prod_{m\in\mathbb{Z}_N\setminus\{k\}}\left(-\frac{1}{N\pi(a_m)}\right) \\ \alpha_1 &= \frac{1}{2}\sum_{k\neq l\in\mathbb{Z}_N}\prod_{m\in\mathbb{Z}_N\setminus\{k,l\}}\left(-\frac{1}{N\pi(a_m)}\right) \end{aligned}$$

(the factor $1/2$ is due to the fact that the couple (k,l) also appears as (l,k)). Multiplying the numerator and the denominator by $\prod_{m\in\mathbb{Z}_N}(-N\pi(a_m))$, we get that

$$-\frac{\alpha_1}{\alpha_0} \;=\; \frac{\frac{N^2}{2}\sum_{k\neq l\in\mathbb{Z}_N}\pi(a_k)\pi(a_l)}{N\sum_{m\in\mathbb{Z}_N}\pi(a_m)}$$

and this leads to the announced result.

■

In particular, for $\pi=\upsilon$, the uniform probability measure on V, we get:

Corollary 20 *For $\pi=\upsilon$, we have for any Hamiltonian cycle A,*

$$F(L_A) \;=\; \frac{N-1}{2}.$$

The next result is the crucial step in the proof of Theorem 5. For its statement, introduce for any $A \in \mathcal{H}$ and $\epsilon \in (0,1)$,

$$\mathcal{N}_{A,\epsilon} \;\coloneqq\; \{L = (1-t)L_A + t\widetilde{L} \,:\, t \in [0,\epsilon) \text{ and } \widetilde{L} \in \bar{\mathcal{L}}(\pi)\} \tag{39}$$

This set is a neighborhood of L_A in $\mathcal{L}(\pi)$ and observe that we would have ended with the same set if we had required in this definition that $\widetilde{L}$ belong to the convex hull generated by the $L_{\widetilde{A}}$, for $\widetilde{A} \in \mathcal{A}\backslash\{A\}$.

Define

$$\begin{aligned}
\epsilon_1(N,\pi_\wedge) &\coloneqq \pi_\wedge^4 \ln\left(1 + \frac{1}{N\pi_\wedge^2}\right),\\
\epsilon_2(\pi_\wedge) &\coloneqq \frac{1}{56}\pi_\wedge^{12},\\
\epsilon(N,\pi_\wedge) &\coloneqq \epsilon_1(N,\pi_\wedge) \wedge \epsilon_2(\pi_\wedge).
\end{aligned}$$

Lemma 21 *For $N \geqslant 2$ and any $A \in \mathcal{H}$, L_A is the unique minimizer of F over $\mathcal{N}_{A,\epsilon(N,\pi_\wedge)}$.*

Proof

Assume that for some given $A \in \mathcal{H}$, L_A is not the unique minimizer of F over $\mathcal{N}_{A,\epsilon(N,\pi_\wedge)}$. Then we can find $t \in (0, \epsilon(N,\pi_\wedge))$ and a probability p on $\mathcal{A}\backslash\{A\}$, such that $F(L_t) \leqslant F(L_A)$, with

$$\begin{aligned}
L_t &\coloneqq (1-t)L_A + t\widetilde{L},\\
\widetilde{L} &\coloneqq \sum_{\widetilde{A}\in\mathcal{A}\backslash\{A\}} p(\widetilde{A}) L_{\widetilde{A}}.
\end{aligned}$$

Applying Taylor-Lagrange formula to the function $[0,t] \ni s \mapsto F(L_s)$, we get there exists $s \in [0,t]$ such that

$$F(L_t) \;=\; F(L_A) + tD_{\widetilde{L}}F(L_A) + \frac{t^2}{2}D^2_{\widetilde{L}}F(L_s).$$

Taking into account Propositions 16 and 17 and (38), we obtain

$$F(L_t) \;\geqslant\; F(L_A) + t\frac{N-1}{2N} - t^2\left(\frac{F(L_s)}{\pi_\wedge^2} + \frac{F(L_s)^2}{\pi_\wedge^4} + \frac{F(L_s)^3}{\pi_\wedge^6}\right). \tag{40}$$

To evaluate $F(L_s)$, note that for $s \in (0,t)$,

$$\begin{aligned}
\partial_s F(L_s) &= D_{\widetilde{L}}F(L_s)\\
&\leqslant M(L_s) + M(L_s)^2\\
&\leqslant \frac{F(L_s)}{\pi_\wedge^2} + \frac{F(L_s)^2}{\pi_\wedge^4}.
\end{aligned}$$

Classical computations show that if a $\mathcal{C}^1$ function $f : [0,t] \to (0,+\infty)$ satisfies $\partial_s f(s) \leqslant af(s) + bf^2(s)$ for all $s \in [0,t]$, where $a, b > 0$, then assuming $f(0)\exp(bt) < f(0) + b/a$, we get

$$\forall\; s \in [0,t], \qquad f(s) \;\leqslant\; \frac{bf(0)\exp(bt)}{b + af(0)(1-\exp(bt))}.$$

In particular, if

$$\exp(bt) \;<\; 1 + b/(2af(0)), \tag{41}$$

then

$$\forall\ s \in [0,t], \qquad f(s) \ \leqslant\ 2\left(f(0) + \frac{b}{2a}\right).$$

Let us apply this observation with the mapping $[0,t] \ni s \mapsto F(L_s)$ and $a \coloneqq 1/\pi_\wedge^2$, $b = 1/\pi_\wedge^4$. Since

$$\begin{aligned} F(L_0) &= F(L_A) \\ &= \frac{N}{2} \sum_{x \sharp y \in V} \pi(x)\pi(y) \\ &\leqslant \frac{N}{2}, \end{aligned}$$

we get that condition (41) is satisfied, due to the definition of $\epsilon_1(N, \pi_\wedge)$ and to the fact that $t \in (0, \epsilon_1(N, \pi_\wedge))$. It follows that,

$$\begin{aligned} \forall\ s \in [0,t], \qquad F(L_s) &\leqslant N + \frac{1}{\pi_\wedge^2} \\ &\leqslant \frac{2}{\pi_\wedge^2}. \end{aligned}$$

since $\pi_\wedge \leqslant 1/N$. Substituting this bound in (40), we deduce that

$$\begin{aligned} F(L_t) &\geqslant F(L_A) + t\frac{N-1}{2N} - t^2\left(\frac{2}{\pi_\wedge^4} + \frac{4}{\pi_\wedge^8} + \frac{8}{\pi_\wedge^{12}}\right) \\ &\geqslant F(L_A) + t\frac{1}{4} - \frac{14}{\pi_\wedge^{12}}t^2 \end{aligned}$$

The r.h.s. is strictly larger than $F(L_A)$ if $t < \epsilon_2(\pi_\wedge)$ and this is in contradiction with our initial assumption.

■

Denote for any $\pi \in \mathcal{P}_+(V)$,

$$F_\wedge(\pi) \ \coloneqq\ \inf\{F(L)\ :\ L \in \mathcal{L}(\pi)\}. \tag{42}$$

Another ingredient in the proof of Theorem 5 is:

Lemma 22 *The mapping $\mathcal{P}_+(V) \ni \pi \mapsto F_\wedge(\pi)$ is continuous.*

Proof

Let $\mathcal{L}$ be the set of irreducible and normalized Markov generators (so that $\mathcal{L} = \sqcup_{\pi \in \mathcal{P}_+(V)} \mathcal{L}(\pi)$), endowed with the topology inherited from $\mathbb{R}^{V^2}$. The functional F is defined on $\mathcal{L}$ and (15) is valid on $\mathcal{L}$. As a consequence, F is continuous on $\mathcal{L}$. Indeed, if $(L_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathcal{L}$ converging to $L \in \mathcal{L}$, then according to Paragraph 5 of Chapter 2 of Kato [8], we have $\lim_{n\to\infty} \Lambda(L_n) = \Lambda(L)$ and so $\lim_{n\to\infty} F(L_n) = F(L)$. Next consider a sequence $(\pi_n)_{n\in\mathbb{N}}$ of elements from $\mathcal{P}_+(V)$ converging to $\pi \in \mathcal{P}_+(V)$ and such that the sequence $(F_\wedge(\pi_n))_{n\in\mathbb{N}}$ admits a limit. For all $n \in \mathbb{N}$, let L_n be an element from $\mathcal{L}(\pi_n)$ such that

$$F_\wedge(\pi_n) \ \leqslant\ F(L_n) \ \leqslant\ F_\wedge(\pi_n) + \frac{1}{n}$$

Due to the normalization condition and to the belonging of π to $\mathcal{P}_+(V)$, we can extract a subsequence (still denoted $(L_n)_{n\in\mathbb{N}}$ below) from $(L_n)_{n\in\mathbb{N}}$ converging to some generator L. It is clear that L is normalized and that π is invariant for L. Let us check that L is irreducible. Fix $x \in V$.

For any $n \in \mathbb{N}$, let $X^{(n)} \coloneqq (X_t^{(n)})_{t \geqslant 0}$ be a Markov process starting from x and whose generator is L_n. It is not difficult to deduce from the corresponding martingale problems, that $X^{(n)}$ converges in law (with respect to the Skorokhod topology) to a Markov process starting from x and whose generator is L. Thus for any $y \in V$ and $T \geqslant 0$,

$$\lim_{n \to \infty} \mathbb{E}_x[T \wedge \tau_y^{(n)}] \quad = \quad \mathbb{E}_x[T \wedge \tau_y]$$

(with an obvious notation). It follows that

$$\begin{aligned} \mathbb{E}_x[T \wedge \tau_y] &\leqslant \liminf_{n \to \infty} \mathbb{E}_x[\tau_y^{(n)}] \\ &\leqslant \liminf_{n \to \infty} \frac{1}{\pi_n(x)\pi_n(y)} F(L_n) \\ &= \frac{1}{\pi(x)\pi(y)} \liminf_{n \to \infty} F_\wedge(\pi_n) \\ &\leqslant \frac{N}{2\pi(x)\pi(y)} \end{aligned}$$

according to Lemma 19. Letting T go to infinity, we get $\mathbb{E}_x[\tau_y] \leqslant N/(2\pi_\wedge^2)$. This bound, valid for all $x, y \in V$, implies that L is irreducible and thus $L \in \mathcal{L}(\pi)$. Furthermore, the above arguments show that

$$\begin{aligned} \lim_{n \to \infty} F_\wedge(\pi_n) &= \lim_{n \to \infty} F(L_n) \\ &= F(L) \\ &\geqslant F_\wedge(\pi). \end{aligned}$$

So $F_\wedge$ is lower continuous on $\mathcal{P}_+(V)$. By considering the sequence $(\pi_n)_{n \in \mathbb{N}}$ identically equal to π, we also get that the infimum defining $F_\wedge(\pi)$ is attained.

To show that $F_\wedge$ is upper continuous on $\mathcal{P}_+(V)$, let again $(\pi_n)_{n \in \mathbb{N}}$ be a sequence of elements from $\mathcal{P}_+(V)$ converging to some $\pi \in \mathcal{P}_+(V)$ and such that the sequence $(F_\wedge(\pi_n))_{n \in \mathbb{N}}$ admits a limit. According to the previous remark, there exists $L \in \mathcal{L}(\pi)$ such that $F(L) = F_\wedge(\pi)$. For any $n \in \mathbb{N}$, consider the matrix $\widetilde{L}_n$ given by

$$\forall\ x, y \in V, \qquad \widetilde{L}_n(x,y) \quad \coloneqq \quad \frac{\pi(x)}{\pi_n(x)} L(x,y).$$

It is immediate to prove that $\widetilde{L}_n$ is an irreducible Markov generator leaving π_n invariant. But it may not be normalized, so let $\kappa_n > 0$ be such that $L_n \coloneqq \kappa_n \widetilde{L}_n$ belongs to $\mathcal{L}(\pi_n)$. There is no difficulty in checking that L_n converges to L and thus that $\lim_{n \to \infty} F(L_n) = F(L) = F_\wedge(\pi)$. Thus passing into the limit in $F(L_n) \geqslant F_\wedge(\pi_n)$, we deduce that

$$F_\wedge(\pi) \quad \geqslant \quad \lim_{n \to \infty} F_\wedge(\pi_n)$$

as desired.

■

With all these ingredients, we can now come to the

Proof of Theorem 5

Note that it is sufficient to consider the case where G is the complete graph over V, since $F_\wedge(\pi) \leqslant \min\{F(L) : L \in \mathcal{L}(G, \pi)\}$, for any graph G and positive probability measure π on V.

The main argument is by contradiction. Assuming that the statement of Theorem 5 is not true, we can find a sequence $(\pi_n)_{n \in \mathbb{N}}$ converging to υ, such that for all $n \in \mathbb{N}$, there exists $L_n \in \mathcal{L}(\pi_n) \backslash \{L_{\pi_n, A} : A \in \mathcal{H}\}$ with $F(L_n) = F_\wedge(\pi_n)$. (Here we have included π_n in the index of $L_{\pi_n, A}$

to underscore the fact that this generator, associated to a Hamiltonian cycle A, also depends on the underlying invariant probability π_n.) As seen in the proof of Lemma 22, a subsequence (still denoted $(L_n)_{n\in\mathbb{N}}$) converging toward some $L \in \mathcal{L}(\upsilon)$ can be extracted from $(L_n)_{n\in\mathbb{N}}$. We furthermore have

$$\lim_{n\to\infty} F(L_n) \quad = \quad F(L)$$

and by Lemma 22

$$\lim_{n\to\infty} F_\wedge(\pi_n) \quad = \quad F_\wedge(\upsilon).$$

It follows that $F(L) = F_\wedge(\upsilon)$. From Theorem 4, we deduce that there exists $A \in \mathcal{H}$ such that $L = L_{\upsilon,A}$. Using again the fact that

$$\lim_{n\to\infty} \pi_n \quad = \quad \upsilon, \tag{43}$$

we get that $\lim_{n\to\infty} L_{\pi_n,A} = L_{\upsilon,A}$ and thus

$$\lim_{n\to\infty} (L_n - L_{\pi_n,A}) \quad = \quad 0. \tag{44}$$

Consider $r \coloneqq \min_{n\in\mathbb{N}} \pi_{n,\wedge}$, which is positive due to (43), and let $\epsilon \coloneqq \epsilon(N, r)$, with the notation introduced before Lemma 21. From (44), we deduce that for $n \in \mathbb{N}$ large enough, L_n belongs to $\mathcal{N}(\pi_n, A, \epsilon)$, defined as in (39), with π replaced by π_n. Then Lemma 21 asserts that $L_n = L_{\pi_n,A}$, because L_n is a minimizer of L over $\mathcal{L}(\pi_n)$. This is in contradiction with our initial assumption.

■

To finish this section, we consider large perturbations of the uniform probability measure υ.

Proof of Theorem 6

Let $G = (V, E)$ be a finite oriented connected graph which is not a Hamiltonian cycle. Then we can find a cycle $A \coloneqq (a_0, a_1, ..., a_{n-1}) \in \mathcal{A}(G)$ with $n < \mathrm{card}(V)$. Denote $\widetilde{V} \coloneqq \{a_0, a_1, ..., a_{n-1}\}$ and $\widehat{V} \coloneqq V \backslash \widetilde{V}$. By the strong connectivity of G, we can find a subset $\widehat{E}$ of oriented edges from E, such that $\mathrm{card}(\widehat{E}) = \mathrm{card}(\widehat{V})$ and for any $x \in \widehat{V}$ we can find exactly one $y \in V$ with $(x, y) \in \widehat{E}$. Putting together the edges from A and those from $\widehat{E}$, we get a graph $\breve{G}$ on V looking like the following picture, where the cycle is oriented clockwise and the trees are oriented toward the cycle.

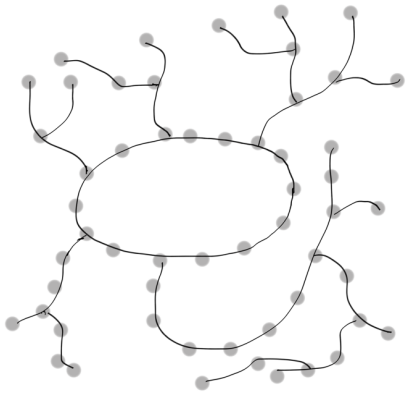

Figure 1: the graph $\breve{G}$

For $r > 0$, consider the Markov generator L_r defined by

$$\forall\ x \neq y \in V, \qquad L_r(x, y) \quad \coloneqq \quad \begin{cases} 1 & \text{, if there exists } l \in [\![0, n-1]\!] \text{ such that } x = a_l \text{ and } y = a_{l+1} \\ r & \text{, if } (x, y) \in \widehat{E} \\ 0 & \text{, otherwise.} \end{cases}$$

This generator is not irreducible, since it does not allow the chain to go from the cycle A to $\widehat{V}$. Nevertheless, its unique invariant probability measure π is the uniform probability measure on $\widetilde{V}$. The generator L_r then satisfies an **extended normalization condition**, in the sense that

$$\sum_{x \neq y \in V} \pi(x) L_r(x,y) \quad = \quad 1$$

The interest in L_r is because it is easy to find its eigenvalues:

$$\Lambda(L_r) \quad = \quad \Lambda(\widetilde{L}_A) \sqcup \{r[|\widehat{V}|]\}$$

where $\widetilde{L}_A$ is the generator corresponding to the Hamiltonian cycle given by A on $\widetilde{V}$ and $\{r[|\widehat{V}|]\}$ is the multiset consisting of the value r with the multiplicity $|\widehat{V}|$. This identity is an immediate consequence of following decomposition of L_r, where all the elements of $\widetilde{V}$ have been put before those of $\widehat{V}$ and where the elements of $\widehat{V}$ have been ordered so that the (oriented) distance to $\widetilde{V}$ is non-decreasing (in particular the last element corresponds to a leaf of $\breve{G}$):

$$L_r \quad = \quad \begin{pmatrix} \widetilde{L}_A & 0 \\ C & D \end{pmatrix}.$$

In the r.h.s., the $\widehat{V} \times \widehat{V}$ matrix D is sub-diagonal and its diagonal consists only of $-r$. Formula (15) enables to extend the functional F to L_r and we get

$$F(L_r) \quad = \quad F(\widetilde{L}_A) + \frac{|V|}{r}.$$

In particular, it follows that

$$\lim_{r \to +\infty} F(L_r) \;=\; F(\widetilde{L}_A) \;=\; \frac{n-1}{2} \;<\; F(L_H)$$

for any Hamiltonian cycle $H \in \mathcal{H}(G)$, where we used twice Corollary 20. From now on, we fix $r > 0$ large enough, so that

$$F(L_r) \quad < \quad F(L_H) \tag{45}$$

for any Hamiltonian cycle $H \in \mathcal{H}(G)$.

For any $\epsilon > 0$, consider the Markov generator

$$L_{r,\epsilon} \quad \coloneqq \quad Z_{r,\epsilon}^{-1}(L_r + \epsilon L_G)$$

where

• the Markov generator L_G is defined by

$$\forall\; x \neq y \in V, \qquad L_G(x,y) \quad \coloneqq \quad \left\{ \begin{array}{ll} 1 & \text{, if } (x,y) \in E \\ 0 & \text{, otherwise} \end{array} \right.$$

• the constant $Z_{r,\epsilon} > 0$ is such that $L_{r,\epsilon}$ is normalized (this is possible because $L_r + \epsilon L_G$ is irreducible on V).

For $r, \epsilon > 0$, denote $\pi_{r,\epsilon}$ the invariant probability measure of $L_{r,\epsilon}$. It is clear that as ϵ goes to 0_+, $\pi_{r,\epsilon}$ converges toward π. It follows that

$$\begin{aligned} \lim_{\epsilon \to 0_+} Z_{r,\epsilon} &= 1 \\ \lim_{\epsilon \to 0_+} L_{r,\epsilon} &= L_r \end{aligned}$$

From the general theory of perturbation of spectra of finite operators (see e.g. the beginning of the second chapter of the book of Kato [8]), we have

$$\lim_{\epsilon \to 0_+} F(L_{r,\epsilon}) \quad = \quad F(L_r).$$

Taking into account (45), we can thus find $\epsilon > 0$ small enough so that

$$F(L_{r,\epsilon}) \quad < \quad F(L_H)$$

for any Hamiltonian cycle $H \in \mathcal{H}(G)$. Namely the probability measure $\pi_{r,\epsilon}$ satisfies the statement of Theorem 6. One would have remarked that this probability measure $\pi_{r,\epsilon}$ is quite far away from υ, because it gives very small weight to the elements of $\widehat{V}$.

■

4 The discrete time framework

Here we discuss the links between the search of the fastest continuous-time Markov processes with the analogous problem in discrete time.

Let a graph $G = (V, E)$ and a positive probability measure π on V be fixed and denote by $\mathcal{K}(G, \pi)$ the set of irreducible Markov kernels K on V whose permitted transitions are edges from E (plus self-loops, i.e., the possibility to stay at the same place) and leaving π invariant, namely satisfying $\pi K = \pi$. For any $K \in \mathcal{K}(G, \pi)$, let $\mathfrak{X} \coloneqq (\mathfrak{X}_n)_{n \in \mathbb{Z}_+}$ be a Markov chain whose transitions are dictated by K. For any $y \in V$, recall (see (1)) that

$$\tau_y \quad \coloneqq \quad \inf\{n \in \mathbb{Z}_+ \, : \, \mathfrak{X}_n = y\}.$$

On $\mathcal{K}(G, \pi)$, we consider the functional $\mathfrak{F}$ defined by

$$\forall \; K \in \mathcal{K}(G, \pi), \qquad \mathfrak{F}(K) \quad \coloneqq \quad \sum_{x,y \in V} \pi(x) \pi(y) \mathbb{E}_x[\tau_y],$$

where subscript x in the expectation indicates that $\mathfrak{X}$ is starting from $x \in V$.

To any $K \in \mathcal{K}(G, \pi)$, we associate $\Theta(K)$ the multiset consisting of the spectrum of K, removing the eigenvalue 1 (of multiplicity 1). It is a priori a sub(multi)set of the closed unit disk centered at 0 of $\mathbb{C}$ left invariant by conjugation. Analogously to the continuous-time situation, we have the eigentime relation

$$\forall \; K \in \mathcal{K}(G, \pi), \qquad \mathfrak{F}(K) \quad = \quad \sum_{\theta \in \Theta(K)} \frac{1}{1 - \theta}.$$

To any $L \in \mathcal{L}(G, \pi)$, associate

$$\begin{aligned} l \quad &\coloneqq \quad \max\{L(x) \, : \, x \in V\}, \\ K \quad &\coloneqq \quad I + \frac{L}{l}. \end{aligned}$$

It is immediate to check that $K \in \mathcal{K}(G, \pi)$. Furthermore, we have $\Theta(K) = 1 - \Lambda(L)/l$, so that

$$\mathfrak{F}(K) \quad = \quad l F(L). \tag{46}$$

Taking into account that

$$l \; \geqslant \; \sum_{x \in V} \pi(x) L(x) \; = \; 1,$$

it follows that $\mathfrak{F}(K) \geqslant F(L)$. We will denote $\Phi : \mathcal{K}(G,\pi) \to \mathcal{L}(G,\pi)$ the mapping $L \to K$ defined above.

Conversely, to any $K \in \mathcal{K}(G,\pi)$, associate

$$\begin{array}{rcl} k & \coloneqq & \dfrac{1}{\sum_{x\in V} \pi(x)(1-K(x,x))}, \\ L & = & k(K-I). \end{array}$$

It is immediate to check that $L \in \mathcal{L}(G,\pi)$. Furthermore, we get $\Lambda(L) = k(1-\Theta(K))$ and it follows that

$$F(L) \;=\; \mathfrak{F}(K)/k.$$

Taking into account that

$$k \;\geqslant\; \frac{1}{\sum_{x\in V}\pi(x)} \;= 1,$$

we get that $F(L) \leqslant \mathfrak{F}(K)$. Denote $\Psi : \mathcal{L}(G,\pi) \to \mathcal{K}(G,\pi)$ the mapping $L \to K$ as above.

Remark 23 The mappings Φ and Ψ are not inverse of each other, because the image of $\mathcal{L}(G,\pi)$ by Φ is included into $\mathcal{K}_0(G,\pi) \coloneqq \{K \in \mathcal{K}(G,\pi) \,:\, \exists\, x \in V \text{ with } K(x,x) = 0\}$. Nevertheless, we have that Φ and Ψ_0 are inverse of each other, where Ψ_0 is the restriction of Ψ to $\mathcal{K}_0(G,\pi)$.

When one is looking for the minimal value of $\mathfrak{F}$ on $\mathcal{K}(G,\pi)$, one can restrict attention to $\mathcal{K}_0(G,\pi)$, because

$$\min\{\mathfrak{F}(K) \,:\, K \in \mathcal{K}(G,\pi)\} \;=\; \min\{\mathfrak{F}(K) \,:\, K \in \mathcal{K}_0(G,\pi)\}.$$

Indeed, for any $K \in \mathcal{K}(G,\pi)$, there exist a unique $\widetilde{K} \in \mathcal{K}_0(G,\pi)$ and $\alpha \in [0,1)$ such that $K = (1-\alpha)\widetilde{K} + \alpha I$. Then we get $\Theta(K) = (1-\alpha)\Theta(\widetilde{K}) + \alpha$, i.e. $\Theta - 1 = (1-\alpha)(\Theta(\widetilde{K}) - 1)$. This implies that

$$\begin{array}{rcl} \mathfrak{F}(\widetilde{K}) & = & (1-\alpha)\mathfrak{F}(K) \\ & \leqslant & \mathfrak{F}(K). \end{array}$$

□

As in (42), denote

$$\begin{array}{rcl} F_\wedge(G,\pi) & \coloneqq & \inf\{F(L) \,:\, L \in \mathcal{L}(G,\pi)\}, \\ \mathfrak{F}_\wedge(G,\pi) & \coloneqq & \inf\{\mathfrak{F}(K) \,:\, K \in \mathcal{K}(G,\pi)\}. \end{array}$$

From the above considerations, we deduce:

Proposition 24 *We always have*

$$F_\wedge(G,\pi) \;\leqslant\; \mathfrak{F}_\wedge(G,\pi)$$

(in particular, when looking for the fastest Markov chain in the sense we have defined, it is preferable to resort to continuous time rather than to discrete time).

Furthermore, assume that there is a minimizer $L \in \mathcal{L}(G,\pi)$ of F such that $L(x)$ does not depend on $x \in V$ (it is then equal to 1). Then $F_\wedge(G,\pi) = \mathfrak{F}_\wedge(G,\pi)$.

Proof

Consider $K \in \mathcal{K}(G,\pi)$. We have seen that

$$\begin{aligned} \mathfrak{F}(K) &\geqslant F(\Psi(K)) \\ &\geqslant F_\wedge(G,\pi), \end{aligned}$$

so taking the infimum over $K \in \mathcal{K}(G,\pi)$, we get the first bound.

Conversely, if $L \in \mathcal{L}(G,\pi)$ is a minimizer of F whose diagonal is constant, then $l = 1$ in (46), namely $\mathfrak{F}(\Phi(L)) = F(L) = F_\wedge(G,\pi)$. From the previous inequality, it follows that $\Phi(L)$ is indeed a minimizer of $\mathfrak{F}$ on $\mathcal{K}(G,\pi)$ and we conclude that $F_\wedge(G,\pi) = \mathfrak{F}_\wedge(G,\pi)$.

■

In association with Theorem 4, the above proposition also enables us to recover the result of Litvak and Ejov [10] stating that for any Hamiltonian graph G, the permutation matrices associated to the Hamiltonian cycles of G are the unique minimizers of $\mathfrak{F}$ on $\mathcal{K}(G,\upsilon)$. But Proposition 24 does not enable us to extend directly Theorem 6 to the discrete time setting, because the diagonal of the generator associated to a Hamiltonian cycle is constant if and only if the underlying invariant probability measure is uniform. This extension is nevertheless true. To show it, note that the differentiation technique of Section 3 can be adapted to $\mathcal{K}(G,\pi)$ in a straightforward manner.

A APPENDIX: Computations on the simplest example of non-Hamiltonian connected graph

The length 2 segment $S_2 \coloneqq (\{0,1,2\}, \{(0,1),(1,0),(1,2),(2,1)\})$ is the simplest non-Hamiltonian (strongly) connected graph. We compute here the minimizer of F on $\mathcal{L}(S_2,\pi)$, for any positive probability measure π on $\{0,1,2\}$. We hope this example will motivate further investigation of the minimizers of F in the challenging non-Hamiltonian framework.

To simplify the notation, write $x = \pi(0), y = \pi(1)$ and $z = \pi(2)$, by assumption we have that $x,y,z > 0$ and $x+y+z = 1$. Up to exchanging the vertices 0 and 2, we assume that $|x - 1/2| \geqslant |z - 1/2|$.

Any Markov generator L from $\mathcal{L}(S_2,\pi)$ has the form

$$L \coloneqq \begin{pmatrix} -a & a & 0 \\ \alpha & -\alpha-\beta & \beta \\ 0 & b & -b \end{pmatrix}$$

where the coefficients $a,\alpha,\beta,b > 0$ satisfy,

$$\begin{aligned} xa &= y\alpha, \\ y\beta &= zb, \\ xa + y(\alpha+\beta) + zb &= 1. \end{aligned}$$

The first two equalities correspond to the invariance of π for L (here π is even reversible for the birth and death generator L) and the third one is the normalization condition, it can be rewritten

$$2xa + 2zb = 1 \tag{47}$$

Denote $\Lambda(L) = \{\lambda_1, \lambda_2\}$, its elements are the non-zero roots in X of the polynomial $\det(X+L)$. We compute that

$$\det(X+L) = X(X^2 - (a+\alpha+\beta+b)X + ab + a\beta + \alpha b),$$

so that

$$\begin{aligned} \lambda_1 + \lambda_2 &= a + \alpha + \beta + b, \\ \lambda_1 \lambda_2 &= ab + a\beta + \alpha b. \end{aligned}$$

From (15), we have

$$\begin{aligned} F(L) &= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} \\ &= \frac{\lambda_1 + \lambda_2}{\lambda_1 \lambda_2} \\ &= \frac{a + \alpha + \beta + b}{ab + a\beta + \alpha b} \\ &= \frac{a\left(1 + \frac{x}{y}\right) + b\left(1 + \frac{z}{y}\right)}{ab\left(1 + \frac{x}{y} + \frac{z}{y}\right)} \\ &= \frac{a(x+y) + b(y+z)}{ab} \\ &= \frac{x+y}{b} + \frac{y+z}{a} \\ &= \frac{1-z}{b} + \frac{1-x}{a}. \end{aligned}$$

Taking into account (47), the minimizer of F on $\mathcal{L}(S_2, \pi)$ corresponds to the minimizer of

$$(0, 1/(2x)) \ni a \quad \mapsto \quad 2z\frac{1-z}{1-2xa} + \frac{1-x}{a}. \tag{48}$$

We are thus led to the second order equation in a:

$$4x(z(1-z) - x(1-x))a^2 + 4x(1-x)a - (1-x) \quad = \quad 0. \tag{49}$$

Due to the assumption $|x - 1/2| \geqslant |z - 1/2|$, the first coefficient is non-negative. We consider two cases.

• If $|x - 1/2| = |z - 1/2|$, then (49) degenerates into a first order equation and $a \coloneqq 1/(4x)$ is the minimizer of the mapping (48). It follows that the minimizer of F on $\mathcal{L}(S_2, \pi)$ is

$$L_\wedge \quad \coloneqq \quad \begin{pmatrix} -\frac{1}{4x} & \frac{1}{4x} & 0 \\ \frac{1}{4y} & -\frac{1}{2y} & \frac{1}{4y} \\ 0 & \frac{1}{4z} & -\frac{1}{4z} \end{pmatrix}$$

and the minimal value $F_\wedge(S_2, \pi)$ of F on $\mathcal{L}(S_2, \pi)$ is

$$F(L_\wedge) \;=\; 4(1-z)z + 4(1-x)x \;=\; 8x(1-x).$$

In particular, for $\pi = \upsilon$, the uniform distribution on $\{0, 1, 2\}$, we get

$$L_\wedge \quad \coloneqq \quad \frac{1}{4}\begin{pmatrix} -3 & 3 & 0 \\ 3 & -6 & 3 \\ 0 & 3 & -3 \end{pmatrix}$$

and $F_\wedge(S_2, \upsilon) = 16/9$.

• If $|x-1/2| > |z-1/2|$, then (49) admits two solutions

$$a_\pm \;\coloneqq\; \frac{-x(1-x) \pm \sqrt{x(1-x)z(1-z)}}{2x(z(1-z)-x(1-x))},$$

but only a_+ belongs to $(0, 1/(2x))$ and is in fact the minimizer of the mapping (48). This value can be simplified into

$$a_+ \;=\; \frac{1}{2x}\frac{\sqrt{x(1-x)}}{\sqrt{x(1-x)}+\sqrt{z(1-z)}}.$$

It follows that the minimizer of F on $\mathcal{L}(S_2,\pi)$ is

$$\begin{aligned} L_\wedge \;&\coloneqq\; \frac{1}{\sqrt{x(1-x)}+\sqrt{z(1-z)}} \begin{pmatrix} -\frac{\sqrt{x(1-x)}}{2x} & \frac{\sqrt{x(1-x)}}{2x} & 0 \\ \frac{\sqrt{x(1-x)}}{2y} & -\frac{\sqrt{x(1-x)}+\sqrt{z(1-z)}}{2y} & \frac{\sqrt{z(1-z)}}{2y} \\ 0 & \frac{\sqrt{z(1-z)}}{2z} & -\frac{\sqrt{z(1-z)}}{2z} \end{pmatrix} \\ &=\; pL_{(0,1)} + (1-p)L_{(1,2)} \end{aligned}$$

with the notation introduced in (25) and

$$p \;\coloneqq\; \frac{\sqrt{x(1-x)}}{\sqrt{x(1-x)}+\sqrt{z(1-z)}}.$$

The minimal value $F_\wedge(S_2,\pi)$ of F on $\mathcal{L}(S_2,\pi)$ is

$$F(L_\wedge) \;=\; 2\left(\sqrt{x(1-x)}+\sqrt{z(1-z)}\right)^2.$$

Letting $|x-1/2|$ converge to $|z-1/2|$, we recover the values of $L_\wedge$ and $F(L_\wedge)$ obtained in the previous case.

References

[1] D. Aldous and J. A. Fill. Reversible Markov Chains and Random Walks on Graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[2] H. Cui and Y.-H. Mao. Eigentime identity for asymmetric finite Markov chains. *Front. Math. China*, 5(4):623–634, 2010.

[3] V. Ejov; N. Litvak; G. T. Nguyen and P. G. Taylor. Proof of the Hamiltonicity-trace conjecture for singularly perturbed Markov chains. *Journal of Applied Probability* 48(4): 901-910, 2011.

[4] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.

[5] M. Held and R. M. Karp. A dynamic programming approach to sequencing problems, *Journal of the Society for Industrial and Applied Mathematics* 10(1):196-210. 1962.

[6] J. J. Hunter. Mixing times with applications to perturbed Markov chains, *Linear Algebra and its Applications* 417(1):108-123, 2006.

[7] S. L. Kalpazidou. *Cycle representations of Markov processes.* Springer Verlag, New York, second edition, 2006. Stochastic Modelling and Applied Probability.

[8] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[9] J. G. Kemeny and J. L. Snell. *Finite Markov chains*, van Nostrand, Princeton, NJ, 1960.

[10] N. Litvak and V. Ejov. Markov chains and optimality of the Hamiltonian cycle. *Math. Oper. Res.*, 34(1):71–82, 2009.

[11] L. Miclo. An absorbing eigentime identity. *Markov Process and Related Fields*, 21(2):249–262, 2015.

[†]borkar@ee.iitb.ac.in

Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai 400076, India

[‡]miclo@math.univ-toulouse.fr

Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9, France