On the Markov commutator

Laurent Miclo

Institut de Mathématiques de Toulouse, UMR 5219
Toulouse School of Economics, UMR 5314
Université de Toulouse and CNRS, France

Abstract

The Markov commutator associated to a finite Markov kernel P is the convex semigroup consisting of all Markov kernels commuting with P. Its interest comes from its relation with the hypergroup property and with the notion of Markovian duality by intertwining. In particular, it is shown that the discrete analogue of the Achour-Trimèche's theorem, asserting the preservation of non-negativity by the wave equations associated to certain Metropolis birth and death transition kernels, cannot be extended to all convex potentials. But it remains true for symmetric and monotone potentials which are sufficiently convex.

1   Introduction

1.1   Motivations

The investigation of hypergroups has a long history which developed in various directions, e.g. harmonic analysis, representation theory, orthogonal polynomials and Markov processes. For a general account of their theory, in particular for its probabilistic aspects, see the reference book of Bloom and Heyer [4]. As reflected by its name, a hypergroup V is an extension of the notion of group, where the product $x \cdot y$ of two elements of V is replaced by a probability distribution $\mu_{x,y}(dz)$ over the elements of V. The family $(\mu_{x,y})_{x,y \in V}$ is assumed to satisfy some convolution axioms. Informally, when we are given a basis of some functional space over V, the hypergroup property amounts to the non-negativity of certain multi-products of the elements of the basis. The chosen basis often consists of orthogonal polynomials or other special functions. To check directly this non-negativity turns out to be a tremendous task in practice, see the fundamental papers of Gasper [12, 13]. They lead to a consequent literature to better understand the underlying principles, see in particular the paper of Koornwinder [14]. Subsequently it became important in harmonic analysis to find bases satisfying the hypergroup property. As a very particular example, consider the eigenfunction basis in $\mathbb{L}^2(\exp(-U(x))dx)$ associated to the Langevin operator $\partial^2 - U' \partial$ on $[0, 1]$ with Neumann's condition at 0 and 1, where $U : [0, 1] \rightarrow \mathbb{R}$ is a convex function. Achour and Trimèche [1] showed that this basis satisfies the hypergroup property if U is monotone or if it is symmetric with respect to $1/2$. Bakry conjectured that this property should remain true for any convex function U. In this paper we will show this is wrong for a discrete analogue of this example. Our approach relies on the commutator $\mathcal{K}(P)$ of a Markov kernel P or of a Markov semi-group $(P_t)_{t \geqslant 0}$: it is the set of all Markov kernels commuting with P or with all the elements of the semi-group. The study of $\mathcal{K}(P)$ is in fact the main purpose of this paper.

Let us further motivate the relevance of the commutator of a Markov kernel or a Markov semigroup by a potential application that appeared after this paper was written (and almost lost), in particular it will not be developed here. It concerns intertwining relations. Let $P := (P_t)_{t \geqslant 0}$ and $\widetilde{P} := (\widetilde{P}_t)_{t \geqslant 0}$ be two Markov semi-groups on respective state spaces V and $\widetilde{V}$. We say there is an intertwining relation from P to $\widetilde{P}$ when there exists a Markov from V to $\widetilde{V}$ such that

$$\forall \, t \geqslant 0, \qquad P_t \Lambda \;\; = \;\; \Lambda \widetilde{P}_t \tag{1}$$

Depending on injectivity properties of Λ, some information can be transferred from P to $\widetilde{P}$. Assume furthermore that there is an intertwining relation from $\widetilde{P}$ to P with a Markov kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V (when P and $\widetilde{P}$ are reversible with respect to probability measures, such a relation can be obtained by considering the adjoints in (1)). We get that

$$\forall \, t \geqslant 0, \qquad \widetilde{P}_t \widetilde{\Lambda} \Lambda \;\; = \;\; \widetilde{\Lambda} P_t \Lambda$$
$$= \;\; \widetilde{\Lambda} \Lambda P_t$$

namely $\widetilde{\Lambda} \Lambda$ belongs to the commutator of P. In a recent work [17], we have shown that if $\widetilde{\Lambda} \Lambda$ can be written under the form

$$\widetilde{\Lambda} \Lambda \;\; = \;\; \int_0^{+\infty} \widetilde{P}_s \, \nu(ds) \tag{2}$$

where ν is a probability distribution on $\mathbb{R}_+$, then information about the speed of convergence of $\widetilde{P}$ to its equilibrium can be deduced from the same information for P, up to a random warm-up time distributed according to ν. Thus it would be helpful to know when a representation such as (2) holds, see [17] for particular examples. In general, a better understanding of the Markov commutator of $\widetilde{P}$ is needed.

2

1.2 Presentation of the results

As mentioned in the previous subsection, a motivation for this paper is to disprove, at least in a finite context, a conjecture due to Dominique Bakry, about an extension of Achour-Trimèche's theorem [1] (see also Bakry and Huet [3]). It also provides the opportunity to begin a systematic study of the commutator convex semi-group associated to a Markov kernel.

Here we will only be concerned with state spaces V which are finite and endowed with a Markov kernel P, namely a matrix $(P(x, y))_{x,y \in V}$ whose entries are non-negative and whose row sums are equal to 1. Two classical assumptions on P are:

Irreducibility: all the coefficients of $\sum_{n \in \llbracket |V| \rrbracket} P^n$ are positive ($|V|$ is the cardinality of V and we denote for any $k \leqslant l \in \mathbb{Z}$, $\llbracket k, l \rrbracket := \{k, k+1, ..., l-1, l\}$, and $\llbracket k \rrbracket := \llbracket 1, k \rrbracket$ for $k \in \mathbb{N}$).

Reversibility: there exists a probability measure μ positive on V, such that

$$\forall \; x, y \in V, \qquad \mu(x) P(x, y) \;\; = \;\; \mu(y) P(y, x) \tag{3}$$

Under the reversibility assumption, there exist orthonormal bases of $\mathbb{L}^2(\mu)$ consisting of eigenvectors $\varphi_1, \varphi_2, ..., \varphi_{|V|}$ of P, associated to the eigenvalues $1 = \theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_{|V|} \geqslant -1$. Without loss of generality, we will always choose $\varphi_1 = \mathbb{1}$. We say that P satisfies the **hypergroup property** with respect to a point $x_0 \in V$, if the previous basis can be chosen such that $\varphi_k(x_0) \neq 0$ for all $k \in \llbracket |V| \rrbracket$, and

$$\forall \; x, y, z \in V, \qquad \sum_{k \in \llbracket |V| \rrbracket} \frac{\varphi_k(x) \varphi_k(y) \varphi_k(z)}{\varphi_k(x_0)} \;\; \geqslant \;\; 0 \tag{4}$$

These notions can be immediately extended to Markov generators L on V, namely matrices whose off-diagonal entries are non-negative and whose row sums vanish (for instance by considering the generated semi-group $(P_t)_{t \geqslant 0} := (\exp(tL))_{t \geqslant 0}$ and asking that the above conditions are satisfied by P_t, for some $t > 0$, it does not depend on the choice of $t > 0$). Extensions to more general Markov processes are also possible, but they may require some care, in particular (4) has to be understood in a weak sense. E.g. in [3], Bakry and Huet consider one-dimensional diffusion generators of the form $L_U := \partial^2 - U' \partial$ on $[-1, 1]$, with Neumann conditions on the boundary and where $U : [-1, 1] \to \mathbb{R}$ is a smooth potential. They prove Achour-Trimèche's theorem [1], asserting that if U is convex and either monotonous or symmetric with respect to 0, then L_U satisfies the hypergroup property. In a personal communication, Dominique Bakry was wondering if this result would remain true if the assumption "monotonous or symmetric with respect to 0" was removed. Our main objective is to show that this is wrong, at least in the finite setting.

More precisely, let $N \in \mathbb{N} \backslash \{1\}$ be given and denote by $\mathcal{C}$ the set of functions $U : \llbracket 0, N \rrbracket \to \mathbb{R}$ which are convex (i.e. whose natural piecewise affine extension to $[0, N]$ is convex). For $U \in \mathcal{C}$, let μ_U be the probability on $\llbracket 0, N \rrbracket$ given by

$$\forall \; x \in \llbracket 0, N \rrbracket, \qquad \mu_U(x) \;\; := \;\; Z_U^{-1} \exp(-U(x)) \tag{5}$$

where Z_U is the normalizing constant. For any $U \in \mathcal{C}$, assume we are given an irreducible birth and death Markov transition P_U on $\llbracket 0, N \rrbracket$ whose invariant probability is μ_U. Recall that a **birth and death** Markov transition P on $\llbracket 0, N \rrbracket$ is a Markov kernel such that

$$\forall \; x, y \in \llbracket 0, N \rrbracket, \qquad P(x, y) > 0 \;\; \Rightarrow \;\; |x - y| \leqslant 1$$

An invariant measure of such a kernel necessarily satisfies (3), so that an irreducible birth and death Markov matrix is reversible.

Endowing $\mathcal{C}$ and the set of Markov kernels with the topology inherited respectively from $\mathbb{R}^{\llbracket 0, N \rrbracket}$ and $\mathbb{R}^{\llbracket 0, N \rrbracket^2}$, we say that the above mapping $\mathcal{C} \ni U \mapsto P_U$ is a (birth and death) **generalized**

Metropolis procedure if it is continuous. A classical Metropolis procedure corresponds for instance to the Markov kernel M_U defined by

$$\forall \ x \neq y \in [\![0, N]\!], \qquad M_U(x, y) \ := \ \frac{M_0(x, y)}{\Sigma_U} \exp\left(\frac{U(x) - U(y)}{2}\right) \tag{6}$$

where the exploration Markov kernel M_0 is given by

$$\forall \ x \neq y \in [\![0, N]\!], \qquad M_0(x, y) \ := \ \begin{cases} 1/2 & \text{, if } |x - y| = 1 \\ 0 & \text{, otherwise} \end{cases} \tag{7}$$

and where

$$\Sigma_U \ := \ \max_{x \in [\![0, N]\!]} \sum_{y \in [\![0, N]\!] \setminus \{x\}} M_0(x, y) \exp\left(\frac{U(x) - U(y)}{2}\right) \tag{8}$$

As usual, the diagonal entries of the matrices M_U and M_0 are imposed by the condition that the row sums are equal to 1.

Our main result is:

Theorem 1 *There does not exist a generalized Metropolis procedure $\mathcal{C} \ni U \mapsto P_U$ such that P_U satisfies the hypergroup property for all $U \in \mathcal{C}$.*

In [16], we checked numerically (by appropriate random choices of U in $\mathcal{C}$) that a variant of the classical Metropolis procedure (described as $\mathcal{C} \ni U \mapsto \widehat{M}_U$ with the notation introduced in (35) below) does not satisfy the hypergroup property.

The proof of Theorem 1 is based on properties of the **commutator convex semi-group** $\mathcal{K}(P)$ associated to a Markov kernel P on V: it is the set of Markov kernels K on V commuting with P: $KP = PK$. It is immediate to see that it is convex and that it is a semi-group: if K and K' belong to $\mathcal{K}(P)$, the same is true for their product KK'. It was introduced in [16], because it gives a simple Markovian characterization of the hypergroup property for certain kernels. More precisely, let us introduce the following objects:

$$\forall \ x \in V, \qquad \mathcal{K}(P, x) \ := \ \{K(x, \cdot) \ : \ K \in \mathcal{K}(P)\} \ \subset \ \mathcal{P}(V)$$

where $\mathcal{P}(V)$ is the convex set of probability measures on V, and

$$\mathcal{H}(P) \ = \ \{x \in V \ : \ \mathcal{K}(P, x) = \mathcal{P}(V)\}$$

Furthermore, say that a Markov kernel is **uniplicit** if it is reversible and if all its eigenvalues are of multiplicity 1 (in particular the eigenvalue 1 is of multiplicity 1, so that uniplicity implies irreducibility). The interest of these notions is:

Lemma 2 *An uniplicit Markov kernel P on V satisfies the hypergroup property with respect to $x_0 \in V$ if and only if $x_0 \in \mathcal{H}(P)$.*

Let us give succinctly some underlying arguments, since this is the only place in the paper where Definition (4) will play a role.

Proof

The reverse implication was observed in [16] and the direct implication is a consequence of the considerations of Bakry and Huet [3], the uniplicit assumption is not even needed, as the following reminder show. Let P be a reversible Markov kernel P on V with an associated orthonormal basis of eigenvectors $\varphi_1, \varphi_2, ..., \varphi_{|V|}$ as above. Assume that P satisfies the hypergroup property with respect to $x_0 \in V$. Let $x \in V$ be given and consider the kernel K_x given by

$$\forall \ y, z \in V, \qquad K_x(y, z) \ := \ \sum_{k \in [\![|V|]\!]} \frac{\varphi_k(x)\varphi_k(y)\varphi_k(z)}{\varphi_k(x_0)} \mu(z)$$

By assumption it is non-negative and for any fixed $y \in V$, we have by orthonormality,

$$
\begin{aligned}
\sum_{z \in V} K_x(y,z) &= \sum_{z \in V} K_x(y,z)\varphi_1(z) \\
&= \sum_{k \in [\![|V|]\!]} \frac{\varphi_k(x)\varphi_k(y)}{\varphi_k(x_0)} \sum_{z \in V} \varphi_k(z)\varphi_1(z)\,\mu(z) \\
&= \frac{\varphi_1(x)\varphi_1(y)}{\varphi_1(x_0)} \\
&= 1
\end{aligned}
$$

Thus K_x is a Markov kernel. A similar computation shows that for any $k \in [\![2,|V|]\!]$, φ_k is also an eigenfunction of K_x associated to the eigenvalue $\varphi_k(x)/\varphi_k(x_0)$. It follows that K_x shares with P the same basis of eigenvectors, so that $K_x \in \mathcal{K}(P)$. Furthermore, we have that for any $l \in [\![|V|]\!]$,

$$
\begin{aligned}
K_x[\varphi_l](x_0) &:= \sum_{z \in V} K_x(x_0,z)\varphi_l(z) \\
&= \sum_{z \in V} \sum_{k \in [\![|V|]\!]} \varphi_k(x)\varphi_k(z)\varphi_l(z)\,\mu(z) \\
&= \varphi_l(x)
\end{aligned}
$$

It implies that $K_x(x_0,\cdot) = \delta_x$. So for any $x \in V$, $\delta_x \in \mathcal{K}(P,x_0)$. Taking into account that $\mathcal{K}(P,x_0)$ is always a convex set, we get that $x_0 \in \mathcal{H}(P)$.

∎

Remark 3 (a) The uniplicity assumption cannot be removed for the reverse implication of Lemma 2. Consider P the transition kernel of the random walk on $V := \mathbb{Z}/(n\mathbb{Z})$, with $n \in \mathbb{N}\backslash\{1,2\}$. At the end of Section 2.5 from [3], Bakry and Huet show that P does not satisfy the hypergroup property. Nevertheless, consider for $v \in \mathbb{Z}/(n\mathbb{Z})$, the translation by v kernel K defined by

$$
\forall\ x,y \in \mathbb{Z}/(n\mathbb{Z}), \qquad K(x,y) := \delta_{x+v}(y)
$$

Clearly $K \in \mathcal{K}(P)$ and $K(0,\cdot) = \delta_v$, so that $\delta_v \in \mathcal{K}(P,0)$ for all $v \in \mathbb{Z}/(n\mathbb{Z})$. It follows that $0 \in \mathcal{H}(P)$. More precisely, we have $\mathcal{H}(P) = \mathbb{Z}/(n\mathbb{Z})$.

(b) The example in (a) satisfies the **complex hypergroup property** with respect to any point $x_0 \in \mathbb{Z}/(n\mathbb{Z})$ (see Proposition 2.10 of Bakry and Huet [3]), in the sense that we can find an unitary basis $(\varphi_1,\ \varphi_2,...,\ \varphi_{|V|})$ of $\mathbb{L}^2(\mu,\mathbb{C})$ consisting of eigenvectors of P such that $\varphi_k(x_0) \neq 0$ for all $k \in [\![|V|]\!]$, and

$$
\forall\ x,y,z \in V, \qquad \sum_{k \in [\![|V|]\!]} \frac{\varphi_k(x)\varphi_k(y)\overline{\varphi_k(z)}}{\varphi_k(x_0)}\ \geqslant\ 0 \tag{9}
$$

So maybe the condition

$$
\mathcal{H}(P)\ \neq\ \varnothing \tag{10}
$$

is related to the complex hypergroup property. But here we will not investigate this question. We will mainly be interested in (10), seen as a generalization of the hypergroup property, because it could be considered for Markov kernels which are not reversible (or defined on abstract measurable spaces: (10) enables to avoid the technical difficulties related to the summations appearing in (4) or (9) when the state space is not finite).

□

An irreducible birth and death kernel is necessarily uniplicit, so in the context of Theorem 1, the hypergroup property for a Markov kernel P is equivalent to (10). We are thus lead to investigate the corresponding Markov commutator convex semi-group and will do it using general arguments. The two properties we will need are

Proposition 4 *Assume that P is an irreducible Markov kernel and let μ be its invariant probability. Then we have*

$$\forall\ x \in \mathcal{H}(P), \qquad \mu(x) \ = \ \min_V \mu$$

For the second property, we need to introduce the **symmetry group** $\mathcal{S}_P$ associated to P: it is the set of bijective mappings $g : V \to V$ such that

$$\forall\ x, y \in V, \qquad P(g(x), g(y)) \ = \ P(x, y) \tag{11}$$

For instance, one recovers the permutation group $\mathcal{S}_V$ of V if P is either the identity matrix I (no move is permitted) or the matrix whose all off-diagonal entries are equal to $1/(|V|-1)$ (all "true" moves are equally permitted). Indeed $\mathcal{S}_P = \mathcal{S}_V$ if and only if P is a convex combination of the two previous matrices, situations where all the elements of V are indistinguishable with respect to the evolution dictated by P.

Proposition 5 *Assume that P is an uniplicit Markov kernel and let $x_0, x_1 \in \mathcal{H}(P)$. Then there exists $g \in \mathcal{S}_P$ such that $g(x_1) = x_0$. Conversely, any $g \in \mathcal{S}_P$ stabilizes $\mathcal{H}(P)$, so that $\mathcal{H}(P)$ is the orbit of any of its element under $\mathcal{S}_P$.*

Another natural question in the finite birth and death setting is the transposition of the Achour-Trimèche's theorem known in the continuous framework. We did not succeed in getting a really satisfactory answer in this direction. The next result is obtained by adapting the arguments of Bakry and Huet [3]. Let $\widetilde{\mathcal{C}}$ be the subset of $U \in \mathcal{C}$ such that $U(x+2) - U(x+1) \geqslant U(x+1) - U(x) + 2\ln(2)$ for all $x \in [\![0, N-2]\!]$ (equivalently, U is the restriction to $[\![0, N]\!]$ of a $\mathcal{C}^2$ function on $[0, N]$ satisfying $U'' \geqslant 2\ln(2)$). Let $\widetilde{\mathcal{C}}_{\mathrm{m}}$ be the subset of $\widetilde{\mathcal{C}}$ consisting of monotonous mappings such that $|U(N) - U(N-1)| \wedge |U(1) - U(0)| \geqslant 2\ln(2)$. Consider also $\widetilde{\mathcal{C}}_{\mathrm{s}}$ the subset of $\widetilde{\mathcal{C}}$ consisting of mappings symmetric with respect to $N/2$.

Proposition 6 *For any $U \in \widetilde{\mathcal{C}}_{\mathrm{m}} \cup \widetilde{\mathcal{C}}_{\mathrm{s}}$, the Metropolis kernel M_U defined in (6) satisfies the hypergroup property. Thus the mapping $\widetilde{\mathcal{C}}_{\mathrm{m}} \cup \widetilde{\mathcal{C}}_{\mathrm{s}} \ni U \mapsto M_U$ is a birth and death Metropolis procedure satisfying the hypergroup property.*

In the one-dimensional diffusive setting, the result corresponding to $\widetilde{\mathcal{C}}_{\mathrm{m}}$ is due to Chebli [6].

Note that from Propositions 4 and 5, we deduce that in the symmetric situation, $\mathcal{H}(M_U) = \{0, N\}$, and that in the monotonous case with U non-constant, $\mathcal{H}(M_U)$ is the singleton consisting of the boundary element with the smallest weight with respect to the reversible measure μ_U.

Remark 32 (d) gives another example of a generalized Metropolis procedure satisfying the hypergroup property for some convex potentials (more general than those considered in Proposition 6). It would be very interesting to find other closed subsets $\mathcal{C}' \subset \mathcal{C}$ for which we can find a generalized Metropolis procedure $\mathcal{C}' \ni U \mapsto P_U$ satisfying the hypergroup property (or to describe $\mathcal{C}' := \{U \in \mathcal{C} : \mathcal{H}(M_U) \neq \varnothing\}$). Especially to try to deduce the analogous results in the continuous framework, in order to recover Gasper's example [12, 13], see also Bakry and Huet [3] and Carlen, Geronimo and Loss [5].

From general considerations related to the Markov commutator convex semi-groups, we will also deduce the following criterion. Let $\bar{P}$ be a Markov kernel on the finite set $\bar{V}$, consider $\bar{G}$ a subgroup of $\mathcal{S}_{\bar{P}}$ and denote by $\equiv$ the equivalence relation it induces on $\bar{V}$ via

$$\forall\ \bar{x}, \bar{y} \in \bar{V}, \qquad \bar{x} \equiv \bar{y} \ \Leftrightarrow \ \exists\ g \in \bar{G} : g(\bar{x}) = \bar{y}$$

Denote by V the set of equivalence classes for $\equiv$ and by $\pi : \bar{V} \to V$ the associated projection mapping. It is immediate to check that a Markov kernel P is well-defined on V through the formula

$$\forall\ x, y \in V, \qquad P(x, y) \ := \ \bar{P}(\bar{x}, \pi^{-1}(y))$$

6

where $\bar{x}$ is any point of $\bar{V}$ such that $\pi(\bar{x}) = x$. This construction corresponds to a reduction of the symmetries of $\bar{P}$. The next result shows that some properties of $\bar{P}$ are preserved under this operation. It will be used to check the hypergroup property of M_U for $U \in \widetilde{\mathcal{C}}_{\mathrm{m}}$, knowing it for $U \in \widetilde{\mathcal{C}}_{\mathrm{s}}$.

Proposition 7 *Assume that $\bar{P}$ is uniplicit and satisfies Condition* (10). *Then the same remains true for P.*

If the uniplicity of $\bar{P}$ could be removed from this statement and be replaced by the weaker uniplicity of P, Proposition 7 would provide an abstract rewriting in the finite context of the Carlen, Geronimo and Loss method [5]. This conjectured extension seems quite challenging, some assumptions could be required on the subgroup $\bar{G}$. Maybe they do not appear here, because when $\bar{P}$ is uniplicit, $\mathcal{S}_{\bar{P}}$ is commutative, see Remark 20 (a) below.

In the next section we will study the Markov commutator convex semi-group in the general finite framework, obtaining in particular Propositions 4, 5 and 7. We will take advantage of the relations between the Markov commutator convex semi-group and the theory of Markov intertwining as it was developed by Diaconis and Fill [7]. In the last section we consider more specifically the birth and death case and prove Theorem 1 and Proposition 6.

2 General properties

This is the beginning of a systematic investigation of the Markov commutator convex semigroup $\mathcal{K}(P)$ associated to a finite Markov kernel P.

We start by recalling some elements of the theory of Markov intertwining due to Diaconis and Fill [7]. Let $X := (X_n)_{n \in \mathbb{Z}_+}$ and $\bar{X} := (\bar{X}_n)_{n \in \mathbb{Z}_+}$ be two Markov chains, respectively on the finite state spaces V and $\bar{V}$. The respective transition kernels are denoted P and $\bar{P}$, and the initial distributions m_0 and $\bar{m}_0$. We say that X is **intertwined** with $\bar{X}$ **through the Markov link** Λ (which is a Markov kernel from $\bar{V}$ to V, seen as a $\bar{V} \times V$ matrix), if there is a coupling $(X, \bar{X})$ such that the two following conditions are met:

$$\forall\, n \in \mathbb{Z}_+, \qquad \mathcal{L}(\bar{X}_{[\![0,n]\!]}|X) \;\;=\;\; \mathcal{L}(\bar{X}_{[\![0,n]\!]}|X_{[\![0,n]\!]}) \tag{12}$$

where as usual this identity of conditional laws has to be understood a.s. with respect to the probability measure underlying the coupling. The trajectorial notation $X_{[\![0,n]\!]} := (X_p)_{p \in [\![0,n]\!]}$ was used.

$$\forall\, n \in \mathbb{Z}_+, \qquad \mathcal{L}(X_n|\bar{X}_{[\![0,n]\!]}) \;\;=\;\; \Lambda(\bar{X}_n, \cdot) \tag{13}$$

When these assumptions are satisfied, we write $X \prec_\Lambda \bar{X}$ and $\bar{X}$ is also said to be a **dual chain** of X through Λ. The notation $X \prec \bar{X}$ will notify there exists Λ such that $X \prec_\Lambda \bar{X}$.

We say that (m_0, P) is **intertwined** with $(\bar{m}_0, \bar{P})$ **through the Markov link** Λ if

$$m_0 \;=\; \bar{m}_0 \Lambda \quad \text{and} \quad \bar{P}\Lambda \;=\; \Lambda P \tag{14}$$

We denote this relation by $(m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P})$ and as above, $(m_0, P) \prec (\bar{m}_0, \bar{P})$ means there exists a kernel Λ such that (14) is satisfied.

Diaconis and Fill [7] have shown that these notions of intertwining coincide, at least if $\bar{X}$ visits the whole state space $\bar{V}$ (in particular if $\bar{P}$ is irreducible):

Proposition 8 *With the above notations, we have*

$$(m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P}) \;\;\Rightarrow\;\; X \prec_\Lambda \bar{X}$$

Furthermore if for any $\bar{x} \in \bar{V}$, there exists $n \in \mathbb{Z}_+$ such that $\mathbb{P}[\bar{X}_n = \bar{x}] > 0$, then

$$X \prec_\Lambda \bar{X} \;\;\Rightarrow\;\; (m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P})$$

Proof:

More specifically, the construction of the coupling of X and $\bar{X}$ satisfying the conditions (12) and (13) under the assumption $(m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P})$ is described in Theorem 2.17 of Diaconis and Fill [7]. The other implication can also be deduced from their considerations. For the sake of completeness, here are some arguments, directly based on the hypotheses (12) and (13).

From (13), we deduce that for all $n \in \mathbb{Z}_+$, $\mathcal{L}(X_n | \bar{X}_n) = \Lambda(\bar{X}_n, \cdot)$ so that by integration with respect to $\bar{X}_n$, we get $\mathcal{L}(X_n) = \mathcal{L}(\bar{X}_n)\Lambda$. In particular for $n = 0$, we obtain $m_0 = \bar{m}_0\Lambda$. Let f and $\bar{f}$ two test functions defined respectively on V and $\bar{V}$. For fixed $n \in \mathbb{Z}_+$, we compute $\mathbb{E}[\bar{f}(\bar{X}_n)f(X_{n+1})]$ in two ways. First, using (13) and the Markov property of $\bar{X}$,

$$
\begin{aligned}
\mathbb{E}[\bar{f}(\bar{X}_n)f(X_{n+1})] &= \mathbb{E}[\bar{f}(\bar{X}_n)\mathbb{E}[f(X_{n+1})|\bar{X}_{[\![0,n+1]\!]}]] \\
&= \mathbb{E}[\bar{f}(\bar{X}_n)\Lambda[f](\bar{X}_{n+1})] \\
&= \mathbb{E}[\bar{f}(\bar{X}_n)(\bar{P}\Lambda)[f](\bar{X}_n)]
\end{aligned}
$$

Second, using (12) and the Markov property of X,

$$
\begin{aligned}
\mathbb{E}[\bar{f}(\bar{X}_n)f(X_{n+1})] &= \mathbb{E}[\mathbb{E}[\bar{f}(\bar{X}_n)|X]f(X_{n+1})] \\
&= \mathbb{E}[\mathbb{E}[\bar{f}(\bar{X}_n)|X_{[\![0,n]\!]}]f(X_{n+1})] \\
&= \mathbb{E}[\mathbb{E}[\bar{f}(\bar{X}_n)|X_{[\![0,n]\!]}]P[f](X_n)] \\
&= \mathbb{E}[\bar{f}(\bar{X}_n)P[f](X_n)] \\
&= \mathbb{E}[\bar{f}(\bar{X}_n)\mathbb{E}[P[f](X_n)|\bar{X}_{[\![0,n]\!]}]] \\
&= \mathbb{E}[\bar{f}(\bar{X}_n)(\Lambda P)[f](\bar{X}_n)]
\end{aligned}
$$

Since this is true for any $\bar{f}$, we deduce that a.s.,

$$
(\Lambda P)[f](\bar{X}_n) = (\bar{P}\Lambda)[f](\bar{X}_n)
$$

and due to the assumption on $\bar{X}$,

$$
\forall\, \bar{x} \in \bar{V}, \qquad (\Lambda P)[f](\bar{x}) = (\bar{P}\Lambda)[f](\bar{x})
$$

Since it is true for all f, it follows that $\Lambda P = \bar{P}\Lambda$. $\blacksquare$

Remark 9 (a) The relation $\prec$ is clearly reflexive (through the identity link) and it can be easily checked to be transitive (for instance at the level of the Markov chains, if $X \prec_\Lambda X'$ and $X' \prec_{\Lambda'} X''$ then $X \prec_{\Lambda\Lambda'} X''$). Thus $\prec$ is a pre-order, e.g. on the trajectorial laws of finite Markov chains (whose state space is a subset of $\mathbb{N}$, to work on a defined set). It is then tempting to verify if it would not be an equivalence or an order relation. To see that $\prec$ is none, consider Y the trivial Markov chain on a singleton. For any finite Markov chain X, we have $Y \prec X$, but $X \prec Y$ is equivalent to the stationarity of X (namely the initial distribution of X is invariant for its transition kernel). It follows that $\prec$ is neither symmetrical nor anti-symmetrical. Next, one can define an equivalence relation $X \sim X'$ via $X \prec X'$ and $X' \prec X$. On the corresponding equivalence classes, $\prec$ defines a partial order relation, in some sense it should compare the difficulty of reaching an equilibrium (see also Remark 11 below). The "stationarity" class of the trivial chain Y is minimal for this order.

(b) Similar conditions are valid for the algebraic intertwining between couples consisting of a probability measure and a Markov kernel. If the finite state set V and the Markov kernel P are fixed, we induce a relation on $\mathcal{P}(V)$ via $m_0 \prec \bar{m}_0$ if and only if $(m_0, P) \prec (\bar{m}_0, P)$. It can be transformed into an order relation on $\mathcal{P}(V)/\sim$ by introducing an equivalence relation $\sim$ as above. It heuristically corresponds to the proximity to the set of invariant measures for P, which are the minimal elements. Note that the semigroup $(P^n)_{n\in\mathbb{Z}_+}$ is non-increasing with respect to $\prec$, since we have $(m_0P, P) \prec_P (m_0, P)$.

$\square$

The main interest of associating a dual chain $\bar{X}$ to a given Markov chain X is that it enables to construct strong times (see for instance Diaconis and Fill [7], Fill [11], Diaconis and Miclo [8] and [15]). A stopping time τ for X (with respect to a filtration containing the filtration generated by X) is a **strong time** if it is a.s. finite and if τ and X_τ are independent. The basic principle of the construction is the following well-known result, whose proof is given for the sake of completeness.

Lemma 10 *Let $(X, \bar{X})$ be a coupling satisfying* (13), *then this equality can be extended to any a.s. finite stopping time τ for $\bar{X}$, namely*

$$\mathcal{L}(X_\tau | \bar{X}_{[\![0,\tau]\!]}) \ = \ \Lambda(\bar{X}_\tau, \cdot)$$

If in addition $(X, \bar{X})$ satisfies (12), *then τ is a strong time if $\Lambda(\bar{X}_\tau, \cdot)$ is independent from τ (for instance if $\Lambda(\bar{X}_\tau, \cdot)$ "is not really depending on" $\bar{X}_\tau$, e.g. if $\bar{X}_\tau$ is a.s. equal to a fixed point).*

Proof

The first assertion is an outcome of the notion of a stopping time: Let f be a function defined on V and $\bar{F}$ a bounded functional measurable with respect to the stopped trajectory $\bar{X}_{[\![0,\tau]\!]}$. We compute that

$$
\begin{aligned}
\mathbb{E}[f(X_\tau)\bar{F}] &= \sum_{n \in \mathbb{Z}_+} \mathbb{E}[f(X_n)\bar{F}\mathbb{1}_{\tau=n}] \\
&= \sum_{n \in \mathbb{Z}_+} \mathbb{E}[\mathbb{E}[f(X_n)|\bar{X}_{[\![0,n]\!]}]\bar{F}\mathbb{1}_{\tau=n}] \\
&= \sum_{n \in \mathbb{Z}_+} \mathbb{E}[\Lambda[f](\bar{X}_n)\bar{F}\mathbb{1}_{\tau=n}] \\
&= \mathbb{E}[\Lambda[f](\bar{X}_\tau)\bar{F}]
\end{aligned}
$$

where the second equality comes from the fact that $\bar{F}\mathbb{1}_{\tau=n}$ is measurable with respect to $\bar{X}_{[\![0,n]\!]}$. The first wanted result follows, since this is true for all f and $\bar{F}$ as above.

For the second assertion, note that (12) implies that a stopping time for $\bar{X}$ is also a stopping time for X. Let f be a function defined on V and let g be a bounded measurable mapping on $\mathbb{R}_+$. Since τ is measurable with respect to $\bar{X}_{[\![0,\tau]\!]}$, we have

$$
\begin{aligned}
\mathbb{E}[f(X_\tau)g(\tau)] &= \mathbb{E}[\mathbb{E}[f(X_\tau)|\bar{X}_{[\![0,\tau]\!]}]g(\tau)] \\
&= \mathbb{E}[\Lambda[f](\bar{X}_\tau)g(\tau)] \\
&= \mathbb{E}[\Lambda[f](\bar{X}_\tau)]\mathbb{E}[g(\tau)] \\
&= \mathbb{E}[f(X_\tau)]\mathbb{E}[g(\tau)]
\end{aligned}
$$

where the third equality comes from the assumption made on $\Lambda(\bar{X}_\tau, \cdot)$. The independence of τ and X_τ follows, since f and g were arbitrary. $\blacksquare$

For the purpose of proving Proposition 4, we will only use the first part of the above lemma, even if the stopping times we will consider are indeed strong times.

Indeed, it is time to come back to the Markov commutator convex semigroup $\mathcal{K}(P)$ associated to an irreducible finite Markov kernel P. Denote $X^{m_0} := (X_t^{m_0})_{t \geqslant 0}$ a Markov chain with P as transition kernel and $m_0 \in \mathcal{P}(V)$ for initial law. From the definitions and Proposition 8, we have for any $K \in \mathcal{K}(P)$ and any initial distribution m_0, $X^{m_0 K} \prec_K X^{m_0}$. Thus it appears that $x_0 \in \mathcal{H}(P)$ if and only if for any $m_0 \in \mathcal{P}(V)$, there exists a Markov kernel K on V such that $X^{m_0} \prec_K X^{x_0}$ (as customary, X^{x_0} is a shorthand for $X^{\delta_{x_0}}$). In particular, if P is uniplicit, then P satisfies the

hypergroup property if and only there exists $x_0 \in V$ such that for any $m_0 \in \mathcal{P}(V)$, $X^{m_0} \prec X^{x_0}$. More generally, we get the following interpretation:

$$\forall \ x \in V, \qquad \mathcal{K}(P, x) \quad = \quad \{m \in \mathcal{P}(V) \ : \ X^m \prec X^x\}$$

All preliminaries are now in place for the

Proof of Proposition 4

Consider $x_0 \in \mathcal{H}(P)$ and let x_1 be any point of V. We want to show that $\mu(x_0) \leqslant \mu(x_1)$.
By definition of $\mathcal{H}(P)$, there exists $K \in \mathcal{K}(P)$ such that $K(x_0, \cdot) = \delta_{x_1}$, so that from Proposition 8, $X^{x_1} \prec_K X^{x_0}$, i.e. we can construct a coupling of X^{x_0} and X^{x_1} satisfying (12) and (13) with $\Lambda := K$.
Let $(\tau_n)_{n \in \mathbb{Z}_+}$ be the sequence of stopping times for X^{x_0} defined by iteration through $\tau_0 = 0$ and

$$\forall \ n \in \mathbb{Z}_+, \qquad \tau_{n+1} \quad := \quad \inf\{p > \tau_n \ : \ X_p = x_0\}$$

According to Lemma 10, for any $n \in \mathbb{Z}_+$,

$$\mathcal{L}(X_{\tau_n}^{x_1} | X_{[\![0, \tau_n]\!]}^{x_0}) \quad = \quad \delta_{x_1}$$

It means that each time X^{x_0} is in x_0, then X^{x_1} is in x_1. It remains to apply the ergodic theorem to get

$$
\begin{aligned}
\mu(x_0) \quad &= \quad \lim_{n \to \infty} \frac{1}{n+1} \sum_{p \in [\![0, n]\!]} \mathbb{1}_{x_0}(X^{x_0}) \\
&\leqslant \quad \lim_{n \to \infty} \frac{1}{n+1} \sum_{p \in [\![0, n]\!]} \mathbb{1}_{x_1}(X^{x_1}) \\
&= \quad \mu(x_1)
\end{aligned}
$$

where the (in)equalities are valid a.s.

■

The elements of $\mathcal{H}(P)$ satisfies other optimization properties, they are for instance points from which it is the most difficult to reach equilibrium in the separation discrepancy sense:

Remark 11 Recall that the **separation discrepancy** $s(m, \mu)$ between two probability measures on V is defined by

$$s(m, \mu) \quad := \quad \sup_{x \in V} 1 - \frac{m(x)}{\mu(x)}$$

(with the usual convention: $r/0 = +\infty$ for any $r > 0$, but $0/0 = 0$).
A **stationary time** τ for an irreducible Markov chain $X^{m_0} := (X_n^{m_0})_{n \in \mathbb{Z}_+}$ (m_0 still stands for the initial distribution) is a strong time such that $X_\tau^{m_0}$ is distributed according to the associated invariant measure μ. Aldous and Diaconis [2] have shown that if the transition kernel is aperiodic and irreducible, then for any initial distribution m_0, there exists a stationary time τ^{m_0} associated to X^{m_0} satisfying

$$\forall \ n \in \mathbb{Z}_+, \qquad \mathbb{P}[\tau^{m_0} > n] \quad = \quad s(m_0 P^n, \mu)$$

Furthermore τ^{m_0} is stochastically smaller than any stationary time associated to X^{m_0}.
The proof of Proposition 4 can be slightly modified to show that if $x_0 \in \mathcal{H}(P)$, then τ^{x_0} is stochastically larger than τ^{m_0} for any initial distribution m_0. Indeed, if $K \in \mathcal{K}(P)$ is such that $K(x_0, \cdot) = m_0$, then considering a coupling of X^{x_0} and X^{m_0} realizing the relation $X^{m_0} \prec_K X^{x_0}$, it appears that τ^{x^0} is a stationary time for X^{m_0}. It is a consequence of the fact that all the elements of $\mathcal{K}(P)$ admit μ for invariant measure, as it was seen in [16] (only the irreducibility of P is needed

for this property). The stochastic domination of τ^{m_0} by τ^{x_0} ensures that for any initial distribution m_0 (or equivalently for any Dirac mass $m_0 = \delta_{x_1}$, with x_1 in the state space V),

$$\forall\, n \in \mathbb{Z}_+, \qquad s(m_0 P^n, \mu) \;\leqslant\; s(P^n(x_0, \cdot), \mu)$$

□

To go in the direction of Proposition 5, we begin by a simple technical result:

Lemma 12 *Let K and K' be two Markov kernels on V such that $K'K = I$, the identity kernel. Then there exist $g \in \mathcal{S}_V$ such that*

$$\forall\, x, y \in V, \qquad \left\{ \begin{array}{rcl} K(x,y) & = & \delta_{g(x)}(y) \\ K'(x,y) & = & \delta_{g^{-1}(x)}(y) \end{array} \right.$$

Proof

By contradiction, assume there exists $x \in V$ such that $K(x, \cdot)$ is not a Dirac mass. Then for any $y \in V$, if $K'(y, x) > 0$ then $K'K(y, \cdot)$ cannot be a Dirac mass. This is not compatible with $K'K = I$, so we must have $K'(y, x) = 0$ for all $y \in V$. It implies that K' is not invertible, in contradiction again with our assumption. So for any $x \in V$, $K(x, \cdot)$ is a Dirac mass $\delta_{g(x)}$ for some $g(x) \in V$. Since K is invertible, necessarily the mapping g is also invertible. The announced result follows at once.

∎

In addition, we will need the following consequence of the uniplicit assumption.

Lemma 13 *Assume that P is uniplicit, then for any fixed $x_0 \in \mathcal{H}(P)$, the affine mapping*

$$\mathcal{K}(P) \ni K \quad \mapsto \quad K(x_0, \cdot) \in \mathcal{K}(P, x_0)$$

is one-to-one.

Proof

Fix $x_0 \in \mathcal{H}(P)$ and $m_0 \in \mathcal{P}(V)$, it is sufficient to see there is exactly one matrix K solution to the equations

$$\begin{array}{rcl} K(x_0, \cdot) & = & m_0 \\ KP & = & PK \end{array}$$

Indeed, consider μ the reversible probability for P and let $\varphi_1, \varphi_2, ..., \varphi_{|V|}$ be an orthonormal (in $\mathbb{L}2(\mu)$) basis of eigenvectors associated to P as in the introduction. By the commutation of K with P, this is also a basis of eigenvectors for K. Thus we can find numbers $a_1, a_2, ..., a_{|V|}$ such that

$$\forall\, x, y \in V, \qquad K(x,y) \;=\; \sum_{l \in [\![1,|V|]\!]} a_l \varphi_l(x) \varphi_l(y) \mu(y)$$

The first condition then reads

$$\forall\, y \in V, \qquad \frac{m_0}{\mu}(y) \;=\; \sum_{l \in [\![1,|V|]\!]} a_l \varphi_l(x_0) \varphi_l(y)$$

namely $(a_l \varphi_l(x_0))_{l \in [\![1,|V|]\!]}$ are the coefficients of m_0/μ in the basis $(\varphi_1, \varphi_2, ..., \varphi_{|V|})$. Since $\varphi_l(x_0) \neq 0$ for all $l \in [\![1, |V|]\!]$, according to Lemma 2, we get that the $a_1, a_2, ..., a_{|V|}$ are uniquely determined.

∎

In particular if P is an uniplicit kernel satisfying the hypergroup property, then $\mathcal{K}(P)$ is a simplex. It is sometimes possible to go further:

Remark 14 In fact the above proof shows that if $x_0 \in V$ is any point such that $\varphi_l(x_0) \neq 0$ for all $l \in [\![1, |V|]\!]$, then the conclusion of Lemma 13 still holds if P is uniplicit. If furthermore (10) holds, then $\mathcal{K}(P)$ is a simplex as well as each of the $\mathcal{K}(P, x)$, for $x \in V$.

Let $\mathcal{R}$ be the set of Markov kernels which are irreducible and reversible. It can be easily seen that the subset of elements of $\mathcal{R}$ which are uniplicit and whose eigenvectors never vanish is a dense open subset of $\mathcal{R}$. But since $\mathcal{H}$, the subset of $\mathcal{R}$ consisting of kernels satisfying the hypergroup property, is very slim in $\mathcal{R}$, it is no longer clear whether or not the subset of elements of $\mathcal{H}$ which are uniplicit and whose eigenvectors never vanish is a dense open subset of $\mathcal{H}$. If it was true, it could be concluded that "generically", $\mathcal{K}(P, x)$ is a simplex for $P \in \mathcal{H}$ and $x \in V$.

$\square$

We have all the ingredients for the

Proof of Proposition 5

Let be given $x_0, x_1 \in \mathcal{H}(P)$. Then there exist $K', K \in \mathcal{K}(P)$ such that

$$
\begin{aligned}
K'(x_0, \cdot) &= \delta_{x_1} \\
K(x_1, \cdot) &= \delta_{x_0}
\end{aligned}
\tag{15}
$$

Thus we get that $K'K(x_0, \cdot) = \delta_{x_0}$. Since P is assumed to be uniplicit, we get from Lemma 13 that $K'K \in \mathcal{K}(P)$ is uniquely determined by this relation. It appears there is no alternative: $K'K = I$. Lemma 12 enables to find a permutation $g \in \mathcal{S}_V$ such K is the Markov kernel induced by g. Note that (15) translates into $g(x_1) = x_0$. The commutation of K and P then implies that

$$
\forall\, x, y \in V, \qquad P(g(x), y) = P(x, g^{-1}(y))
$$

which can be rewritten under the form (11) namely $g \in \mathcal{S}_P$. The remaining assertions of Proposition 5 are straightforward.

■

We are now going in the direction of Proposition 7 through a sequence of general arguments, in the hope they present in a clear way the problems one will encounter in trying to generalize it. We start by recalling some considerations from [16]. A Markov kernel Λ from $\bar{V}$ to V can be interpreted as an operator sending any function f defined on V to the mapping $\Lambda[f]$ defined on $\bar{V}$ by

$$
\forall\, \bar{x} \in \bar{V}, \qquad \Lambda[f](\bar{x}) := \sum_{x \in V} \Lambda(\bar{x}, x) f(x)
$$

Let $\bar{\mu}$ be a probability measure given on $\bar{V}$ and consider $\mu := \bar{\mu}\Lambda$ its image by Λ. Then Λ can be seen as an operator from $\mathbb{L}^2(\mu)$ to $\mathbb{L}^2(\bar{\mu})$ (because $\Lambda[f]$ is $\bar{\mu}$-negligible if f is μ-negligible). It enables to define Λ^* its dual operator from $\mathbb{L}^2(\bar{\mu})$ to $\mathbb{L}^2(\mu)$, which is Markovian in the sense that

$$
\begin{aligned}
\Lambda^*[\mathbb{1}_{\bar{V}}] &= \mathbb{1}_V \\
\forall\, f \in \mathbb{L}^2(\bar{\mu}), \quad f \geqslant 0 &\Rightarrow \Lambda^*[f] \geqslant 0
\end{aligned}
$$

where the relations have to be understood $\bar{\mu}$- or μ-a.s.

If $\bar{\mu}$ and μ give positive weights to all points of $\bar{V}$ and V respectively, then Λ^* can be seen as a Markov kernel from V to $\bar{V}$.

Remark 15 In the intertwining framework, similar considerations are valid for $\bar{P}$ and P, in order to define $\bar{P}^*$ and P^*, seen as Markov operators on $\mathbb{L}^2(\bar{\mu})$ and $\mathbb{L}^2(\mu)$, when $\bar{\mu}$ and μ are invariant probability measures, respectively for $\bar{P}$ and P, i.e. $\bar{\mu}\bar{P} = \bar{\mu}$ and $\mu P = \mu$. Thus to be able to

consider $\bar{P}^*$ and P^* as Markov matrices, it is convenient to make the following assumption: we say that the couple $(\bar{P}, \Lambda)$ is **positive**, if $\bar{P}$ admits a positive invariant measure $\bar{\mu}$ and if $\mu := \bar{\mu}\Lambda$ is also positive. Up to reducing $\bar{V}$ and V respectively to the support of $\bar{\mu}$ and μ, it is always possible to come back to this case. Note that the commutation relation

$$\bar{P}\Lambda = \Lambda P \qquad (16)$$

implies that μ is an invariant probability for P.

Under the hypotheses that $(\bar{P}, \Lambda)$ is positive and that (16) is satisfied, we get a dual commutation relation:

$$P^*\Lambda^* = \Lambda^*\bar{P}^*$$

If furthermore we assume that $(m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P})$ and that

$$\bar{m}_0\Lambda\Lambda^* = \bar{m}_0 \qquad (17)$$

then we get the intertwining relation

$$(m_0, \bar{P}^*) \prec_{\Lambda*} (\bar{m}_0, P^*)$$

The reversibility assumption for $\bar{P}$ with respect to $\bar{\mu}$ amounts to $\bar{P}^* = \bar{P}$ and similarly for P. These considerations lead to a restricted symmetry property for the relation $\prec$: $(m_0, P) \prec_\Lambda (\bar{m}_0, \bar{P})$ implies $(\bar{m}_0, \bar{P}) \prec_{\Lambda*} (m_0, P)$ under the assumptions that $(\bar{P}, \Lambda)$ is positive, that $\bar{P}$ and P are reversible and that (17) is satisfied. This is an instance of the equivalence relation $\sim$ introduced in Remark 9.

We give below in Remark 22 (b) a natural condition under which (17) is true.

□

Beyond reversibility or uniplicity, an important assumption will be

$$\Lambda\Lambda^*\bar{P}\Lambda = \bar{P}\Lambda \qquad (18)$$

(this condition for the Markov kernel $\bar{P}$ is an analogue of (17) for the probability measure $\bar{m}_0$). Define

$$P := \Lambda^*\bar{P}\Lambda \qquad (19)$$

From (18), it appears that $\bar{P}$ and P are intertwined through Λ, namely (16) is satisfied. We can go further in the exploration of $\mathcal{K}(P)$ with the help of $\mathcal{K}(\bar{P})$: the next result is a slight modification of Proposition 3 of [16], where $\mathcal{K}(\bar{P})$ was replaced by the smaller set

$$\mathcal{K}(\bar{P}, \Lambda) := \{K \in \mathcal{K}(\bar{P}) : \Lambda\Lambda^*\bar{K}\Lambda = \bar{K}\Lambda\}$$

namely the set of elements from $\mathcal{K}(\bar{P})$ satisfying the condition (18). It is also a convex semigroup and in Lemma 19 some conditions will be given so that $\mathcal{K}(\bar{P}, \Lambda) = \mathcal{K}(\bar{P})$.

Lemma 16 *Assume that $\bar{P}$ is reversible with respect to $\bar{\mu}$ and that (18) holds, then we have*

$$\Lambda^*\mathcal{K}(\bar{P})\Lambda \subset \mathcal{K}(P)$$

Proof

For any $\bar{K} \in \mathcal{K}(\bar{P})$, we compute that

$$\begin{aligned} \Lambda^*\bar{K}\Lambda P &= \Lambda^*\bar{K}\bar{P}\Lambda \\ &= \Lambda^*\bar{P}\bar{K}\Lambda \\ &= \Lambda^*\bar{P}\Lambda\Lambda^*\bar{K}\Lambda \\ &= P\Lambda^*\bar{K}\Lambda \qquad (20) \end{aligned}$$

where for the third equality, we have used the dual relation of (18) asserting that $\Lambda^* \bar{P}^* \Lambda \Lambda^* = \Lambda^* \bar{P}^*$, namely $\Lambda^* \bar{P} \Lambda \Lambda^* = \Lambda^* \bar{P}$, since $\bar{P} = \bar{P}^*$. Relation (20) shows that $\Lambda^* \bar{K} \Lambda$ belongs to $\mathcal{K}(P)$.
∎

Condition (18) seems quite strange at first view and we would have liked to only work with (16). Lemma 21 below will show this is possible when Λ is deterministic.

The motivation for Proposition 3 of [16] was to give an abstract version in the finite context of a method of Carlen, Geronimo and Loss [5] to recover the hypergroup property in the context of Jacobi polynomials, result initially due to Gasper [12, 13]. The underlying idea is equally conveyed by Lemma 16: to prove (10), one tries to find a Markov model (or several ones) $\bar{P}$, above P in the sense of intertwining (namely according to the order relation induced by $\prec$ as in Remark 9), such that $\mathcal{K}(\bar{P})$ is relatively easy to apprehend. If it appears that $\mathcal{K}(\bar{P})$ is quite big, then the inclusion of Lemma 16 gives an opportunity to show that $\mathcal{K}(P)$ is also big, leading us toward (10). But to guess such a nice Markov kernel $\bar{P}$ from P may not be an easy task! That is why we now go in the reverse direction, starting with $\bar{P}$. In particular it is natural to wonder when does

$$\mathcal{H}(\bar{P}) \neq \varnothing \tag{21}$$

imply (10). Before partially answering this question, let us mention a construction of Markov kernels satisfying (21).

Remark 17 (a) Any irreducible Markov kernel P on $\{0, 1\}$ satisfies (10). Indeed, let μ be the associated invariant measure and by symmetry, assume that $\mu(0) \leqslant \mu(1)$. Then there exists $a \in [-\mu(0)/\mu(1), 1]$ such that $P = aI + (1 - a)\mu$, where μ is seen as the Markov kernel whose two rows are equal to μ. Any Markov kernel $K := bI + (1 - b)\mu$, with $b \in [-\mu(0)/\mu(1), 1]$, belongs to $\mathcal{K}(P)$. Taking $b = -\mu(0)/\mu(1)$ (respectively $b = 1$), the first row of K is $(0, 1)$ (resp. $(1, 0)$). This shows that $0 \in \mathcal{H}(P)$.

(b) If P_1 and P_2 are two Markov kernels on V_1 and V_2, then $P_1 \otimes P_2$ is a Markov kernel on $V_1 \times V_2$. It appears that $\mathcal{K}(P_1) \otimes \mathcal{K}(P_2) \subset \mathcal{K}(P_1 \otimes P_2)$ and in particular $\mathcal{H}(P_1) \times \mathcal{H}(P_2) \subset \mathcal{H}(P_1 \otimes P_2)$.

(c) From the two points above, it follows that if P is an irreducible Markov kernel on $\{0, 1\}$, then for any $N \in \mathbb{N}$, $\bar{P} := P^{\otimes N}$ satisfies (21). Such Markov kernels were used in [16] to recover the hypergroup property of the biased Ehrenfest model (initially due to Eagleson [10]).
□

We introduce now three assumptions which are helpful in the direction of deducing (10) from (21).

First, the surjectivity of Λ as an operator on $\mathcal{P}(\bar{V})$:

$$\mathcal{P}(\bar{V})\Lambda = \mathcal{P}(V) \tag{22}$$

Second, the determinism of Λ on $\mathcal{H}(\bar{P})$:

$$\forall \, \bar{x}_0 \in \mathcal{H}(\bar{P}), \qquad \Lambda(\bar{x}_0, \cdot) = \delta_{\pi(\bar{x}_0)} \tag{23}$$

where $\pi(\bar{x}_0)$ is an element of V. Denote $\pi(\mathcal{H}(\bar{P}))$ the image by π of $\mathcal{H}(\bar{P})$. The last hypothesis is an extension of (18) to the identity kernel:

$$\Lambda \Lambda^* \Lambda = \Lambda \tag{24}$$

Note that by multiplication on the left or on the right by Λ^*, this implies that $\Lambda \Lambda^*$ and $\Lambda^* \Lambda$ are projection operators in their respective spaces $\mathbb{L}^2(\bar{\mu})$ and $\mathbb{L}^2(\mu)$.

Proposition 18 *Assume $\bar{P}$ is uniplicit and (18), (21), (22), (23) and (24) hold. Then (10) is satisfied with P given by (19) and more precisely $\pi(\mathcal{H}(\bar{P})) \subset \mathcal{H}(P)$.*

Before proving this statement, let us give another important consequence of uniplicity. If P is a Markov kernel on V, let $\mathcal{A}(P)$ be the algebra generated by P, namely the set of finite combinations of the form $a_0 I + a_1 P + a_2 P^2 + \cdots + a_n P^n$, where $n \in \mathbb{Z}_+$ and $a_0, a_1, a_2, ..., a_n \in \mathbb{R}$. Denote also by $\mathcal{K}(V)$ the convex set of Markov kernels on V.

Lemma 19 *Assume that P is uniplicit. Then we have*

$$\mathcal{K}(P) \;=\; \mathcal{A}(P) \cap \mathcal{K}(V)$$

In particular if (24) holds and $\bar{P}$ is uniplicit and satisfies (18), then the latter property can be extended to $\mathcal{K}(\bar{P})$:

$$\forall \; \bar{K} \in \mathcal{K}(\bar{P}), \qquad \Lambda\Lambda^*\bar{K}\Lambda \;=\; \bar{K}\Lambda \tag{25}$$

Proof

Let $(\varphi_1, \; \varphi_2, ..., \; \varphi_{|V|})$ be an orthonormal basis of eigenvectors of P and let $\lambda_1, \; \lambda_2, ..., \; \lambda_{|V|}$ be the corresponding eigenvalues. Consider $K \in \mathcal{K}(P)$, by commutativity, $(\varphi_1, \; \varphi_2, ..., \; \varphi_{|V|})$ is also a basis of eigenvectors of K, denote by $\theta_1, \; \theta_2, ..., \; \theta_{|V|}$ the associated eigenvalues. Since the $\lambda_1, \; \lambda_2, ..., \; \lambda_{|V|}$ are all distinct, we can find a polynomial R of degree at most $|V|$ such that

$$\forall \; l \in [\![|V|]\!], \qquad R(\lambda_l) \;=\; \theta_l$$

It follows that $K = R(P)$, showing that $\mathcal{K}(P) \subset \mathcal{A}(P) \cap \mathcal{K}(V)$. The reverse inclusion is obviously always true.

The second assertion of the lemma comes from the fact that (18) implies that

$$\forall \; n \in \mathbb{N}, \qquad \Lambda\Lambda^*\bar{P}^n\Lambda \;=\; \bar{P}^n\Lambda$$

Indeed, this is shown by iteration on $n \in \mathbb{N}$:

$$\begin{aligned}
\Lambda\Lambda^*\bar{P}^{n+1}\Lambda &= \Lambda\Lambda^*\bar{P}^n(\bar{P}\Lambda) \\
&= \Lambda\Lambda^*\bar{P}^n(\Lambda\Lambda^*\bar{P}\Lambda) \\
&= (\Lambda\Lambda^*\bar{P}^n\Lambda)\Lambda^*\bar{P}\Lambda \\
&= (\bar{P}^n\Lambda)\Lambda^*\bar{P}\Lambda \\
&= \bar{P}^n(\Lambda\Lambda^*\bar{P}\Lambda) \\
&= \bar{P}^{n+1}\Lambda
\end{aligned}$$

The case $n = 0$ corresponds to assumption (24). So we get that for any $\bar{A} \in \mathcal{A}(\bar{P})$,

$$\Lambda\Lambda^*\bar{A}\Lambda \;=\; \bar{A}\Lambda$$

from which we deduce (25) if $\bar{P}$ is uniplicit.

$\blacksquare$

Remark 20 (a) The inclusion $\mathcal{A}(P) \cap \mathcal{K}(V) \subset \mathcal{K}(P)$ is always true, but it is not necessarily an equality. Indeed, if $\mathcal{K}(P) = \mathcal{A}(P) \cap \mathcal{K}(V)$, then the elements of $\mathcal{K}(P)$ commute. But $\mathcal{S}_P$ is naturally included into $\mathcal{K}(P)$ via the representation $\mathcal{S}_P \ni g \mapsto T_g \in \mathcal{K}(V)$ where T_g is the deterministic Markov kernel given by

$$\forall \; x \in V, \qquad T_g(x, \cdot) \;=\; \delta_{g(x)}$$

If the elements of $\mathcal{K}(P)$ commute, then $\mathcal{S}_P$ is itself commutative. This is not always true, one can e.g. consider the transition kernel of the random walk generated by the transpositions on the permutation group $\mathcal{S}_N$, with $N \geqslant 3$.

(b) The example of Remark 17 is equally such that $\mathcal{S}_{\bar{P}}$ is not commutative for $N \geqslant 3$. Indeed, consider for $\sigma \in \mathcal{S}_N$ the mapping g on $\{0,1\}^N$ obtained by shuffling the coordinates according to σ. Then T_g, defined as above, belongs to $\mathcal{S}_{\bar{P}}$. It follows that $\mathcal{S}_{\bar{P}}$ contains $\mathcal{S}_N$ as a subgroup and thus cannot be commutative. Despite the fact that $\bar{P}$ is not uniplicit, it was proven in [16] that the conclusion of Proposition 7 is true, where $G := S_N$. In this case P is a birth and death chain and is thus uniplicit.

(c) Even if it outside the finite framework, the example of the Laplacian L on the sphere $\mathbb{S}^N \subset \mathbb{R}^{N+1}$, with $N \geqslant 1$, is also such that $\mathcal{K}(L)$ (rigorously, one should define it with respect to the associated heat kernel at a positive time) is not commutative, because $\mathcal{S}_L$ contains all the isometric transformations of $\mathbb{S}^N$, namely the orthogonal group $O(N+1)$. Note nevertheless that since $\mathcal{K}(L)$ is big, the same is true for $\mathcal{H}(L)$: it is the whole sphere! We mention this case, because it plays an important role in Carlen, Geronimo and Loss [5]. At first view, it has some similarities with the situation of (b) above: L is not uniplicit but formally the conclusion of Proposition 7 is true when G is the subset of $O(N+1)$ conserving the norm of the n first coordinates of $\mathbb{R}^N$, with $n \in [\![N-1]\!]$.

(d) Despite what we just said, it seems there is an important difference between the cases (b) and (c) above. In the latter it can be checked that $\mathcal{K}(L,\Lambda) \neq \mathcal{K}(L)$, while in the former we think that $\mathcal{K}(\bar{P},\Lambda) = \mathcal{K}(\bar{P})$. That is why Proposition 18 could be applied to such $\bar{P}$ without the assumption of uniplicity, thus explaining the validity of Proposition 7 for this example. In [16], it was rather used that $\mathcal{H}(\bar{P},\Lambda) \neq \varnothing$, where $\mathcal{H}(\bar{P},\Lambda) := \{x \in \{0,1\}^N : \delta_x \mathcal{K}(\bar{P},\Lambda) = \mathcal{P}(\{0,1\}^N)\}$.

□

With these observations, we can come to the

Proof of Proposition 18

Consider $x_0 \in \mathcal{H}(\bar{P})$. Taking into account (25), we have

$$\begin{aligned}
\delta_{x_0}\Lambda\Lambda^*\mathcal{K}(\bar{P})\Lambda &= \delta_{x_0}\mathcal{K}(\bar{P})\Lambda \\
&= \mathcal{P}(\bar{V})\Lambda \\
&= \mathcal{P}(V)
\end{aligned}$$

where we used (22). Assumption (23) ensures that $\delta_{x_0}\Lambda = \delta_{\pi(x_0)}$, so we get

$$\delta_{\pi(x_0)}\Lambda^*\mathcal{K}(\bar{P})\Lambda = \mathcal{P}(V)$$

Finally we use Lemma 16 to see that

$$\mathcal{P}(V) \subset \delta_{\pi(x_0)}\mathcal{K}(P)$$

which is the wanted result.

∎

It is time now to consider the purely determinist case for Λ, which simplifies most of the previous hypotheses. More precisely, assume that there exists a surjective mapping π from $\bar{V}$ to V such that Λ is given by

$$\forall\, x \in \bar{V}, \qquad \Lambda(x,\cdot) := \delta_{\pi(x)}(\cdot) \tag{26}$$

Lemma 21 *Under* (26)*, if $\bar{P}$ is a Markov kernel on $\bar{V}$ such that $(\bar{P},\Lambda)$ is positive and if P is a Markov kernel on V satisfying the intertwining relation* (16) *(called Dynkin's condition in this situation, see [9]), then* (18)*,* (22)*,* (23) *and* (24) *are true. Furthermore, $\Lambda^*\Lambda = I$ and P is given by* (19)*.*

Proof

Under Assumption (26), it was seen in Lemma 5 of [16] that $\Lambda\Lambda^*$ is the conditional expectation with respect to the sigma-algebra $\mathcal{T}$ generated by π.

Consider (18), which amounts to

$$\forall\, f \in \mathbb{L}^2(\mu), \qquad \Lambda\Lambda^*\bar{P}\Lambda[f] \;=\; \bar{P}\Lambda[f]$$

Note that the relation $\bar{P}\Lambda[f] = \Lambda P[f] = P[f] \circ \pi$ implies that $\bar{P}\Lambda[f]$ is $\mathcal{T}$-measurable for any $f \in \mathbb{L}^2(\mu)$, so the above equality holds. Similarly, using that $\Lambda[f]$ is $\mathcal{T}$-measurable for any $f \in \mathbb{L}^2(\mu)$, we get (24). It follows that $\Lambda^*\Lambda$ is a projection in $\mathbb{L}^2(\mu)$ and to see that $\Lambda^*\Lambda = I$, it is sufficient to check that $\Lambda^*\Lambda$ is injective. So let $f \in \mathbb{L}^2(\mu)$ be such that $\Lambda^*\Lambda[f] = 0$, we get that

$$f \circ \pi \;=\; \Lambda[f] \;=\; \Lambda\Lambda^*\Lambda[f] \;=\; 0$$

Since π is surjective, it appears that $f = 0$.

It follows that P is given by (19):

$$P \;=\; \Lambda^*\Lambda P \;=\; \Lambda^*\bar{P}\Lambda$$

Condition (26) implies obviously (23), and (22) due to the surjectivity of π.

∎

Remark 22 (a) The deterministic case (26) is not the only one where (18) is satisfied. Indeed, assume that π is surjective but not injective in (18). Let $\bar{P}$ be a Markov kernel on $\bar{V}$ such that $(\bar{P}, \Lambda)$ is positive. From Lemme 21, it appears that $\Lambda\Lambda^* = I$, so we get

$$\Lambda\Lambda^*\bar{P}^*\Lambda^* \;=\; \bar{P}^*\Lambda^*$$

namely (18) for $\bar{P}^*$ and Λ^*. But since π is not injective, the conditional expectation $\Lambda^*\Lambda$ is not the identity, thus Λ^* does not satisfy (26).

(b) Under Assumption (26), Condition (17) is also simple to understand: it asks that the conditional expectations with respect to $\mathcal{T}$ (the sigma-algebra generated by π) with respect to $\bar{\mu}$ and $\bar{m}_0$ coincide. Namely, if $(A_1, ..., A_l)$ is the partition of $\bar{V}$ generating $\mathcal{T}$ (corresponding to the equivalence relation between $x, y \in \bar{V}$ given by $\pi(x) = \pi(y)$), then $\bar{m}_0$ satisfies (17) if and only if it is of the form

$$\forall\, x \in \bar{V}, \qquad m_0(x) \;=\; \sum_{k \in [\![l]\!]} \frac{a_k}{\bar{\mu}(A_k)} \mathbb{1}_{A_k}(x)\bar{\mu}(x)$$

where $(a_1, ..., a_l)$ is a probability measure on $[\![l]\!]$.

□

From Proposition 18 and Lemma 21, we deduce:

Corollary 23 *Assume that the Markov kernel $\bar{P}$ is uniplicit and that $\mathcal{H}(\bar{P}) \neq \varnothing$. Let P be a Markov kernel satisfying Relation (16) with a link Λ given by (26) with π surjective. Then P is uniplicit and satisfies (10) as well as the hypergroup property.*

Proof

The above results show that $\mathcal{H}(P) \neq \varnothing$. According to Lemma 2, it is then sufficient to check that P is uniplicit. By duality, we have $P^*\Lambda^* = \Lambda^*\bar{P}^* = \Lambda^*\bar{P}$, it implies, via the equality $\Lambda^*\Lambda = I$ of Lemma 21,

$$\begin{aligned}
P^* \;&=\; P^*\Lambda^*\Lambda \\
&=\; \Lambda^*\bar{P}\Lambda \\
&=\; P
\end{aligned}$$

where we used (18), which is true due to Lemma 21 again. This shows that P is reversible. Consider θ an eigenvector of P and $\varphi, \widetilde{\varphi}$ two associated eigenvectors. From the intertwining relation (16) we get

$$\theta \Lambda[\varphi] = \bar{P}[\Lambda[\varphi]]$$

and similarly for $\widetilde{\varphi}$. By uniplicity of $\bar{P}$, $\Lambda[\varphi]$ and $\Lambda[\widetilde{\varphi}]$ are then co-linear. Remembering that Λ is injective by surjectivity of π, we get that φ and $\widetilde{\varphi}$ are co-linear as wanted.

∎

Proposition 7 is itself a consequence of the previous corollary. Indeed, it is immediate to check that $\bar{P}$, P and π given before Proposition 7 satisfy the intertwining relation (16) where Λ is defined by (26).

To end this section, we mention some (upper) semi-continuity properties associated to the Markov commutator convex semi-groups, suggesting the easy handling of this notion. Note that for any Markov kernel P on the finite set V and $x \in V$, the sets $\mathcal{K}(P)$ and $\mathcal{K}(P,x)$ are compact subsets, respectively of the set of Markov kernels and of probability measures on V (endowed with the topologies inherited from those of $\mathbb{R}^{V^2}$ and $\mathbb{R}^V$), themselves being compact. As usual, consider the Hausdorff topology on the compact subsets of a compact set, it turns it into a compact set itself. The following properties are elementary and their proofs are left to the reader.

Lemma 24 *Let $(P_n)_{n \in \mathbb{N}}$ be a sequence of Markov kernels on V converging to P. We have for any $x \in V$,*

$$\limsup_{n \to \infty} \mathcal{K}(P_n) \subset \mathcal{K}(P)$$
$$\limsup_{n \to \infty} \mathcal{K}(P_n, x) \subset \mathcal{K}(P, x)$$
$$\limsup_{n \to \infty} \mathcal{H}(P_n) \subset \mathcal{H}(P)$$

As a consequence, the set of Markov kernels P on V satisfying the generalized hypergroup property (10) is closed.

Let us remark that the above last inclusion can be strict. Anticipating a little on the next section, consider $V := \{0, 1\}$ and let $(U_n)_{n \in \mathbb{N}}$ be a sequence of functions on V satisfying $U_n(0) > U_n(1)$ for all $n \in \mathbb{N}$ and $\lim_{n \to \infty} U_n = 0$. With the notation of (6), we have

$$\forall\, n \in \mathbb{N}, \qquad \mathcal{H}(M_{U_n}) = \{0\}$$
$$\lim_{n \to \infty} M_{U_n} = M_0$$
$$\mathcal{H}(M_0) = \{0, 1\}$$

3 On the discrete Achour-Trimèche's theorem

Here the specific birth and death situation is considered in a more detailed way. The diffusive Achour-Trimèche's theorem will be partially translated into the discrete case, but first we show it cannot be extended to all convex potentials. It corresponds respectively to the proofs of Proposition 6 and Theorem 1.

The previous section provided all the ingredients necessary to the

Proof of Theorem 1

Recall the setting described in the introduction. Theorem 1 is proven by a contradictory argument: assume there exists a generalized Metropolis procedure $\mathcal{C} \ni U \mapsto P_U$ such that P_U satisfies the hypergroup property for all $U \in \mathcal{C}$.

Since $N \geqslant 2$, there exists $U \in \mathcal{C}$ such that $U(0) = U(1)$ and which is not symmetric with respect to the mapping $[\![0, N]\!] \ni x \mapsto N - x$. For $\epsilon > 0$, consider the function U_ϵ defined on $[\![0, N]\!]$ by

$$\forall \, x \in [\![0, N]\!], \qquad U_\epsilon(x) \;\; := \;\; \begin{cases} U(0) + \epsilon & \text{, if } x = 0 \\ U(x) & \text{, otherwise} \end{cases}$$

It is clear that $U_\epsilon \in \mathcal{C}$. Furthermore, due to the convexity of U and the assumption $U(0) = U(N)$, it appears that $U_\epsilon(0) > U_\epsilon(x)$ for all $x \in [\![N]\!]$. By Definition (5), the minimum of μ_{U_ϵ} is only attained at 0. Taking into account Proposition 4, it follows that $\mathcal{H}(P_{U_\epsilon}) = \{0\}$. By letting $\epsilon > 0$ go to zero, Lemma 24 implies that $0 \in \mathcal{H}(P_U)$. The same reasoning, where the value of $U(N)$ is a little increased, equally enables to conclude that $N \in \mathcal{H}(P_U)$. So we get that $\{0, N\} \subset \mathcal{H}(P_U)$. Since P_U is a birth and death, it is uniplicit, and according to Proposition 5, we can find $g \in \mathcal{S}_{P_U}$ with $g(0) = N$. Note that under the action of any element of the symmetry group $\mathcal{S}_P$, the graph of the transitions permitted by P is preserved (not taking into account the self-loops). For birth and death transitions on $[\![0, N]\!]$, this graph is the usual linear graph structure of $[\![0, N]\!]$. There are only two graph morphisms preserving this structure, the identity and the mapping $[\![0, N]\!] \ni x \mapsto N - x$. So we end up with a contradiction, because g can be neither of them.

∎

We now come to the proof of Proposition 6. We begin by reducing the problem to symmetric potentials. Recall that the classical Metropolis procedure $\mathcal{C} \ni U \mapsto M_U$ is defined by (6).

Lemma 25 *If for all $N \in \mathbb{N} \backslash \{1\}$, the Metropolis kernel M_U satisfies the hypergroup property for $U \in \widetilde{\mathcal{C}}_{\mathrm{s}}$, then it is also true for $U \in \widetilde{\mathcal{C}}_{\mathrm{m}}$.*

Proof

This is a consequence of Proposition 7. Indeed, let $U \in \widetilde{\mathcal{C}}_{\mathrm{m}}$, up to reversing the discrete segment $[\![0, N]\!]$, assume that U is non-increasing. Consider $\bar{V} := [\![0, 2N + 1]\!]$, on which we construct the potential $\bar{U}$ by symmetrization of U with respect to $N + 1/2$. Note that $\bar{U}$ is convex and more precisely that $\bar{U} \in \widetilde{\mathcal{C}}_{\mathrm{m}}$, due to the assumption $U(N - 1) - U(N) \geqslant 2\ln(2)$, which implies

$$\begin{aligned} \bar{U}(N + 2) - \bar{U}(N + 1) \;\; &\geqslant \;\; 2\ln(2) \\ &= \;\; \bar{U}(N + 1) - \bar{U}(N) + 2\ln(2) \\ &\geqslant \;\; \bar{U}(N) - \bar{U}(N - 1) + 4\ln(2) \end{aligned}$$

Associate to $\bar{U}$ the classical Metropolis kernel $\bar{M}_{\bar{U}}$ on $\bar{V}$. Let $\bar{G} = \mathcal{S}_{\bar{M}_{\bar{U}}}$ be the group consisting of the identity and of the involution $[\![0, 2N + 1]\!] \ni x \mapsto 2N + 1 - x$. The reduction presented before Proposition 7 transforms $\bar{M}_{\bar{U}}$ into M_U (up to a modification of the constant Σ_U given in (8), which has no impact on the hypergroup property, since it amounts to change M_U into a convex combination of M_U and I). Again, since $\bar{M}_{\bar{U}}$ is a birth and death chain, it is uniplicit. Thus Proposition 7 enables to see that M_U satisfies Condition (10), because by assumption this is true for $\bar{M}_{\bar{U}}$. Applying once more Lemma 2 shows that M_U satisfies the hypergroup property.

∎

Remark 26 In the above proof, another symmetrization could have been considered: let $\bar{V} := [\![0, 2N]\!]$ and $\bar{U}$ be obtained from U by symmetry with respect to N (U being non-increasing). Applying the same arguments under the relaxed assumption $U(N - 1) - U(N) \geqslant \ln(2)$ (implying $\bar{U}(N+1) - \bar{U}(N) \geqslant \bar{U}(N) - \bar{U}(N-1) + 2\ln(2)$) , we get in the end that $\widetilde{M}_U$ satisfies the hypergroup property, where $\widetilde{M}_U$ is defined as M_U in (6), but with M_0 replaced by the exploration kernel $\widetilde{M}_0$ given by

$$\forall \, x \neq y \in [\![0, N]\!], \qquad \widetilde{M}_0(x, y) \;\; := \;\; \begin{cases} 1/2 & \text{, if } |x - y| = 1 \text{ and } x \neq N \\ 1 & \text{, if } x = N \text{ and } y = N - 1 \\ 0 & \text{, otherwise} \end{cases}$$

19

It remains to prove that for $U \in \widetilde{\mathcal{C}}_{\mathrm{s}}$, M_U satisfies the hypergroup property. We did not find general arguments to obtain this result. Instead, we will adapt to the discrete case the proof presented by Bakry and Huet [3] in the context of symmetric one-dimensional diffusions.

Proposition 27 *For any $U \in \widetilde{\mathcal{C}}_{\mathrm{s}}$, the Metropolis kernel M_U satisfies the hypergroup property, with respect to the points 0 and N.*

By uniplicity of M_U and its symmetry with respect to the mapping

$$ s \;:\; [\![0, N]\!] \ni x \;\; \mapsto \;\; N - x \tag{27} $$

it is sufficient to check that $0 \in \mathcal{H}(M_U)$ for given $U \in \widetilde{\mathcal{C}}_{\mathrm{s}}$. Let us consider more generally the problem of showing that $0 \in \mathcal{H}(P)$, when P is an irreducible birth and death Markov transition on $[\![0, N]\!]$, left invariant by the symmetry s. By definition, it amounts to show that for any given probability $m_0 \in \mathcal{P}([\![0, N]\!])$, there is a Markov kernel K commutating with P and such that $K(0, \cdot) = m_0$. This question is equivalent to the fact that a wave equation starting from a non-negative condition remains non-negative, as it was shown by Bakry and Huet [3] in the diffusive situation and in Remark 6 of [16] for the discrete case. More precisely, there is a unique matrix K commuting with P such that $K(0, \cdot) = m_0$ (due to the uniplicity of M_U, see the proof of Lemma 13 or Lemma 10 of [16]), our problem is to check that its entries are non-negative. Denote $L = P - I$, the Markovian generator matrix associated to P and

$$ \forall \; x, y \in [\![0, N]\!], \qquad k(x, y) \;:=\; \frac{K(x, y)}{\mu(y)} $$

The commutation of K with P can be rewritten as the wave equation

$$ \forall \; x, y \in [\![0, N]\!], \qquad L^{(1)}[k](x, y) \;=\; L^{(2)}[k](x, y) \tag{28} $$

where for $i \in \{1, 2\}$, $L^{(i)}$ stands for the generator acting on the i-th variable as L.
Consider the discrete triangle

$$ \triangle \;:=\; \{(x, y) \in [\![0, N]\!]^2 \;:\; y \leqslant x \leqslant N - y\} $$

For $z_0 := (x_0, y_0) \in \triangle$, let $p_{z_0}^- := (p_{z_0}^-(n))_{n \in [\![0, 2y_0]\!]}$ be the path defined by iteration through

$$ p_{z_0}^-(0) \;:=\; z_0 $$

$$ \forall \; n \in [\![0, 2y_0 - 1]\!], \qquad p_{z_0}^-(n + 1) \;:=\; \begin{cases} p_{z_0}^-(n) - (0, 1) & \text{, if } n \text{ is even} \\ p_{z_0}^-(n) - (1, 0) & \text{, if } n \text{ is odd} \end{cases} $$

Note that the path $p_{z_0}^-$ stays in $\triangle$ and that $p_{z_0}^-(2y_0)$ belongs to the segment $[\![0, N]\!] \times \{0\}$.
Similarly, for $z_0 \in \triangle$, we define the path $p_{z_0}^+ := (p_{z_0}^+(n))_{n \in [\![0, 2y_0]\!]}$, which is symmetric to $p_{z_0}^-$ with respect to the axe $x = x_0$. The interest of these paths is:

Lemma 28 *Assume that the mapping $k \;:\; [\![0, N]\!]^2 \to \mathbb{R}$ satisfies the wave equation (28). Then for any $z_0 := (x_0, y_0) \in \triangle$, we have, if $y_0 \geqslant 1$,*

$$
\begin{aligned}
\omega(z_0, p_{z_0}^-(1))k(z_0) \;=\; & [\omega(z_0, p_{z_0}^-(1)) - \omega(p_{z_0}^-(1), p_{z_0}^-(2)) - \omega(p_{z_0}^+(1), p_{z_0}^+(2))]k(p_{z_0}^-(1)) \\
& + \omega(p_{z_0}^-(2y_0 - 1), p_{z_0}^-(2y_0))k(p_{z_0}^-(2y_0)) + \omega(p_{z_0}^+(2y_0 - 1), p_{z_0}^+(2y_0))k(p_{z_0}^+(2y_0)) \\
& + \sum_{n \in [\![2, 2y_0 - 1]\!]} [\omega(p_{z_0}^-(n - 1), p_{z_0}^-(n)) - \omega(p_{z_0}^-(n), p_{z_0}^-(n + 1))]k(p_{z_0}^-(n)) \\
& + \sum_{n \in [\![2, 2y_0 - 1]\!]} [\omega(p_{z_0}^+(n - 1), p_{z_0}^+(n)) - \omega(p_{z_0}^+(n), p_{z_0}^+(n + 1))]k(p_{z_0}^+(n))
\end{aligned}
$$

where for any $(z, z') := ((x, y), (x', y')) \in [\![0, N]\!]^4$, we take

$$\omega(z, z') \; := \; \begin{cases} \mu(x)\mu(y)L(x, x') & , \text{ if } z' - z \in \{(1, 0), (-1, 0)\} \\ \mu(x)\mu(y)L(y, y') & , \text{ if } z' - z \in \{(0, 1), (0, -1)\} \end{cases}$$

Proof

From the reversibility of L with respect to μ, we deduce the discrete integration by part formula: for any functions f, g on $[\![0, N]\!]$, we have

$$\mu[fL[g]] \;\; = \;\; - \sum_{0 \leqslant x < y \leqslant N} \mu(x)L(x, y)[f(y) - f(x)][g(y) - g(x)]$$

In particular, if f is the indicator function of a segment $[\![q, r]\!] \subset [\![0, N]\!]$, we get

$$\mu[\mathbb{1}_{[\![q,r]\!]}L[g]] \;\; = \;\; [g(r + 1) - g(r)]\mu(r)L(r, r + 1) + [g(q - 1) - g(q)]\mu(r)L(q, q - 1) \quad (29)$$

with the convention (Neumann boundary) that $g(-1) = g(0)$ and $g(N + 1) = g(N)$.
For $z_0 \in \triangle$, define the discrete triangle

$$\triangle(z_0) \;\; := \;\; \{(x, y) \in [\![0, N]\!]^2 \; : \; y - y_0 + x_0 - 1 \leqslant x \leqslant -y + y_0 + x_0 - 1\} \quad (30)$$

Applying (29) horizontally and vertically, we get, for k satisfying the wave equation (28),

$$\begin{aligned} 0 \;\; &= \;\; \mu^{\otimes 2}[\mathbb{1}_{\triangle(z_0)}(L^{(1)} - L^{(2)})[k]] \\ &= \;\; \sum_{e \in \partial\triangle(z_0)} dk(e)\chi(e)\omega(e) \end{aligned} \quad (31)$$

where the boundary $\partial\triangle(z_0)$ of $\triangle(z_0)$ is defined by

$$\partial\triangle(z_0) \;\; := \;\; \{(z, z') \in \triangle(z_0) \times ([\![0, N]\!]^2 \backslash \triangle(z_0)) \; : \; z' - z \in \{(1, 0), (-1, 0), (0, 1), (0, -1)\}\}$$

and where for any $e := (z, z') \in \partial\triangle(z_0)$, $\omega(e)$ was defined in the statement of the lemma and

$$\begin{aligned} dk(e) \;\; &:= \;\; k(z') - k(z) \\ \chi(z, z') \;\; &:= \;\; \begin{cases} 1 & , \text{ if } z' - z \in \{(1, 0), (-1, 0)\} \\ -1 & , \text{ if } z' - z \in \{(0, 1), (0, -1)\} \end{cases} \end{aligned}$$

It is easy (but a picture can help) that (31) can written under the form

$$\begin{aligned} 0 \;\; = \;\; & \sum_{n \in [\![0, 2y_0 - 1]\!]} [k(p_{z_0}^-(n + 1)) - k(p_{z_0}^-(n))]\omega(p_{z_0}^-(n), p_{z_0}^-(n + 1)) \\ & + \sum_{n \in [\![1, 2y_0 - 1]\!]} [k(p_{z_0}^+(n + 1)) - k(p_{z_0}^+(n))]\omega(p_{z_0}^+(n), p_{z_0}^+(n + 1)) \end{aligned} \quad (32)$$

Observe that the first sum can be transformed (via discrete integration by parts, also known as Abel's trick) into

$$\begin{aligned} & \sum_{n \in [\![0, 2y_0 - 1]\!]} [k(p_{z_0}^-(n + 1)) - k(p_{z_0}^-(n))]\omega(p_{z_0}^-(n), p_{z_0}^-(n + 1)) \\ = \;\; & k(p_{z_0}^-(2y_0))\omega(p_{z_0}^-(2y_0 - 1), p_{z_0}^-(2y_0)) - k(z_0)\omega(z_0, p_{z_0}^-(1)) \\ & - \sum_{n \in [\![1, 2y_0 - 1]\!]} k(p_{z_0}^-(n))[\omega(p_{z_0}^-(n), p_{z_0}^-(n + 1)) - \omega(p_{z_0}^-(n - 1), p_{z_0}^-(n))] \end{aligned}$$

A similar manipulation is possible for the second sum (32) and we end up with the result announced in the lemma.

$$\blacksquare$$

As a consequence, we get

Proposition 29 *Assume that P is a birth and death transition kernel on $[\![0, N]\!]$ such that*

$$\forall \; z := (x, y) \in \triangle, \qquad P(y-1, y) \;\;\geqslant\;\; P(x, x-1) + P(x, x+1)$$
$$\forall \; z := (x, y) \in \widetilde{\triangle}, \qquad P(y, y-1) \;\;\leqslant\;\; P(x-1, x) \wedge P(x+1, x)$$

where $\widetilde{\triangle}$ is the "interior" of $\triangle$:

$$\widetilde{\triangle} \;\; := \;\; \{(x, y) \in [\![0, N]\!]^2 \; : \; y+1 \leqslant x \leqslant N-y-1\}$$

Let k be a solution of (28) *such that $k(\cdot, 0)$ is non-negative. Then k remains non-negative on $\triangle$.*

Proof

We begin by showing that the condition of the proposition (which can be written identically in terms of L), implies that for any $z_0 := (x_0, y_0) \in \triangle$ and $n \in [\![2, 2y_0-1]\!]$, we have

$$\omega(p_{z_0}^-(n-1), p_{z_0}^-(n)) - \omega(p_{z_0}^-(n), p_{z_0}^-(n+1)) \;\;\geqslant\;\; 0$$
$$\omega(p_{z_0}^+(n-1), p_{z_0}^+(n)) - \omega(p_{z_0}^+(n), p_{z_0}^+(n+1)) \;\;\geqslant\;\; 0$$

It amounts to see that for any $(x, y) \in \triangle$,

$$\begin{cases} \omega((x, y), (x, y-1)) - \omega((x, y-1), (x-1, y-1)) \;\;\geqslant\;\; 0 \\ \omega((x, y), (x, y-1)) - \omega((x, y-1), (x+1, y-1)) \;\;\geqslant\;\; 0 \end{cases} \tag{33}$$

and that for any $(x, y) \in \widetilde{\triangle}$,

$$\begin{cases} \omega((x, y), (x-1, y)) - \omega((x-1, y), (x-1, y-1)) \;\;\geqslant\;\; 0 \\ \omega((x, y), (x+1, y)) - \omega((x+1, y), (x+1, y-1)) \;\;\geqslant\;\; 0 \end{cases} \tag{34}$$

Concerning (33), let $\varepsilon \in \{-1, +1\}$, we have

$$\begin{aligned} & \omega((x, y), (x, y-1)) - \omega((x, y-1), (x+\varepsilon, y-1)) \\ =\;\; & \mu(x)\mu(y)L(y, y-1) - \mu(x)\mu(y-1)L(x, x+\varepsilon) \\ =\;\; & \mu(x)\mu(y-1)[L(y-1, y) - L(x, x+\varepsilon)] \end{aligned}$$

where we used the reversibility of μ with respect to L. By the first assumed inequality, we have in particular $P(y-1, y) \geqslant P(x, x-1) \vee P(x, x+1)$, so that the last r.h.s. is non negative, as wanted. The treatment of (34) is similar, taking into account the second assumed inequality:

$$\begin{aligned} & \omega((x, y), (x+\varepsilon, y)) - \omega((x+\varepsilon, y), (x+\varepsilon, y-1)) \\ =\;\; & \mu(x)\mu(y)L(x, x+\varepsilon) - \mu(x+\varepsilon)\mu(y)L(y, y-1) \\ =\;\; & \mu(x+\varepsilon)\mu(y)[L(x+\varepsilon, x) - L(y, y-1)] \\ \geqslant\;\; & 0 \end{aligned}$$

Next we want to show that

$$\omega(z_0, p_{z_0}^-(1)) - \omega(p_{z_0}^-(1), p_{z_0}^-(2)) - \omega(p_{z_0}^+(1), p_{z_0}^+(2)) \;\;\geqslant\;\; 0$$

Writing $(x, y) := p_{z_0}^-(1)$, it means that

$$\omega((x, y+1), (x, y)) - \omega((x, y), (x-1, y)) - \omega((x, y), (x+1, y)) \;\;\geqslant\;\; 0$$

namely

$$\mu(x)\mu(y)[L(y, y+1) - L(x, x+1) - L(x, x-1)] \;\;\geqslant\;\; 0$$

22

condition which is satisfied by the first assumed inequality of the lemma (since $z_0 = (x, y + 1)$). Thus all the coefficients in front of values of k in the equality of Lemma 28 are non-negative. Assume that k does not remain non-negative on $\triangle$. We can then consider y_0 the minimal value of $y \in [\![0, N]\!]$ such that there exists $y \leqslant x \leqslant N - y$ such that $k(x, y) < 0$. Next, let x_0 the minimal value of $x \in [\![y, N - y]\!]$ such that $k(x, y_0) < 0$. In particular, $z_0 := (x_0, y_0) \in \triangle$ and $k(z_0) < 0$, fact which is in contradiction with the equality of Lemma 28, whose r.h.s. is non-negative.

∎

Assume now that P is furthermore left invariant by the symmetry s defined in (27). One important consequence is that the conclusion of Proposition 29 is valid on the whole discrete square $[\![0, N]\!]^2$:

Proposition 30 *Assume that the birth and death transition P on $[\![0, N]\!]$ is invariant by s. Let k be a solution of* (28)*. Then k is left invariant by the following symmetries of the discrete square:*

$$
\begin{aligned}
[\![0, N]\!]^2 \ni (x, y) &\mapsto (y, x) \\
[\![0, N]\!]^2 \ni (x, y) &\mapsto (N - x, N - y) \\
[\![0, N]\!]^2 \ni (x, y) &\mapsto (N - y, N - x)
\end{aligned}
$$

As a consequence, if k is non-negative on $\triangle$, then it is non-negative on $[\![0, N]\!]^2$.

Proof

Consider $\widetilde{k} : \triangle \to \mathbb{R}$ satisfying the wave equation (28) on $\overset{\circ}{\triangle}$. Extend $\widetilde{k}$ to the discrete triangle $\triangle_2 := \{(x, y) \in [\![0, N]\!]^2 : y \leqslant N - x\}$ by symmetry with respect to the line $y = x$. Let us check that $\widetilde{k}$ satisfies (28) on $\overset{\circ}{\triangle}_2 := \{(x, y) \in [\![0, N]\!]^2 : y \leqslant N - x - 1\}$. By symmetry of P, it is obvious on the image of $\overset{\circ}{\triangle}$ by the mapping $(x, y) \mapsto (y, x)$. Thus it is sufficient to show that (28) is also valid on the points $(x, x) \in \overset{\circ}{\triangle}_2$. Indeed, we compute that

$$
\begin{aligned}
&L^{(1)}[\widetilde{k}](x, x) - L^{(2)}[\widetilde{k}](x, x) \\
&= L(x, x + 1)(\widetilde{k}(x + 1, x) - \widetilde{k}(x, x)) + L(x, x - 1)(\widetilde{k}(x - 1, x) - \widetilde{k}(x, x)) \\
&\quad - L(x, x + 1)(\widetilde{k}(x, x + 1) - \widetilde{k}(x, x)) - L(x, x - 1)(\widetilde{k}(x, x) - \widetilde{k}(x, x - 1)) \\
&= 0
\end{aligned}
$$

due to the construction by symmetrization.
Next we can extend $\widetilde{k}$ to $[\![0, N]\!]^2$ by symmetrization with respect to the line $y = N - x$. The same arguments as above show that this extension satisfies (28) on $[\![0, N]\!]^2$. Observe that the mapping $\widetilde{k}$ constructed in this way is left invariant by the symmetries presented in the lemma.

Now consider $k : [\![0, N]\!]^2 \to \mathbb{R}$ a solution of (28). Let $\widetilde{k}$ be its restriction to $\triangle$. By the above construction, we extend $\widetilde{k}$ to $[\![0, N]\!]^2$ into a function also satisfying (28). Note that $k(\cdot, 0) = \widetilde{k}(\cdot, 0)$, so by uniqueness of the solution of (28) given its value on the discrete segment $\{0\} \times [\![0, N]\!]$, we get $k = \widetilde{k}$.

∎

Consider the following assumption called (H): the mappings $[\![0, \lfloor N/2 \rfloor]\!] \ni x \mapsto 2^x P(x, x + 1)$ and $[\![0, \lfloor N/2 \rfloor]\!] \ni x \mapsto P(x + 1, x)$ are respectively non-increasing and non-decreasing.

Our main result about a partial extension of Achour-Trimèche's theorem to the discrete setting can be stated as

Theorem 31 *Assume that the birth and death transition P on $[\![0, N]\!]$ is invariant by s and that (H) is fulfilled. Then P satisfies the hypergoup property with respect to 0 and N.*

Proof

According to Proposition 30, it is enough to check that (H) implies the assumption of Proposition 29. Note that in the case where N is odd, by symmetry of P through s, we have $P((N-1)/2, (N+1)/2) = P((N+1)/2, (N-1)/2)$. When N is even, we rather get $P(N/2, N/2+1) = P(N/2, N/2-1)$ and $P(N/2-1, N/2) = P(N/2+1, N/2)$. In both situations, it appears that (H) leads to

$$\forall\, y \in [\![0, \lfloor N/2\rfloor]\!] - 1]\!], \, \forall\, x \in [\![y+1, \lfloor N/2\rfloor]\!], \qquad \begin{cases} 2P(x+1,x) \leqslant 2P(x,x+1) \leqslant P(y,y+1) \\ P(y+1,y) \leqslant P(x+1,x) \leqslant P(x,x+1) \end{cases}$$

By symmetry of P through s, it follows that

$$\forall\, y \in [\![0, \lfloor N/2\rfloor]\!] - 1]\!], \, \forall\, x \in [\![y+1, N-y-1]\!], \qquad \begin{cases} P(x+1,x) + P(x,x+1) \leqslant P(y,y+1) \\ P(y+1,y) \leqslant P(x+1,x) \wedge P(x,x+1) \end{cases}$$

which is the assumption of Proposition 29.

∎

As a simple corollary we obtain Proposition 27, because P_U satisfies (H) if $U \in \widetilde{\mathcal{C}}_s$. Indeed, this condition asks for the mappings $[\![0, \lfloor N/2\rfloor]\!] \ni x \mapsto U(x) - U(x+1) + 2\ln(2)x$ and $[\![0, \lfloor N/2\rfloor]\!] \ni x \mapsto U(x+1) - U(x)$ to be respectively non-increasing and non-decreasing. This is valid, by the definition of $\widetilde{\mathcal{C}}$ given before Proposition 6.

Remarks 32

(a) One can replace the exploration kernel M_0 given in (7) by $\widehat{M}_0$ defined via

$$\forall\, x \neq y \in [\![0, N]\!], \qquad \widehat{M}_0(x,y) \;:=\; \begin{cases} 1/2 & \text{, if } |x-y| = 1,\, x \neq 0 \text{ and } x \neq N \\ 1 & \text{, if } (x,y) = (0,1) \text{ or } (x,y) = (N, N-1) \\ 0 & \text{, otherwise} \end{cases}$$

The corresponding Metropolis procedure $\widetilde{\mathcal{C}}_s \ni U \mapsto \widehat{M}_U$ (where $\widehat{M}_U$ is defined as in (6), with M_0 replaced by $\widehat{M}_0$) also satisfies the hypergroup property, because (H) is equally true for these birth and death Markovian transitions.
Taking into account Remark 26, this result can be extended to the Metropolis procedure $\widetilde{\mathcal{C}}_m \cup \widetilde{\mathcal{C}}_s \ni U \mapsto \widehat{M}_U$.

Nevertheless, due to the fact that $0 \notin \widetilde{\mathcal{C}}$, we are not able to recover that $\widehat{M}_0$ satisfies the hypergroup property, as it was shown in Example 7 of [16].

(b) For $U \in \mathcal{C}$, consider the variant classical Metropolis procedure $\widehat{M}_U$ given by

$$\forall\, x \neq y \in [\![0, N]\!], \qquad \widehat{M}_U(x,y) \;:=\; \widehat{M}_0(x,y) \exp(-(U(y) - U(x))_+) \tag{35}$$

Simulations suggest that $\widehat{M}_U$ satisfies the hypergroup property if the convex function U is either monotonous or symmetric with respect to the middle point of the discrete segment $[\![0, N]\!]$. It would be a nice discrete extension of the Achour-Trimèche's theorem, but we have not been able to prove this conjecture.

(c) The previous conjecture is not true if in (35), $\widehat{M}_0$ is replaced by M_0 (given by (7)). Indeed, consider the case $N = 2$ and $U = 0$. Let k be the solution of the corresponding wave equation (28) starting from $k(\cdot, 0) := (0, 1, 0)$. Equation (28) at point $(1,1)$ writes:

$$\frac{1}{2}(k(0,0) - k(1,0)) + \frac{1}{2}(k(2,0) - k(1,0)) \;=\; \frac{1}{2}(k(1,1) - k(1,0))$$

namely $k(1,1) = -k(1,0) = -1$. So non-negativity is not preserved by (28) and by consequence M_0 does not satisfy the hypergroup property.

In particular the assumption $U \in \widetilde{\mathcal{C}}$ is not merely technical in Proposition 6. Note this observation is not in contradiction with the conjecture given in (b).

(d) Theorem 31 enables to construct other examples of birth and death Metropolis procedures satisfying the hypergroup property. E.g. consider the exploration kernel $\widetilde{M}_0$ given by

$$\forall\, x \neq y \in [\![0, N]\!], \qquad \widetilde{M}_0(x, y) \;:=\; \begin{cases} 1/2^{x \wedge (N-x)} & \text{, if } |x - y| = 1 \\ 0 & \text{, otherwise} \end{cases}$$

Let $\check{\mathcal{C}}_{\mathrm{s}}$ be the set of potentials U symmetric with respect to $N/2$ and such that $\check{U} \in \mathcal{C}$, where

$$\forall\, x \in [\![0, N]\!], \qquad \check{U}(x) \;:=\; U(x) + \ln(2)(x \wedge (N - x))$$

Define the Markov kernel $\widetilde{M}_U$ via

$$\forall\, x \neq y \in [\![0, N]\!], \qquad \widetilde{M}_U(x, y) \;:=\; \widetilde{M}_0(x, y) \exp(-(\check{U}(y) - \check{U}(x))_+)$$

For $U \in \check{\mathcal{C}}_{\mathrm{s}}$, $\widetilde{M}_U$ satisfies (H) and admits μ_U, the Gibbs measure defined in (5), as reversible measure. Thus $\mathcal{C}_{\mathrm{s}} \ni U \mapsto \widetilde{M}_U$ is a generalized birth and death Metropolis procedure satisfying the hypergroup property. The proof of Lemma 25 enables to deduce a similar construction for monotonous potentials (for instance for convex potentials U such that $[\![0, N]\!] \ni x \mapsto U(x) + \ln(2)x$ is non-increasing).

Note that the potentials from $\check{\mathcal{C}}_{\mathrm{s}}$ are more general than those from $\hat{\mathcal{C}}_{\mathrm{s}}$, since the former ones can grow linearly (away from the middle point of the state space), while the latter ones must grow quadratically. The drawback is that $\widetilde{M}_U$ is further away from the continuous model $\partial^2 - U'\partial$ than M_U defined in (6).

□

To finish, let us mention a non-negativity preservation on edges rather than on vertices under a natural relaxation of the assumption of Proposition 29:

Proposition 33 *Assume that P is a birth and death transition kernel on $[\![0, N]\!]$ such that*

$$\forall\, z := (x, y) \in \triangle, \qquad P(y - 1, y) \;\geqslant\; P(x, x - 1) \vee P(x, x + 1)$$
$$\forall\, z := (x, y) \in \widetilde{\triangle}, \qquad P(y, y - 1) \;\leqslant\; P(x - 1, x) \wedge P(x + 1, x)$$

Let k be a solution of (28) *such that $k(\cdot, 0)$ is non-negative. Then for any $(x, y) \in \triangle$, we have*

$$(x, y + 1) \in \triangle \;\;\Rightarrow\;\; k(x, y) + k(x, y + 1) \geqslant 0$$
$$(x + 1, y) \in \triangle \;\;\Rightarrow\;\; \mu(x)k(x, y) + \mu(x + 1)k(x + 1, y) \geqslant 0$$

Proof

Note that the equality of Lemma 28 can be rewritten under the form:

$$\begin{aligned}
&\omega(z_0, p_{z_0}^-(1))[k(z_0) + k(p_{z_0}^-(1))] \\
&= \;\; \omega(p_{z_0}^-(2y_0 - 1), p_{z_0}^-(2y_0))k(p_{z_0}^-(2y_0)) + \omega(p_{z_0}^+(2y_0 - 1), p_{z_0}^+(2y_0))k(p_{z_0}^+(2y_0)) \\
&\quad + \sum_{n \in [\![1, 2y_0 - 1]\!]} [\omega(p_{z_0}^-(n - 1), p_{z_0}^-(n)) - \omega(p_{z_0}^-(n), p_{z_0}^-(n + 1))]k(p_{z_0}^-(n)) \\
&\quad + \sum_{n \in [\![1, 2y_0 - 1]\!]} [\omega(p_{z_0}^+(n - 1), p_{z_0}^+(n)) - \omega(p_{z_0}^+(n), p_{z_0}^+(n + 1))]k(p_{z_0}^+(n))
\end{aligned}$$

So the first implication of the above proposition can be shown as in the proof of Proposition 29, which enables to see that $k(z_0) + k(p_{z_0}^-(1)) \geqslant 0$, if $y_0 \geqslant 1$.

For the second implication, rather consider for $z_0 \in \triangle$ such that $z_0 + (1,0) \in \triangle$, the path $p_{z_0}^+ \coloneqq (p_{z_0}^+(n))_{n \in [\![0, 2y_0+1]\!]}$ defined by iteration through

$$
\begin{aligned}
p_{z_0}^+(0) &\coloneqq z_0 \\
\forall\, n \in [\![0, 2y_0]\!], \qquad p_{z_0}^+(n+1) &\coloneqq \begin{cases} p_{z_0}^+(n) + (1,0) & \text{, if } n \text{ is even} \\ p_{z_0}^+(n) - (0,1) & \text{, if } n \text{ is odd} \end{cases}
\end{aligned}
$$

The set $\triangle(z_0)$ defined in (30) must be modified into the "almost triangle"

$$
\triangle(z_0) \quad \coloneqq \quad \{(x,y) \in [\![0,N]\!]^2 \,:\, x \leqslant y - y_0 + x_0 - 1 \text{ and } x \leqslant -y + y_0 + x_0\}
$$

The proof of Lemma 28 then leads to

$$
\begin{aligned}
& \omega(z_0, p_{z_0}^-(1))k(z_0) + \omega(p_{z_0}^+(1), p_{z_0}^+(2))k(p_{z_0}^+(1)) \\
&= \quad \omega(p_{z_0}^-(2y_0 - 1), p_{z_0}^-(2y_0))k(p_{z_0}^-(2y_0)) + \omega(p_{z_0}^+(2y_0), p_{z_0}^+(2y_0 + 1))k(p_{z_0}^+(2y_0 + 1)) \\
&\quad + \sum_{n \in [\![1, 2y_0 - 1]\!]} [\omega(p_{z_0}^-(n-1), p_{z_0}^-(n)) - \omega(p_{z_0}^-(n), p_{z_0}^-(n+1))]k(p_{z_0}^-(n)) \\
&\quad + \sum_{n \in [\![2, 2y_0]\!]} [\omega(p_{z_0}^+(n-1), p_{z_0}^+(n)) - \omega(p_{z_0}^+(n), p_{z_0}^+(n+1))]k(p_{z_0}^+(n))
\end{aligned}
$$

The proof of Proposition 29 now implies that $\omega(z_0, p_{z_0}^-(1))k(z_0) + \omega(p_{z_0}^+(1), p_{z_0}^+(2))k(p_{z_0}^+(1)) \geqslant 0$, namely $\mu(x_0)k(z_0) + \mu(x_0 + 1)k(x_0 + 1, y_0) \geqslant 0$. ∎

The advantage of Proposition 33 over Proposition 29 is that it enables to recover by approximation (with N going to infinity) the result of Bakry and Huet [3] concerning the preservation of non-negativity by the wave equation in the context of the diffusive Achour-Trimèche theorem.

Acknowledgments:

References

[1] Abdennebi Achour and Khalifa Trimeche. Opérateurs de translation généralisée associés à un opérateur différentiel singulier sur un intervalle borné. *C. R. Acad. Sci. Paris Sér. A-B*, 288(7):A399–A402, 1979.

[2] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987.

[3] D. Bakry and N. Huet. The hypergroup property and representation of Markov kernels. In *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 295–347. Springer, Berlin, 2008.

[4] Walter R. Bloom and Herbert Heyer. *Harmonic analysis of probability measures on hypergroups*, volume 20 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1995.

[5] Eric A. Carlen, Jeffrey S. Geronimo, and Michael Loss. On the Markov sequence problem for Jacobi polynomials. *Adv. Math.*, 226(4):3426–3466, 2011.

[6] Houcine Chebli. Opérateurs de translation généralisée et semi-groupes de convolution. In *Théorie du potentiel et analyse harmonique (Journées Soc. Math. France, Inst. Recherche Math. Avancée, Strasbourg, 1973)*, pages 35–59. Lecture Notes in Math., Vol. 404. Springer, Berlin, 1974.

[7] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[8] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.

[9] E. B. Dynkin. *Markov processes. Vols. I, II*, volume 122 of *Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wissenschaften, Bände 121*. Academic Press Inc., Publishers, New York, 1965.

[10] G. K. Eagleson. A characterization theorem for positive definite sequences on the Krawtchouk polynomials. *Austral. J. Statist.*, 11:29–38, 1969.

[11] James Allen Fill. The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, 22(3):543–557, 2009.

[12] George Gasper. Positivity and the convolution structure for Jacobi series. *Ann. of Math. (2)*, 93:112–118, 1971.

[13] George Gasper. Banach algebras for Jacobi series and positivity of a kernel. *Ann. of Math. (2)*, 95:261–280, 1972.

[14] Tom Koornwinder. Jacobi polynomials. III. An analytic proof of the addition formula. *SIAM. J. Math. Anal.*, 6:533–543, 1975.

[15] Laurent Miclo. On absorption times and Dirichlet eigenvalues. *ESAIM Probab. Stat.*, 14:117–150, 2010.

[16] Laurent Miclo. On the hypergroup property. *Ann. Fac. Sci. Toulouse Math. (6)*, 26(2):417–435, 2017.

[17] Laurent Miclo and Pierre Patie. On completely monotone intertwining relations. Work in progress, 2019.

miclo@math.cnrs.fr

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