On interweaving relations

Laurent Miclo[†*] and Pierre Patie[‡]

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

[‡] School of Operations Research and Information Engineering,
Cornell University

Abstract

Interweaving relations are introduced and studied here in a general Markovian setting as a strengthening of usual intertwining relations between semigroups, obtained by adding a randomized delay feature. They provide a new classification scheme of the set of Markovian semigroups which enables to transfer from a reference semigroup and up to an independent warm-up time, some ergodic, analytical and mixing properties including the φ-entropy convergence to equilibrium, the hyperboundedness and when the warm-up time is deterministic the cut-off phenomena. We also present several useful transformations that preserve interweaving relations. We provide a variety of examples of interweaving relations ranging from classical, discrete, and non-local Laguerre and Jacobi semigroups to degenerate hypoelliptic Ornstein-Uhlenbeck semigroups and some non-colliding particle systems.

Keywords: interweaving relations, Laguerre processes, hypoelliptic diffusions, entropic convergence to equilibrium, hyperboundedness.

MSC2010: primary: 47D07, secondary: 60J25, 60J27, 46E30, 37A25, 60G18, 33C45.

Contents

*LM acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010.

1 Introduction and main results

Comparison and classification are traditional mathematical tools to transfer information from a reference object to more complex ones. The goal of this paper is to develop this framework in the study of Markov semigroups by introducing the notion of interweaving as a refinement of the usual concept of intertwining. Anticipating the formal definition given below, an interweaving relation between two Markov semigroups can be seen as a symmetric (or a two-sided) intertwining relations between them with the additional feature that the two Markovian intertwining kernels factorize one of the semigroup considered at a random time.

The recent years have witnessed the ubiquity and usefulness of intertwining relations in the study of Markov processes. Indeed, this concept which traces back to the works of Dynkin [21] and Rogers and Pitman [42] yielding, in that later case, at the relationship between a Brownian motion in $\mathbb{R}^n$ and its radial part, the Bessel process of dimension n, has been, for instance, used by Diaconis and Fill [19] in relation with strong stationary times, by Carmona, Petit and Yor [15] in relation to the so-called self-similar saw tooth-processes, extended by Patie and Savov in [36, 37] to general self-similar positive Markov processes, by Miclo [31] in connection with the algebraic concept of similarity transform, by Fill [23] for an elegant characterization of the distribution of the first passage time of some Markov chains, by Borodin and Olshanski [13, 14] for the construction of Markov processes on infinite dimensional spaces, by S. Pal and M. Shkolnikov [35] for diffusions, by Patie and Simon [39] and Patie and Zhao [41] in relation with fractional operators.

The concept of interweaving will reinforce this line of research by proposing further developments in the investigation of general Markov processes. Although additional applications can certainly be developed, we will primarily focused on the study of ergodic, analytical and mixing properties of Markov semigroups including, for instance, convergence to equilibrium in the sense of φ-entropy, hyperboundness properties and cut-off phenomena. Our range of examples will be very broad as it encompasses some discrete Markov chains, classical linear diffusions, some denegenerate hypoelliptic diffusions, stochastic dynamics on partitions and some Markov processes with jumps.

Let us now proceed with the formal definition of interweaving relations between Markov semigroups. Consider a (measurable) Markov kernel semigroup $P \coloneqq (P_t)_{t\geq 0}$ on a measurable state space $(V, \mathcal{V})$. Namely, P is a Markov kernel from $\mathbb{R}_+ \times V$ to V: for any $A \in \mathcal{V}$, the function $\mathbb{R}_+ \times V \ni (t,x) \mapsto P_t(x,A)$ is measurable and for any $(t,x) \in \mathbb{R}_+ \times V$, the mapping $\mathcal{V} \ni A \mapsto P_t(x,A)$ is a probability measure. The semigroup property asserts that for any $t, s \geq 0$, $P_t P_s = P_{t+s}$, in the sense of the composition of Markov kernels from V to V. Let now $\widetilde{P} \coloneqq (\widetilde{P}_t)_{t\geq 0}$ be another Markov semigroup on a measurable state space $(\widetilde{V}, \widetilde{\mathcal{V}})$. We say there is a (Markov) **intertwining relation** from P to $\widetilde{P}$ when there exists a Markov kernel Λ from V to $\widetilde{V}$ such that

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

It will be convenient to denote $P \overset{\Lambda}{\curvearrowright} \widetilde{P}$ this commutation property (or $P_t \overset{\Lambda}{\curvearrowright} \widetilde{P}_t$ for the relation between Markov kernels for a fixed $t \geq 0$). Such a link may not say much. For instance when $\widetilde{P}$ admits an invariant probability $\widetilde{\nu}$, (1) is satisfied by considering the Markov kernel $\Lambda = \widetilde{\nu}$ defined by

$$\forall\, x \in V, \forall\, \widetilde{A} \in \widetilde{\mathcal{V}}, \qquad \Lambda(x, \widetilde{A}) \;=\; \widetilde{\nu}(A). \tag{2}$$

The intertwining relation (1) is said to be **symmetric** when there exists another Markov kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V such that $\widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$. A more meaningful notion is the following one.

Definition 1 We say that P has an **interweaving relation** with $\widetilde{P}$ if there exist two Markov kernels Λ and $\widetilde{\Lambda}$ and a non-negative random variable τ such that

$$P \overset{\Lambda}{\curvearrowright} \widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P \tag{3}$$

$$\Lambda\widetilde{\Lambda} \;=\; P_\tau = \int_0^\infty P_t \,\mathbb{P}(\tau \in dt). \tag{4}$$

We call τ the **warm-up time** or the **delay** and we write $P \hookleftarrow\!\!\!\!\rightarrow \widetilde{P}$ or $P \overset{\tau}{\hookleftarrow\!\!\!\!\rightarrow} \widetilde{P}$ to emphasize the dependency on τ. Note that when $\tau = \delta_{t_0}$ is the degenerate random variable at $t_0 > 0$, we may simply write, when there is no confusion, $P \overset{t_0}{\hookleftarrow\!\!\!\!\rightarrow} \widetilde{P}$.
When τ is in addition infinitely divisible we say that P admits an **interweaving relation with an infinitely divisible** warm-up time (for short **IRID**) with $\widetilde{P}$ and we write $P \overset{\boldsymbol{\tau}}{\hookleftarrow\!\!\!\!\rightarrow} \widetilde{P}$.
Finally, when we also have

$$\widetilde{\Lambda}\Lambda = \widetilde{P}_\tau \tag{5}$$

we say that there is a **symmetric interweaving relation** between P and $\widetilde{P}$ and we write $P \overset{\tau}{\leftrightsquigarrow} \widetilde{P}$ (resp. $P \overset{\boldsymbol{\tau}}{\leftrightsquigarrow} \widetilde{P}$ when τ is infinitely divisible).

□

Note that due to our measurability assumption above on the kernel P, the integrand in the r.h.s. of (4) is measurable with respect to $t > 0$ and the identity can be understood as the Markov kernel on $(V, \mathcal{V})$ defined by

$$\forall\, x \in V, \forall\, A \in \mathcal{V}, \qquad P_\tau(x, A) \;\coloneqq\; \int_0^{+\infty} P_t(x, A)\,\mathbb{P}(\tau \in dt)$$

The notion of interweaving is related to completely monotone functions. Indeed, observe that

$$P_\tau = \int_0^\infty e^{-tL}\mathbb{P}(\tau \in dt) = F(L) \tag{6}$$

where L is the infinitesimal generator of P and F as the Laplace transform of positive measure is, by Bochner classical result, a completely monotone function, i.e. $F \in C^{\infty}(\mathbb{R}_+)$ and $(-1)^n F^{(n)}(x) \geq 0$ for all $n \in \mathbb{N}$ and $x \geq 0$. Next, we recall that a random variable τ is said to be infinitely divisible if for each $N \in \mathbb{N}$, there exits a sequence of i.i.d. random variables $(\tau_n)_{1 \leq n \leq N}$ such that, in distribution, $\tau \overset{(d)}{=} \tau_1 + \ldots \tau_n$. Note that when τ is in addition infinitely divisible then there exists a Bernstein function ϕ, i.e. $\phi(0) \geq 0$ and ϕ' is completely monotone, such that, in (6) above, $F = e^{-\phi}$. Moreover, in such a case, there exists an unique convolution semigroups on $\mathbb{R}^+$ whose transition kernel is the law of a subordinator $\boldsymbol{\tau} = (\tau_t)_{t \geq 0}$, a non-decreasing Lévy process, such that $\tau \overset{(d)}{=} \tau_1$ and $P^{\boldsymbol{\tau}} = (P^{\boldsymbol{\tau}}_t)_{t \geq 0}$ is a Markov semigroup, where for any bounded Borelian function f and $t \geq 0$,

$$P^{\boldsymbol{\tau}}_t f = \int_0^{\infty} P_s f \, \mathbb{P}(\tau_t \in ds). \tag{7}$$

$P^{\boldsymbol{\tau}}$ is the subordination of P in the sense of Bochner and we have $P^{\boldsymbol{\tau}}_1 = P_{\tau}$. The definition of interweaving can be summarized by the following commutative diagram (suggesting the name of interweaving), holding for every $t \geq 0$:

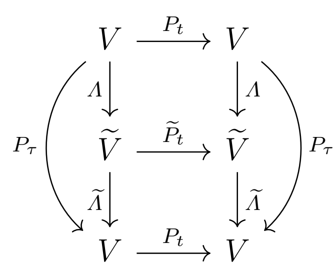

Figure 1: Interweaving relations with $\Lambda\widetilde{\Lambda} = P_{\tau}$

Our objective in this paper is to provide some properties and investigate some applications of interweaving relations in the study of probabilistic and analytical properties of general Markov processes. Before presenting its range of applications, let us present a few general observations about this concept.

Some general comments on interweaving relations

(a) The above Markov framework is quite plain. There are several ways to enrich it, especially to associate a generator L to the semigroup P, since this is in general the simplest way to describe P. Analytically, the semigroup P can be acting on a Banach space, in the sense of Hille-Yosida theory, see e.g. the book of Yosida [50]. One standard choice, when P admits an invariant probability ν, is to consider the Hilbert space $\mathbf{L}^2(\nu)$. Another possibility, when the state space V is endowed with a σ-compact topology, is to consider the space of continuous functions vanishing at infinity, endowed with the supremum norm.
From a probabilistic point of view, the generator L appears in the formulation of an underlying martingale problem for the trajectories $X := (X_t)_{t \geq 0}$ of an associated Markov process (cf. for instance the book of Ethier and Kurtz [22]). Usually the state space V is endowed with a topology and the trajectories are càdlàg, in particular the position X_t converges to X_0 as t goes to 0_+.
The examples considered in this paper will be described through their generators. All will admit an invariant measure which will be a probability measure, except for the squared Bessel processes and some related examples, and thus the $\mathbf{L}^2$ setting and the martingale problems will be equivalent.

As t goes to zero and in the appropriate senses dictated by the above analytical or probabilist frameworks, P_t converges to the identity operator Id, seen as the transition kernel corresponding to no motion.

(b) When the generators L and $\widetilde{L}$ are available for the semigroups P and $\widetilde{P}$, e.g. in one of the meanings seen in (a), the intertwining relation (1) is often equivalent to $L\Lambda = \Lambda\widetilde{L}$, where the Markov kernel Λ has to be seen as an operator from $\mathbf{D}(\widetilde{L})$ to (a subset of) $\mathbf{D}(L)$, the respective domains of the generators. When the intertwining relation is symmetric, see (5), we should have that the image of $\mathbf{D}(L)$ by $\widetilde{\Lambda}$ is included in $\mathbf{D}(\widetilde{L})$, in particular for interweaving relations, the l.h.s. of (4) can also be seen as an operator from $\mathbf{D}(L)$ to itself which can be "extended" into P_τ, a priori acting on $\mathbf{B}(V)$, the space of bounded measurable functions on V.

(c) One way to avoid the case (2) is to ask for Λ to be one-to-one, e.g. as an operator from $\mathbf{B}(\widetilde{V})$ to $\mathbf{B}(V)$ (but when $\widetilde{V}$ is not discrete, this is often requiring too much). Somewhat the requirement (4) also goes in this direction: in the "regular" situations described above in (a), P_t converges to Id for small $t > 0$ and thus should end up being invertible in this asymptotic. This should still be true for P_τ when τ has a distribution concentrated near 0 and in particular Λ would be one-to-one and $\widetilde{\Lambda}$ would be surjective. In the case of a symmetric interweaving relation with a warm-up variable τ on $\mathbb{R}_+$ concentrated near 0, we can expect Λ and $\widetilde{\Lambda}$ to be both invertible. That is why, more generally and heuristically, we see symmetric interweaving as a Markovian formulation of a weak invertibility assumption on Λ and $\widetilde{\Lambda}$, resulting in P and $\widetilde{P}$ being closely related. In the same spirit, the more mass the law of τ gives to neighborhoods of 0_+, the more informative (4) is, as the "invertibility of P_τ should be stronger". Conversely, assuming that P is ergodic with invariant probability measure ν, we have that for large $t \geq 0$, P_t is converging to ν (seen as a Markov kernel as in (2)). It follows that the more the law of τ is concentrated on large values, the less informative (4) becomes. This interpretation will be strengthened when we will see τ as a random warm-up time.

(d) From a spectral point of view and in the regular settings of (a), the meaning of an interweaving relation from P to $\widetilde{P}$ seems to be that the spectrum of the generator L of P is included into the spectrum of the generator $\widetilde{L}$ of $\widetilde{P}$, at least under appropriate ergodicity assumptions on P and when the spectrum is be understood in an extended sense. We will not enter into the underlying technicalities here, so let us just mention a conjecture that we hope to investigate in a future work:

Conjecture 2 Consider two irreducible Markov generators L and $\widetilde{L}$ on finite state spaces V and $\widetilde{V}$. There exists a interweaving relation from $(\exp(tL))_{t\geq 0}$ to $(\exp(tL))_{t\geq 0}$ if and only if the extended spectrum of L is included into that of $\widetilde{L}$. By extended spectrum, we mean the eigenvalues as well as the dimensions of the associated Jordan blocks (inclusion implying smaller or equal dimensions).

□

Such a result and possible extensions to more general state spaces would provide a spectral understanding of why interweaving relations enable to deduce quantitative informations on the convergence to equilibrium for P from similar knowledge from $\widetilde{P}$.

(e) Assume an intertwining relation $P \overset{\Lambda}{\curvearrowright} \widetilde{P}$ and that P and $\widetilde{P}$ admit reversible probability measures μ and $\widetilde{\mu}$, with $\mu\Lambda = \widetilde{\mu}$. Working in the $\mathbf{L}^2$ framework mentioned above in (a), we get by duality an intertwining relation $\widetilde{P} \overset{\Lambda^*}{\curvearrowright} P$. A priori $\Lambda^* : \mathbf{L}^2(\mu) \to \mathbf{L}^2(\widetilde{\mu})$ is an abstract Markov operator, in the sense that it preserves non-negativity and the function always taking the value 1. To get a ($\mathbf{L}^2$-)interweaving relation, it remains to check that $\Lambda\Lambda^* = P_\tau$. Thus in such a reversible setting, interweaving relations are relatively easy to deduce from intertwining relations.

(f) Assume that we have a symmetric intertwining relation between two semigroups P and $\widetilde{P}$, namely $P \overset{\Lambda}{\curvearrowright} \widetilde{P}$ and $\widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$ for some Markov kernels Λ and $\widetilde{\Lambda}$. Then necessary $\Lambda\widetilde{\Lambda}$ commutes with all the P_t for $t \geq 0$. Assume that P admits a generator L which is diagonalizable with eigenvalues of multiplicities one. When functional calculus is available, we deduce that $\Lambda\Lambda^*$ is of the form $F(-L)$, where $F : \mathbb{R}_+ \to \mathbb{R}$ is a measurable mapping. To get a interweaving relation is then equivalent to F being completely monotone.

(g) The symmetric interweaving relation does not correspond to the symmetrization of the interweaving relation, which is only requiring two interweaving relations, one from P to $\widetilde{P}$ and one from $\widetilde{P}$ to P. For the latter, the kernels from $\widetilde{V}$ to V and from V to $\widetilde{V}$ may be different from $\widetilde{\Lambda}$ and Λ, as well as the warm-up time from τ. Some results below can be extended from symmetric to symmetrized interweaving relations But the notion of symmetric interweaving relation is natural because of Proposition 4 below.

1.1 Basic properties of interweaving relations

We present now some useful transformations of semigroups that preserve interweaving relations and postpone their proofs to Section 4. We start with the following result that enables to construct from an IRID with a random warm-up time a interweaving relation with the constant 1 as warm-up time. This observation will be useful in some applications of interweaving relations for which the assumption of deterministic warm-up time is required.

Theorem 3 *Assume that $P \overset{\tau}{\hookrightarrow} \widetilde{P}$, that is the warm-up time τ is infinitely divisible. Then $P^{\boldsymbol{\tau}} \overset{1}{\hookrightarrow} \widetilde{P}^{\boldsymbol{\tau}}$ where $\boldsymbol{\tau} = (\tau_t)_{t\geq 0}$ is the subordinator such that $\tau \overset{(d)}{=} \tau_1$ and the subordinated semigroups are defined as in* (7).

We point out that in Section 2 (resp. Section 3), we present several examples for which the warm-up time τ is a constant (resp. a positive infinitely divisible random variable). In the applications of interweaving relations to ergodic properties, the previous result allows us to compare the approach based on interweaving relations with the classical ones based on functional inequalities.

We proceed with additional properties of interweaving relations. To simplify the forthcoming discussion, we assume that P (resp. $\widetilde{P}$) is a semigroup on some Banach space $\mathbf{B}$ (resp. $\widetilde{\mathbf{B}}$), e.g. if P is a Feller semigroup then $\mathbf{B} = \mathrm{C}_b(V)$ the space of continuous and bounded functions on V endowed with the supremum topology.

Let us now come to symmetric interweaving relations. They are a consequence of interweaving relations under a seemingly mild additional assumption:

Proposition 4 *When the Markov kernel Λ is one-to-one, say from $\mathbf{B}(\widetilde{V})$ to $\mathbf{B}(V)$, then a interweaving relation is symmetric.*

Proof: Indeed, from (4), we deduce, first for a non-negative Borelian function f and then for a general Borelian function f, by writing $f = \max(f, 0) - \max(-f, 0)$, that

$$\Lambda\widetilde{\Lambda}\Lambda f \;=\; P_\tau \Lambda f = \int_0^{+\infty} P_t \Lambda f \, \mathbb{P}(\tau \in dt) = \int_0^{+\infty} \Lambda \widetilde{P}_t \, \mathbb{P}(\tau \in dt) f = \Lambda \widetilde{P}_\tau f$$

where we used Tonnelli theorem for the last identity. The injectivity of Λ implies that $\widetilde{\Lambda}\Lambda = \widetilde{P}_\tau$.
■

The one-to-one assumption of Proposition 4 is quite restrictive, when the state spaces are not denumerable. Nevertheless, the simplicity of the above proof shows it can be weakened when working in the Hille-Yosida framework mentioned in Remark 1(a), by considering the corresponding notion of injectivity, in particular in $\mathbf{L}^2$ spaces.

We now proceed by showing that, under mild conditions, $\hookleftarrow$ is an equivalence relation. This highlights the idea, triggered by this concept, of an original classification scheme which enables to extend in a natural way to general Markov semigroups some ergodic and analytical properties that were attainable only for some specific classes, such as reversible diffusion ones.

Theorem 5 *Assume that the Markov intertwining kernels is one-to-one on a dense subset of* $\mathbf{B}$ *then* $\hookleftarrow$ *is an equivalence relation as*

(i) $\hookleftarrow$ *is reflexive, that is* $P \overset{0}{\hookleftarrow} P$ *with* 0 *the degenerate variable at* 0.

(ii) $\hookleftarrow$ *is symmetric, that is if* $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ *then* $\widetilde{P} \overset{\tau}{\hookleftarrow} P$ *with* $\widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P \overset{\Lambda}{\curvearrowright} \widetilde{P}$ *and* $\widetilde{\Lambda}\Lambda = \widetilde{P}_\tau$.

(iii) $\hookleftarrow$ *is transitive, that is if* $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ *and* $\widetilde{P} \overset{\widetilde{\tau}}{\hookleftarrow} \overline{P}$ *then* $P \overset{\tau+\widetilde{\tau}}{\hookleftarrow} \overline{P}$, *where* τ *and* $\widetilde{\tau}$ *are assumed to be independent.*

Remark 6 *It is not difficult to check that if one restricts the previous theorem to the subset of IRID then* $\hookleftarrow$ *remains an equivalence relation.*

Schematically, the transitivity property of interweaving relations can be described, for any $t \geq 0$, where by rotating to 45 degrees the figure of our previous diagrams:

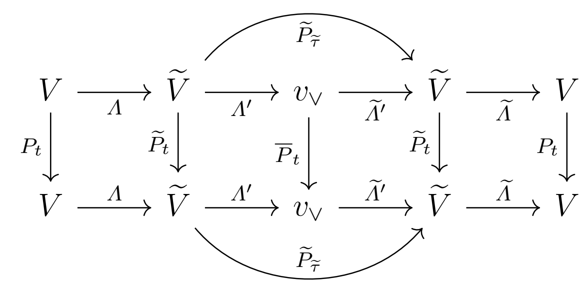

Figure 2: Transitive interweaving relations

We proceed with the following theorem that provides a closure property of interweaving relations by similarity transform as well as a way to *transport* interweaving relations.

Theorem 7 *Let us assume that* $P \overset{\tau}{\hookleftarrow} \widetilde{P}$.

1) Let P^M *be a Markov semigroup acting on the Banach space* $\mathbf{B}^M$ *If the two Markov* P *and* P^M *are similar, that is, for all* $t \geq 0$, $P_t^M = MP_tM^{-1}$ *where* M *and its inverse* M^{-1} *are bounded operators. Then,*

$$P^M \overset{\tau}{\hookleftarrow} \widetilde{P}$$

where, with the obvious notation $\Lambda^M = M\Lambda$ *and* $\widetilde{\Lambda}^M = \widetilde{\Lambda}M^{-1}$.

2) If

$$P \overset{T}{\curvearrowright} \mathbb{P} \text{ and } \widetilde{P} \overset{T}{\curvearrowright} \widetilde{\mathbb{P}} \tag{8}$$

with $\mathbb{P}$ *and* $\widetilde{\mathbb{P}}$ *two Markov semigroups defined on the measurable space* $(\mathbb{V}, \mathscr{V})$ *and* T *an one-to-one Markov operator. Then*

$$\mathbb{P} \overset{\tau}{\hookleftarrow} \widetilde{\mathbb{P}} \qquad \text{and} \qquad \mathbb{P}_\tau = \boldsymbol{\Lambda}\widetilde{\boldsymbol{\Lambda}}$$

where

$$\Lambda \overset{T}{\curvearrowright} \boldsymbol{\Lambda} \text{ and } \Lambda \overset{T}{\curvearrowright} \widetilde{\boldsymbol{\Lambda}}. \tag{9}$$

1.2 Applications of interweaving relations to the theory of Markov semigroups

We now turn to the description of some interesting features and applications of interweaving relations. Thorughout this section, we make the hypothesis that P and $\widetilde{P}$ admit ν and $\widetilde{\nu}$ as invariant probability measures, respectively, and $P \overset{\tau}{\hookleftarrow} \widetilde{P}$. In this case, $\nu\Lambda$ is also an invariant probability measure for $\widetilde{P}$, as shown by multiplying (1) on the left by ν. Similarly, $\widetilde{\nu}\widetilde{\Lambda}$ is invariant for P. We will assume that $\widetilde{\nu} = \nu\Lambda$ and that $\nu = \widetilde{\nu}\widetilde{\Lambda}$, when the invariant probability measures are not unique.

1.2.1 Entropy convergence to equilibrium

We want to deduce estimates on the speed of convergence of P to the equilibrium ν by taking into account a similar knowledge for $\widetilde{P}$ and $\widetilde{\nu}$. First we must specify the way to measure how far a probability measure m on V is from ν and here we choose the entropy (see Subsection 4.2.1 for extension of the result to general φ-entropy). The **(relative) entropy** of m with respect to ν is given by

$$\mathrm{Ent}(m|\nu) \quad := \quad \begin{cases} \displaystyle\int \ln\left(\frac{dm}{d\nu}\right) dm, & \text{if } m \ll \nu \\ +\infty, & \text{otherwise} \end{cases}$$

where $dm/d\nu$ stands for the Radon-Nikodym density of m with respect to ν. As desired, the quantity $\mathrm{Ent}(m|\nu)$ measures the discrepancy between m and ν, in particular we have the Pinsker's bound:

$$\mathrm{Ent}(m|\nu) \quad \geq \quad 2\,\|m-\nu\|_{\mathrm{tv}}^2$$

where the **total variation** distance $\|m-\nu\|_{\mathrm{tv}}$ between m and ν is defined as the supremum of $m(A) - \nu(A)$ over $A \in \mathcal{V}$.

We proceed by assuming that we have some information about the convergence of $\widetilde{P}$ towards $\widetilde{\nu}$, under the following form: there exists a function $\varepsilon : \mathbb{R}_+ \times \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$, with $\overline{\mathbb{R}}_+ := \mathbb{R}_+ \sqcup \{+\infty\}$, which is non-decreasing with respect to the second variable, such that

$$\forall\, \widetilde{m}_0 \in \mathcal{P}(\widetilde{V}),\, \forall\, t \geq 0, \qquad \mathrm{Ent}(\widetilde{m}_0\widetilde{P}_t|\widetilde{\nu}) \quad \leq \quad \varepsilon(t, \mathrm{Ent}(\widetilde{m}_0|\widetilde{\nu})) \tag{10}$$

where $\mathcal{P}(\widetilde{V})$ is the set of all probability measures on $\widetilde{V}$. For this bound to be meaningful, we furthermore require that

$$\forall\, E \in \mathbb{R}_+, \qquad \lim_{t\to+\infty} \varepsilon(t, E) \quad = \quad 0$$

A typical instance of (10) is when $\widetilde{P}$ satisfies (modified) logarithmic Sobolev inequalities (here and below, we refer for instance to the book of Ané et al. [4] for a friendly presentation of these inequalities). Then there exists a constant $\widetilde{\alpha} > 0$ such that (10) holds with the function ε given by

$$\forall\, t \geq 0,\, \forall\, E \in \overline{\mathbb{R}}_+, \qquad \varepsilon(t, E) \quad = \quad \exp(-\widetilde{\alpha}t)E$$

Here is the transfer of the entropic convergence estimate to P:

Theorem 8 *Assume that* $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ *and that* (10) *holds. Then we have*

$$\forall\, m_0 \in \mathcal{P}(V),\, \forall\, t \geq 0, \qquad \mathrm{Ent}(m_0P_{t+\tau}|\nu) \quad \leq \quad \varepsilon(t, \mathrm{Ent}(m_0|\nu)) \tag{11}$$

From a probabilistic point of view (see Remark 1(a) or the definition of a measurable Markov process below), $m_0P_{t+\tau}$ is the distribution of $X_{t+\tau}$, where τ is a random time independent of X and distributed according to τ. The bound (11) says that up to waiting a random warm-up time τ, we get for P the same estimate on the speed of convergence to equilibrium as for $\widetilde{P}$.

1.2.2 Hypercontractivity

Another famous classical application of logarithmic Sobolev inequalities concerns hypercontractivity, which is a kind of regularizing property. Interweaving relations equally enable its transfer from a semigroup to another one, up to a random warm-up time. More precisely, the **hypercontractivity** property of the semigroup $\widetilde{P}$, is the existence of a constant $\widetilde{\alpha} > 0$ (which may be different from the one considered above, for Markov processes which are not diffusions), such that we have for the operator norms,

$$\forall\ t \geq 0, \qquad \|\widetilde{P}_t\|_{\mathbf{L}^2(\widetilde{\nu}) \to \mathbf{L}^{p(\widetilde{\alpha} t)}(\widetilde{\nu})} \ \leq \ 1 \tag{12}$$

where

$$\forall\ t \geq 0, \qquad p(\widetilde{\alpha} t) \ \coloneqq \ 1 + \exp(\widetilde{\alpha} t)$$

Here is the analogue of Theorem 8 for hypercontractivity:

Theorem 9 *Assume that $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ and that* (12) *holds. Then we have*

$$\forall\ t \geq 0, \qquad \|P_{t+\tau}\|_{\mathbf{L}^2(\nu) \to \mathbf{L}^{p(\widetilde{\alpha} t)}(\nu)} \ \leq \ 1 \tag{13}$$

1.2.3 Cut-off phenomenon

Coming back to the convergence to equilibrium, we now explain how a symmetric interweaving relation enables the transfer of the cut-off phenomenon (for a short survey of this notion, see Diaconis [18]). To state our result, we need a family $(P^{(n)})_{n\in\mathbb{Z}_+}$ of Markov semigroups on state spaces $(V^{(n)})_{n\in\mathbb{Z}_+}$ with respective invariant probability measures $(\nu^{(n)})_{n\in\mathbb{Z}_+}$. Defining, for any $n \in \mathbb{Z}_+$,

$$\forall\ t \in \mathbb{R}_+, \qquad \mathfrak{d}^{(n)}(t) \ \coloneqq \ \sup_{m_0 \in \mathcal{P}(V^{(n)})} \left\| m_0 P_t^{(n)} - \nu^{(n)} \right\|_{\mathrm{tv}} \tag{14}$$

we say that the family $(P^{(n)})_{n\in\mathbb{Z}_+}$ has

(1) a (uniform) **cut-off** at the positive **cut-off times** $(t^{(n)})_{n\in\mathbb{Z}_+}$ when for any $r \in (0,1) \sqcup (1,+\infty)$,

$$\lim_{n\to\infty} \mathfrak{d}^{(n)}(rt^{(n)}) \ = \ \mathbb{1}_{\{0<r<1\}}$$

(2) a **window cut-off** (resp. **profile cut-off**) at $(t^{(n)}, w^{(n)})_{n\in\mathbb{Z}_+}$ (resp. and with profile η) if $t^{(n)} \to \infty$, $w^{(n)} = o(t^{(n)})$ as $n \to \infty$, and

$$\lim_{c\to-\infty} \varliminf_{n\to\infty} \mathfrak{d}^{(n)}(t^{(n)} + cw^{(n)}) = 1 \text{ and } \lim_{c\to+\infty} \varlimsup_{n\to\infty} \mathfrak{d}^{(n)}(t^{(n)} + cw^{(n)}) = 0$$

(resp. and for all $c \in \mathbb{R}$, $\eta(c) = \varliminf_{n\to\infty} \mathfrak{d}^{(n)}(t^{(n)} + cw^{(n)}) = \varlimsup_{n\to\infty} \mathfrak{d}^{(n)}(t^{(n)} + cw^{(n)})$).

With these definitions we have the following result.

Theorem 10 *Consider two families of Markov semigroups $(P^{(n)})_{n\in\mathbb{Z}_+}$ and $(\widetilde{P}^{(n)})_{n\in\mathbb{Z}_+}$ on $(V^{(n)})_{n\in\mathbb{Z}_+}$ and $(\widetilde{V}^{(n)})_{n\in\mathbb{Z}_+}$ and with invariant probability distributions $(\nu^{(n)})_{n\in\mathbb{Z}_+}$ and $(\widetilde{\nu}^{(n)})_{n\in\mathbb{Z}_+}$, respectively. Let $(t^{(n)})_{n\in\mathbb{Z}_+}$ be a sequence of positive real numbers and assume that for any $n \in \mathbb{Z}_+$, $P \overset{\boldsymbol{t}_0^{(n)}}{\leftrightsquigarrow} \widetilde{P}$ such that*

$$\lim_{n\to\infty} \frac{\boldsymbol{t}_0^{(n)}}{t^{(n)}} \ = \ 0 \quad (\textit{resp. } \varlimsup_{n\to\infty} \frac{\boldsymbol{t}_0^{(n)}}{w^{(n)}} = 0) \tag{15}$$

Then the cut-off (resp. ***window cut-off*** *and* ***profile cut-off****) phenomenon with cut-off times $(t^{(n)})_{n\in\mathbb{Z}_+}$ (resp. windows $(w^{(n)})_{n\in\mathbb{Z}_+}$ and profile η) for $(P^{(n)})_{n\in\mathbb{Z}_+}$ is equivalent to that of $(\widetilde{P}^{(n)})_{n\in\mathbb{Z}_+}$.*

The remaining part of the paper is organized as follows. In the two forthcoming sections we describe several examples of interweaving relations along with their applications. More specifically, in the next section, we focus on interweaving relations where the warm-up distribution is a Dirac mass: this includes the two points space and the intertwining relations between continuous and discrete Bessel and Laguerre processes and some degenerate hypoelliptic Ornstein-Uhlenbeck processes. In Section 3, we consider interweaving relations between diffusive Laguerre processes of different parameters, as well as some semigroups associated to Markov processes with jumps. Finally we prove extensions of the statements presented in this introduction in Section 4.

2 Deterministic warm-up time examples

Three examples of interweaving relations whose warm-up times are deterministic are presented in the following subsections: there exists $t_0 \geq 0$ such that $\tau = \delta_{t_0}$. In this situation the statements of Theorems 8 and 9 simplify, as (11) and (13) are respectively replaced by

$$\forall\, m_0 \in \mathcal{P}(V),\, \forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 P_{t_0+t}|\mu) \;\leq\; \varepsilon(t, \mathrm{Ent}(m_0|\mu))$$

and

$$\forall\, t \geq 0, \qquad \|P^{(\beta)}_{t_0+t}\|_{\mathbf{L}^2(\mu)\to\mathbf{L}^{p(\widetilde{\alpha}t)}(\mu)} \;\leq\; 1$$

2.1 The two point space

Consider the simplest non-trivial case of the setting of the introduction, where $V = \widetilde{V}$ is the two point space $\{0,1\}$. Let L and $\widetilde{L}$ be two isospectral irreducible Markov generators on V. We can write

$$L \;=\; \lambda(\mu - \mathrm{Id})$$

where $\lambda > 0$ is the non-zero eigenvalue of $-L$, μ is the invariant probability of L, seen as a Markov kernel, and Id is the identity operator. Any non-zero function φ on V such that $\mu[\varphi] = 0$ is an eigenfunction of L associated to the eigenvalue $-\lambda$. Consider the function φ normalized in $\mathbf{L}^2(\mu)$ given by

$$\varphi \;:=\; \begin{pmatrix} \varphi(0) \\ \varphi(1) \end{pmatrix} \;:=\; \begin{pmatrix} l \\ -1/l \end{pmatrix} \qquad \text{with } l \;:=\; \sqrt{\frac{\mu(1)}{\mu(0)}} \tag{16}$$

Since $\widetilde{L}$ is irreducible and isospectral with L, it can be written $\lambda(\widetilde{\mu} - \mathrm{Id})$, where $\widetilde{\mu}$ is the invariant probability of $\widetilde{L}$. Define $\widetilde{\varphi}$ as in (16), with μ replaced by $\widetilde{\mu}$.

For $\epsilon > 0$, define Λ_ϵ the linear mapping sending $\widetilde{\varphi}$ to $\epsilon\varphi$ and preserving the function $\mathbb{1}$. It is immediate to check that $L \overset{\Lambda_\epsilon}{\curvearrowright} \widetilde{L}$. A priori Λ_ϵ is not a Markov kernel. Nevertheless its matrix in the basis $(\mathbb{1}_{\{0\}}, \mathbb{1}_{\{1\}})$ is of the form $\begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix}$ and we compute that

$$a \;=\; \frac{1+\epsilon l\widetilde{l}}{1+\widetilde{l}^2}$$

$$b \;=\; \frac{1-\epsilon\widetilde{l}/l}{1+\widetilde{l}^2}$$

It follows that for $\epsilon > 0$, Λ_ϵ is Markovian if and only if

$$\epsilon \quad \leq \quad \min(l/\widetilde{l}\,,\,\widetilde{l}/l) \tag{17}$$

Choose $\epsilon_0 \coloneqq \min(l/\widetilde{l}\,,\,\widetilde{l}/l)$, the largest value such that Λ_{ϵ_0} is Markovian. Symmetrically, for $\widetilde{\epsilon} > 0$, construct $\widetilde{\Lambda}_{\widetilde{\epsilon}}$ sending φ to $\widetilde{\epsilon}\widetilde{\varphi}$ and preserving $\mathbb{1}$. We have $\widetilde{L} \overset{\widetilde{\Lambda}_{\widetilde{\epsilon}}}{\curvearrowright} L$ and by symmetry of the r.h.s. of (17), $\widetilde{\Lambda}_\epsilon$ is Markovian for $\widetilde{\epsilon} \in (0, \epsilon_0]$. Again choose $\widetilde{\epsilon} = \epsilon_0$, the largest value such that $\widetilde{\Lambda}_{\widetilde{\epsilon}}$ is Markovian. The mapping $\widetilde{\Lambda}_{\epsilon_0}\Lambda_{\epsilon_0}$ is uniquely determined by the fact that it preserves $\mathbb{1}$ and that $(\widetilde{\Lambda}_{\epsilon_0}\Lambda_{\epsilon_0})\varphi = \epsilon_0^2\varphi$. This observation leads us to consider $t_0 \coloneqq t_0(L, \widetilde{L}) \coloneqq -\ln(\epsilon_0^2) \geq 0$, so that $\widetilde{\Lambda}_{\epsilon_0}\Lambda_{\epsilon_0} = \exp(t_0 L)$. We are thus in the framework considered in the introduction. Similarly, we get $\Lambda_{\epsilon_0}\widetilde{\Lambda}_{\epsilon_0} = \exp(t_0\widetilde{L})$, and this can also be deduced from Proposition 4, since Λ_{ϵ_0} is invertible. It seems that t_0 is the smallest warm-up deterministic time enabling to go from estimates of convergence for one of the semigroup to the other one. As in the introduction, let us consider more specifically the traditional case of relative entropy. Diaconis and Saloff-Coste [20] computed the logarithmic Sobolev constant $\alpha(L)$ of L:

$$\alpha(L) \quad = \quad 4\frac{1-2\mu_\wedge}{\ln(1/\mu_\wedge - 1)}\lambda$$

with $\mu_\wedge \coloneqq \mu(0) \wedge \mu(1)$, the smallest value taken by the invariant measure.

We have for any initial distribution m_0 on $\{0, 1\}$,

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 \exp(tL)|\mu) \quad \leq \quad \exp(-\alpha(L)t)\mathrm{Ent}(m_0|\mu) \tag{18}$$

Taking into account Theorem 8, this bound can be improved into

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 \exp(tL)|\mu) \quad \leq \quad \min\{\exp(-\alpha(\widetilde{L})(t - t_0(L, \widetilde{L}))_+) \,:\, \widetilde{L} \in \mathfrak{L}(L)\}\mathrm{Ent}(m_0|\mu)$$

where $\mathfrak{L}(L)$ is the set of irreducible Markov generators isospectral to L. Note that $\alpha(\widetilde{L})$ is strictly decreasing as a function of $\widetilde{\mu}_\wedge$ and thus the logarithmic Sobolev constants of L and $\widetilde{L}$ are distinct when $L \neq \widetilde{L}$ (up to the symmetry exchanging 0 and 1). Furthermore, the bound $\alpha(\widetilde{L}) \leq 2\lambda$ is only attained when $\widetilde{\mu}$ is the uniform distribution on $\{0, 1\}$ (in this case the computation of the logarithmic Sobolev inequality is due to Gross [25]). So it is appealing to try a comparison with this "fastest case" where $\widetilde{\mu} = (1/2, 1/2)$, and we get

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 \exp(tL)|\mu) \quad \leq \quad \exp(-2\lambda(t - \ln(1/\mu_\wedge - 1))_+)\mathrm{Ent}(m_0|\mu) \tag{19}$$

since

$$\epsilon_0 \;=\; \min\left(\sqrt{\frac{\mu(1)}{\mu(0)}}, \sqrt{\frac{\mu(0)}{\mu(1)}}\right) \;=\; \sqrt{\frac{\mu_\wedge}{1-\mu_\wedge}}$$

so that $t_0 = \ln(1/\mu_\wedge - 1)$.

Formula (19) becomes rapidly better than (18). It follows that, for "medium" times, to get good estimates of the relative entropy with respect to μ of the time marginal laws of the Markov evolution generated by L, it is more interesting to intertwine this evolution with the isospectral generator $\widetilde{L}$ corresponding to the uniform distribution than to compute the logarithmic Sobolev constant associated to L.

The existence of Markovian kernels Λ and $\widetilde{\Lambda}$ intertwining two irreducible isospectral (in the extended sense: equality of eigenvalues and dimensions of the Jordan blocks) and finite Markov generators was shown in [31]. We believe these kernels can furthermore be chosen so that a interweaving relation holds, as a subcase of Conjecture 2.

2.2 Classical and discrete squared Bessel processes

The examples described in this subsection and in the following one were the first instances of interweaving relations that we identified in [32]. However, this notion was not properly isolated and investigated there.
For a given $\beta > 0$, consider the classical squared Bessel diffusion generator G_β of index $\beta - 1$ (dimension 2β) on $\mathbb{R}_+$ given by

$$\forall\, x \in (0, +\infty), \qquad G_\beta \;\coloneqq\; x\partial^2 + \beta\partial$$

where ∂ is the usual differentiation operator. This diffusion generator admits μ_β as invariant (even reversible) measure, where

$$\forall\, x \in (0, +\infty), \qquad \mu_\beta(dx) \;\coloneqq\; \frac{x^{\beta-1}}{\Gamma(\beta)}\, dx$$

where Γ is the usual gamma function. For $\beta > 0$, denote $Q^{(\beta)} \coloneqq (Q_t^{(\beta)})_{t\geq 0}$ the Markov semigroup generated by G_β.

An analogue discrete squared Bessel birth-and-death generator $\mathbb{G}_\beta$ is defined by

$$\forall\, n \in \mathbb{Z}_+, \qquad \mathbb{G}_\beta \;\coloneqq\; (n+\beta)\partial_+ + n\partial_-$$

where the operators $\partial_\pm$ act on any function $\boldsymbol{f} : \mathbb{Z}_+ \to \mathbb{R}$ via

$$\forall\, n \in \mathbb{Z}_+, \qquad \partial_\pm \boldsymbol{f}(n) \coloneqq \; \boldsymbol{f}(n \pm 1) - \boldsymbol{f}(n)$$

(with the convention that $\boldsymbol{f}(-1) \coloneqq \boldsymbol{f}(0)$). The birth-and-death generator $\mathbb{G}_\beta$ admits $\boldsymbol{u}_\beta$ as invariant (even reversible) measure, where

$$\forall\, n \in \mathbb{Z}_+, \qquad \boldsymbol{u}_\beta(n) \;\coloneqq\; \frac{(n+\beta-1)(n+\beta-2)\cdots\beta}{n!}.$$

For β, $\sigma > 0$, denote $\mathbb{Q}^{(\beta,\sigma)} \coloneqq (\mathbb{Q}_t^{(\beta,\sigma)})_{t\geq 0}$ the Markov semigroup generated by $\sigma\mathbb{G}_\beta$. For $\sigma > 0$, consider Λ_σ the Markov kernel from $\mathbb{R}_+$ to $\mathbb{Z}_+$ given by the Poisson transition probability measures:

$$\forall\, x \in \mathbb{R}_+, \forall\, n \in \mathbb{Z}_+, \qquad \Lambda_\sigma(x, n) \;\coloneqq\; \frac{(\sigma x)^n}{n!} \exp(-\sigma x)$$

Conversely, for $\beta, \sigma > 0$, consider $\widetilde{\Lambda}_{\beta,\sigma}$ the Markov kernel from $\mathbb{Z}_+$ to $\mathbb{R}_+$ given by the gamma transition probability measures:

$$\forall\, n \in \mathbb{Z}_+, \forall\, x \in (0, \infty), \qquad \widetilde{\boldsymbol{\Lambda}}_{\beta,\sigma}(n, dx) \;=\; \sigma^{n+\beta} \frac{x^{n+\beta-1}}{\Gamma(n+\beta)} \exp(-\sigma x)\, dx$$

In [32], we have shown the following symmetric interweaving relation with deterministic warm-up time $\sigma > 0$:

Proposition 11 *For any $\beta, \sigma > 0$, we have*

$$Q^{(\beta)} \overset{\sigma}{\leftrightsquigarrow} \mathbb{Q}^{(\beta,\sigma)}$$

where $\Lambda = \Lambda_\sigma$ and $\widetilde{\Lambda} = \widetilde{\boldsymbol{\Lambda}}_{\beta,\sigma}$.

For $\beta > 0$, the invariant measures μ_β and $\boldsymbol{u}_\beta$ have infinite weight so the above Bessel processes do not enter in the framework of convergence to equilibrium and we cannot apply the results presented in the introduction. Nevertheless the interweaving relations of Proposition 11 are useful for simulation purposes of one process in terms of the other one, especially in the direction of using the birth-and-death process to simulate the diffusion process, as it was seen in [32].

2.2.1 Non-colliding discrete and continuous squared Bessel processes

We proceed by describing a very elegant extension of the interweaving relations between squared Bessel processes to the multidimensional setting that has been recently proposed by Assiotis [7]. More specifically, for any integer $N \geq 1$ and $\beta > 0$, let $Q^{N,\beta}$ (resp. $\mathbb{Q}^{N,\beta}$) be the semigroup of N independent copies of squared Bessel processes (resp. the discrete squared Bessel process) of index $\beta - 1$ conditioned to never intersect. These semigroups are known to be Feller semigroups acting on the space $C_0(W_+^N)$ and $\mathbb{C}_0(\mathbb{W}_+^N)$ respectively where the Weyl chambers with positive coordinates are defined by

$$W_+^N = \{\mathbf{x} = (x_1, \cdots, x_N) \in \mathbb{R}_+^N : x_1 \leq x_2 \leq \cdots \leq x_N\}$$
$$\mathbb{W}_+^N = \{\mathbf{n} = (n_1, \cdots, n_N) \in \mathbb{Z}_+^N : n_1 < n_2 < \cdots < n_N\}.$$

Then relying on the one-dimensional result that appeared in [32, Proposition 13 and 14], Assiotis obtain the following, see [7, Proposition 1, Theorem 1.4, Remark 1.6].

Proposition 12 *For any integer $N \geq 1$ and $\beta > 0$, we have*

$$Q^{N,(\beta)} \overset{1}{\leftrightsquigarrow} \mathbb{Q}^{N,(\beta)}$$

where $\Lambda = \Lambda_1^N$ and $\widetilde{\Lambda} = \widetilde{\boldsymbol{\Lambda}}_{\beta,1}^N$ are Markov kernels defined respectively, for any $\boldsymbol{n} \in \mathbb{W}_+^N$ and $\boldsymbol{x} \in W_+^N$, by

$$\Lambda_1^N(\boldsymbol{x}, \boldsymbol{n}) = \frac{\Delta_N(\boldsymbol{n})}{\Delta_N(\boldsymbol{x})} \det\left(\Lambda_1(x_i, n_j)\right)_{i,j=1}^N,$$
$$\widetilde{\boldsymbol{\Lambda}}_{\beta,1}^N(\boldsymbol{n}, d\boldsymbol{x}) = \frac{\Delta_N(\boldsymbol{x})}{\Delta_N(\boldsymbol{n})} \det\left(\widetilde{\boldsymbol{\Lambda}}_{\beta,1}(n_i, dx_j)\right)_{i,j=1}^N dx_1 \cdots dx_N,$$

and $\Delta_N(\boldsymbol{x}) = \det\left(x_i^{j-1}\right)_{i,j=1}^N = \prod_{1 \leq i < j \leq N}(x_j - x_i)$ stands for the Vandermonde determinant.

We mention that the Markov realizations of the semigroups $Q^{N,(\beta)}$ and $\mathbb{Q}^{N,(\beta)}$ appear in random matrix theory as the dynamics of the eigenvalues of the so-called continuous and discrete Laguerre ensembles and refer to [7] for further connections between these objects and other algebraic structures.

2.3 Classical and discrete Laguerre processes

A natural way to transform the transient Bessel processes into recurrent processes is recalled in [32] and it leads to the Laguerre processes. This procedure slightly modifies the interweaving relations and we ended up with the following results.

For $\beta, \sigma > 0$, consider the classical **Laguerre differential operator** $L_{\beta,\sigma}$ on $\mathbb{R}_+$ acting on $\mathbf{C}_\mathrm{b}^\infty(\mathbb{R}_+)$, the space of bounded smooth functions with bounded derivatives on $\mathbb{R}_+$, via

$$\forall\, f \in \mathbf{C}_\mathrm{b}^\infty(\mathbb{R}_+), \forall\, x \in (0, +\infty), \qquad L_{\beta,\sigma}[f](x) \;=\; \sigma x \partial^2 f(x) + (\sigma\beta - x)\partial f(x) \tag{20}$$

This operator is a one-dimensional diffusion generator and it is easy to check that its unique invariant (even reversible) probability measure $\nu_{\beta,\sigma}$ on $\mathbb{R}_+$, is the **gamma distribution** of shape parameter β and scale parameter σ, i.e.

$$\forall\, x \in (0, +\infty), \qquad \nu_{\beta,\sigma}(dx) \;=\; \frac{x^{\beta-1}\exp(-x/\sigma)}{\sigma^\beta \Gamma(\beta)}\, dx$$

It follows (via Freidrichs theory, see e.g. the book of Akhiezer and Glazman [1]) that $L_{\beta,\sigma}$ can be extended into a self-adjoint operator on $\mathbf{L}^2(\nu_{\beta,\sigma})$. The associated continuous Markov semigroup is denoted $P^{(\beta,\sigma)} \coloneqq (P_t^{(\beta,\sigma)})_{t\geq 0}$.

An analogue discrete Laguerre birth-and-death generator $\mathbb{L}_{\beta,\sigma}$ is defined by

$$\forall\, n \in \mathbb{Z}_+, \qquad \mathbb{L}_{\beta,\sigma} \;\coloneqq\; \sigma(n+\beta)\partial_+ + (\sigma+1)n\partial_- \tag{21}$$

This generator admits an invariant (even reversible) probability measure $\widetilde{\boldsymbol{v}}_{\beta,\sigma}$ on $\mathbb{Z}_+$, which is the negative binomial distribution of parameters β and $\sigma/(1+\sigma)$, i.e.

$$\forall\, n \in \mathbb{Z}_+, \qquad \widetilde{\boldsymbol{v}}_{\beta,\sigma}(n) \;\coloneqq\; (1+\sigma)^{-\beta}\left(\frac{\sigma}{\sigma+1}\right)^n \frac{(n+\beta-1)(n+\beta-1)\cdots\beta}{n!}$$

Denote $\mathbb{P}^{(\beta,\sigma)} \coloneqq (\mathbb{P}_t^{(\beta,\sigma)})_{t\geq 0}$ the Markov semigroup generated by $\mathbb{L}_{\beta,\sigma}$. In [32], we have shown the following symmetric interweaving relation with deterministic warm-up time.

Proposition 13 *For any $\beta, \sigma, \varsigma > 0$, we have*

$$P^{(\beta,\varsigma)} \overset{\ln(1+\frac{1}{\varsigma\sigma})}{\leftrightsquigarrow} \mathbb{P}^{(\beta,\varsigma\sigma)}$$

where $\Lambda = \Lambda_\sigma$ and $\widetilde{\Lambda} = \widetilde{\Lambda}_{\beta,\sigma+\frac{1}{\varsigma}}$.

The last relation can be seen as a consequence of the last-but-one identity, via Proposition 4, since Λ_σ, from [32, Lemma 2.2], is one-to-one. The relations of Proposition 13 can be summarized by the following diagram:

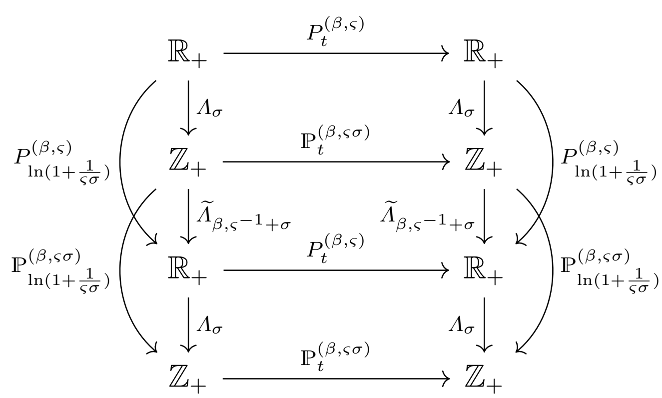

Figure 3: Laguerre intertwining relations

The interweaving relations between the continuous and discrete Laguerre processes enable to deduce links between their speed of convergence to equilibrium. As in the introduction, let us present them in the usual entropy sense (see Section 4 for generalisations). First we recall the logarithmic Sobolev inequalities satisfied by the Laguerre semigroups.

We start with the classical situation. For any $\beta, \varsigma > 0$, the **logarithmic Sobolev constant** $\alpha(\beta,\varsigma)$ associated to the generator $L_{\beta,\varsigma}$ defined in (20) is

$$\alpha(\beta,\varsigma) \;\coloneqq\; \inf_{f\in\mathbf{C}^1_{\mathrm{b}}(\mathbb{R}_+)\,:\,\nu_{\beta,\varsigma}[f^2]=1} \frac{4\varsigma\int_{\mathbb{R}_+} xf'^2(x)\,\nu_{\beta,\varsigma}(dx)}{\int_{\mathbb{R}_+} f^2(x)\ln(f^2(x))\,\nu_{\beta,\varsigma}(dx)} \tag{22}$$

(for any $k \in \mathbb{N}$, $\mathbf{C}_{\mathrm{b}}^k(\mathbb{R}_+)$ is the space of bounded continuously k times differentiable functions on $\mathbb{R}_+$, with bounded derivatives). The numerator in (22) is four times the **Dirichlet form (energy)** $\mathcal{E}_{\beta,\varsigma}(f,f)$ associated to $L^{(\beta,\varsigma)}$ and defined, at least for $f \in \mathbf{C}_{\mathrm{b}}^2(\mathbb{R}_+)$, by

$$\begin{aligned}\mathcal{E}_{\beta,\varsigma}(f,f) &\coloneqq -\nu_{\beta,\varsigma}[fL_{\beta,\varsigma}[f]]\\ &= \varsigma\int_{\mathbb{R}_+} xf'^2(x)\,\nu_{\beta,\varsigma}(dx)\end{aligned}$$

where the last equality is obtained by integration by parts and the last expression enables to extend the domain of definition of $\mathcal{E}_{\beta,\varsigma}$.

It is well-known (see for instance the book of Ané et al. [4]) that the logarithmic Sobolev constant is bounded above by twice the spectral gap of the associated generator. In the present setting, it implies that $\alpha(\beta,\varsigma) \le 2$ for any $\beta,\varsigma > 0$, since the spectrum of $L_{\beta,\varsigma}$ is $-\mathbb{Z}_+$ with eigenvalues of multiplicity 1, and so its spectral gap is 1. In fact the constant $\alpha(\beta,\varsigma)$ does not depend on ς:

Lemma 14 *For any $\beta,\varsigma > 0$, we have $\alpha(\beta,\varsigma) = \alpha(\beta)$, where $\alpha(\beta) \coloneqq \alpha(\beta,1)$.*

Remark 15 The constant $\alpha(\beta)$ has been well-studied. Via the famous Γ_2-criterion, Bakry [8] has shown that $\alpha(\beta) = 1$ for all $\beta \ge 1/2$. Otherwise, the behavior of $\alpha(\beta)$ changes when $\beta > 0$ is going to 0_+, since it converges to zero as $\alpha(\beta) \sim -4/\ln\beta$, see [30]. We also refer to Corollary for an alternative analysis based on the concept of interweaving relation of the convergence to equilibrium in entropy for $0 \le \beta < \frac{1}{2}$.

Proof: For any $\varsigma > 0$, let M_ς be the dilation operator acting on any function f defined on $\mathbb{R}_+$ via

$$M_\varsigma f(x) = f(\varsigma x)$$

An immediate linear change of variable shows that for any $\beta,\varsigma > 0$, we have $\nu_{\beta,\varsigma} = \nu_\beta M_\varsigma$ (where ν_β stands for $\nu_{\beta,1}$). For $f \in \mathbf{C}_{\mathrm{b}}^1(\mathbb{R}_+)$ with $\nu_{\beta,\varsigma}[f^2] = 1$, consider the function $\widetilde{f} \coloneqq M_\varsigma f$. We have on the one hand,

$$\begin{aligned}\nu_{\beta,\varsigma}[f^2] &= \nu_\beta[\widetilde{f}^2]\\ \int_{\mathbb{R}_+} f^2(x)\ln f^2(x)\,\nu_{\beta,\varsigma}(dx) &= \int_{\mathbb{R}_+} \widetilde{f}^2(x)\ln\widetilde{f}^2(x)\,\nu_\beta(dx)\end{aligned}$$

and on the other hand,

$$\int_{\mathbb{R}_+} xf'^2(x)\,\nu_{\beta,\varsigma}(dx) = \frac{1}{\varsigma}\int_{\mathbb{R}_+} x\widetilde{f}'^2(x)\,\nu_\beta(dx)$$

The announced result now follows from the bijectivity of the mapping $f \mapsto \widetilde{f}$ between $\{f \in \mathbf{C}_{\mathrm{b}}^1(\mathbb{R}_+) : \nu_{\beta,\varsigma}[f^2] = 1\}$ and $\{\widetilde{f} \in \mathbf{C}_{\mathrm{b}}^1(\mathbb{R}_+) : \nu_\beta[\widetilde{f}^2] = 1\}$.

■

Here we are interested in $\alpha(\beta)$ since for any initial distribution m_0 on $\mathbb{R}_+$, we have

$$\forall\, t \ge 0, \qquad \mathrm{Ent}(m_0P_t^{(\beta,\varsigma)}|\nu_{\beta,\varsigma}) \le \exp(-\alpha(\beta)t)\mathrm{Ent}(m_0|\nu_{\beta,\varsigma}) \tag{23}$$

(of course, such a bound is only relevant when the initial relative entropy $\mathrm{Ent}(m_0|\nu_{\beta,\varsigma})$ is finite) and $\alpha(\beta)$ is optimal for these equalities to hold for any initial distribution $m_0 \in \mathcal{P}((0,+\infty))$ and for any time $t \ge 0$.

The quantitative convergence to equilibrium in the entropy sense has not been investigated for the discrete Laguerre generators. A priori, we have the following information. For $\beta, \sigma > 0$, the modified logarithmic Sobolev constant $\boldsymbol{\alpha}_m(\beta,\sigma)$ associated to the generator $\mathbb{L}_{\beta,\sigma}$ defined in (21) is

$$\boldsymbol{\alpha}_m(\beta,\sigma) \quad \coloneqq \quad \inf_{f\in \mathbf{F}_{\mathrm{f}}(\mathbb{Z}_+)\,:\,\boldsymbol{v}_{\beta,\sigma}[\boldsymbol{f}^2]=1} \frac{\mathbb{E}_{\beta,\sigma}(\boldsymbol{f}^2, \ln(\boldsymbol{f}^2))}{\boldsymbol{v}_{\beta,\sigma}[\boldsymbol{f}^2\ln(\boldsymbol{f}^2)]} \tag{24}$$

where $\mathbf{F}_{\mathrm{f}}(\mathbb{Z}_+)$ is the space of functions defined on $\mathbb{Z}_+$ which vanish except on a finite subset of points and where the Dirichlet form $\mathbb{E}_{\beta,\sigma}(\boldsymbol{f},\boldsymbol{g})$ of two functions $\boldsymbol{f},\boldsymbol{g}\in \mathbf{F}_{\mathrm{f}}(\mathbb{Z}_+)$ is given by

$$\begin{aligned}\mathbb{E}_{\beta,\sigma}(\boldsymbol{f},\boldsymbol{g}) &\coloneqq -\boldsymbol{v}_{\beta,\sigma}[\boldsymbol{f}\mathbb{L}_{\beta,\sigma}[\boldsymbol{g}]]\\ &= \sum_{n\in\mathbb{Z}_+} (\boldsymbol{f}(n+1)-\boldsymbol{f}(n))(\boldsymbol{g}(n+1)-\boldsymbol{g}(n))\boldsymbol{v}_{\beta,\sigma}(n)\mathbb{L}_{\beta,\sigma}(n,n+1)\end{aligned}$$

Again, the interest of $\boldsymbol{\alpha}_m(\beta,\sigma)$ is the discrete analogue of (23): for any initial distribution $\boldsymbol{m}_0$ on $\mathbb{Z}_+$, we have

$$\forall\, t\geq 0, \qquad \mathrm{Ent}(\boldsymbol{m}_0\mathbb{P}_t^{(\beta,\sigma)}|\boldsymbol{v}_{\beta,\sigma}) \quad\leq\quad \exp(-\boldsymbol{\alpha}_m(\beta,\sigma)t)\mathrm{Ent}(\boldsymbol{m}_0|\boldsymbol{v}_{\beta,\sigma}) \tag{25}$$

(for the deduction of this bound and (23) by differentiating their respective left-hand-side. with respect to the time $t\geq 0$, see again the book of Ané et al. [4]) and $\boldsymbol{\alpha}_m(\beta,\sigma)$ is optimal for these inequalities to hold for any initial distribution $\boldsymbol{m}_0\in\mathcal{P}(\mathbb{Z}_+)$ and for any time $t\geq 0$. We also have that $\boldsymbol{\alpha}_m(\beta,\sigma)$ is bounded above by twice the spectral gap of $\mathbb{L}_{\beta,\sigma}$. Namely $\boldsymbol{\alpha}_m(\beta,\sigma)\leq 2$ for any $\beta,\sigma>0$, since the spectrum of $\mathbb{L}_{\beta,\sigma}$ is $-\mathbb{Z}_+$. Unfortunately, there is no proper way to estimate from below $\boldsymbol{\alpha}_m(\beta,\sigma)$, which is only known in very few situations, especially related to the Poisson distribution, see Wu [49]. That is why $\boldsymbol{\alpha}_m(\beta,\sigma)$ is often replaced by the classical logarithmic Sobolev constant $\boldsymbol{\alpha}(\beta,\sigma)$, given by

$$\boldsymbol{\alpha}(\beta,\sigma) \quad\coloneqq\quad \inf_{\boldsymbol{f}\in \mathbf{F}_{\mathrm{f}}(\mathbb{Z}_+)\,:\,\boldsymbol{v}_{\beta,\sigma}[\boldsymbol{f}^2]=1} \frac{4\mathbb{E}_{\beta,\sigma}(\boldsymbol{f},\boldsymbol{f}))}{\boldsymbol{v}_{\beta,\sigma}[\boldsymbol{f}^2\ln(\boldsymbol{f}^2)]} \tag{26}$$

It can be checked that $\boldsymbol{\alpha}(\beta,\sigma)\leq\boldsymbol{\alpha}_m(\beta,\sigma)$, so that (25) still holds with $\boldsymbol{\alpha}_m(\beta,\sigma)$ replaced by $\boldsymbol{\alpha}(\beta,\sigma)$, with the advantage that the latter ergodic constant can be estimated via discrete Hardy's inequalities (cf. [29]):

Consider the quantity

$$C_{\beta,\sigma} \quad\coloneqq\quad \min_{n\in\mathbb{Z}_+}\max(C^-_{\beta,\sigma}(n), C^+_{\beta,\sigma}(n))$$

where for any $n\in\mathbb{Z}_+$, we take

$$\begin{aligned}C^-_{\beta,\sigma}(n) &\coloneqq \sup_{m<n}\left(\sum_{l=m}^{n-1}\frac{1}{\boldsymbol{v}_{\beta,\sigma}(l)\mathbb{L}_{\beta,\sigma}(l,l+1)}\right)\boldsymbol{v}_{\beta,\sigma}([\![0,m]\!])\ln(1/\boldsymbol{v}_{\beta,\sigma}([\![0,m]\!]))\\ C^+_{\beta,\sigma}(n) &\coloneqq \sup_{m>n}\left(\sum_{l=n}^{m-1}\frac{1}{\boldsymbol{v}_{\beta,\sigma}(l)\mathbb{L}_{\beta,\sigma}(l,l+1)}\right)\boldsymbol{v}_{\beta,\sigma}([\![m,\infty[\![)\ln(1/\boldsymbol{v}_{\beta,\sigma}([\![m,\infty[\![))\end{aligned}$$

We have the general bounds

$$\frac{1}{10}\frac{1}{C_{\beta,\sigma}} \;\leq\; \boldsymbol{\alpha}(\beta,\sigma) \;\leq\; \frac{8}{3}\left(1-\frac{\sqrt{5}}{2\sqrt{2}}\right)^{-1}\frac{1}{C_{\beta,\sigma}} \tag{27}$$

These expressions can be exploited to get reasonably accurate estimates on $\boldsymbol{\alpha}(\beta,\sigma)$ in terms of β and σ, in particular $\boldsymbol{\alpha}_m(\beta,\sigma) \geq \boldsymbol{\alpha}(\beta,\sigma) > 0$ for all $\beta,\sigma > 0$ (insuring that the bound (25) is not trivial).

Nevertheless, the underlying computations are not so nice, while resorting to interweaving relations eventually leads to better bounds on the convergence to equilibrium in the entropy sense. More precisely, as a particular consequence of Theorem 8 applied to the three last lines of Figure 3, with $\varsigma = 1$, we get

Corollary 16 *For any initial probability $\boldsymbol{m}_0$ on $\mathbb{Z}_+$ and for any $\beta,\sigma > 0$ and $t \geq 0$, we have*

$$\mathrm{Ent}(\boldsymbol{m}_0\mathbb{P}_t^{(\beta,\sigma)}|\boldsymbol{v}_{\beta,\sigma}) \quad \leq \quad \left(\frac{\sigma+1}{\sigma}\right)^{\alpha(\beta)} e^{-\alpha(\beta)t}\,\mathrm{Ent}(\boldsymbol{m}_0|\boldsymbol{v}_{\beta,\sigma})$$

where we recall that $\alpha(\beta) = 1$ for any $\beta \geq 1$, see Remark 15.

In particular, for $\beta \geq 1/2$ and up to waiting a warming-up time $\ln(1+\frac{1}{\sigma})$, before which Corollary (16) provides no information and is less good than (25), we get after this period an exponential rate of convergence equal to 1 (the best possible asymptotical one would be 2, i.e. twice the spectral gap of $\mathbb{L}_{\beta,\sigma}$). Corollary 16 is also relevant for small $\beta > 0$, since one cannot hope for an estimate so simple via (27).

Applying the bounds from Theorem 8 to the three first lines of Figure 3, we get for any initial probability $\boldsymbol{m}_0$ on $\mathbb{Z}_+$, any $\beta,\varsigma,\sigma > 0$ and any $t \geq 0$,

$$\mathrm{Ent}(\boldsymbol{m}_0\mathbb{P}_t^{(\beta,\varsigma)}|\boldsymbol{v}_{\beta,\varsigma}) \quad \leq \quad \exp(-\boldsymbol{\alpha}_m(\beta,\varsigma\sigma)[t-\ln(1+\frac{1}{\varsigma\sigma})]_+)\mathrm{Ent}(\boldsymbol{m}_0|\boldsymbol{v}_{\beta,\varsigma})$$

Letting $\sigma > 0$ go to infinity and recalling that $\alpha(\beta)$ is optimal in (23), we deduce that

$$\forall\ \beta > 0, \qquad \overline{\alpha}(\beta) \ \coloneqq \ \overline{\lim_{\sigma\to+\infty}}\,\boldsymbol{\alpha}_m(\beta,\sigma) \quad \leq \quad \alpha(\beta) \tag{28}$$

In particular $\overline{\alpha}(\beta)$ is going to zero as β goes to 0_+ (in fact we believe that $\overline{\alpha}(\beta) = \alpha(\beta)$, as suggested by the remark about approximations at the end of this subsection).

Similar relations between the classical and discrete Laguerre semigroups are equally valid concerning hyperboundedness via Theorem 9. Indeed, the logarithmic Sobolev inequalities imply that for any $\beta,\varsigma > 0$, we have

$$\forall\ t \geq 0, \qquad \|P_t^{(\beta,\varsigma)}\|_{\mathbf{L}^2(\nu_{\beta,\varsigma})\to\mathbf{L}^{p(\alpha(\beta)t)}(\nu_{\beta,\varsigma})} \quad \leq \quad 1$$

where $\alpha(\beta)$ is defined in Lemma 14 and

$$\forall\ t \geq 0, \qquad p(\alpha(\beta)t) \quad \coloneqq \quad 1+\exp(\alpha(\beta)t)$$

and for any $\beta,\sigma > 0$

$$\forall\ t \geq 0, \qquad \|\mathbb{P}_t^{(\beta,\sigma)}\|_{\mathbf{L}^2(\boldsymbol{v}_{\beta,\sigma})\to\mathbf{L}^{p(\boldsymbol{\alpha}(\beta,\sigma)t)}(\boldsymbol{v}_{\beta,\sigma})} \quad \leq \quad 1$$

where $\boldsymbol{\alpha}(\beta,\sigma)$ is defined in (26) and

$$\forall\ t \geq 0, \qquad p(\boldsymbol{\alpha}(\beta,\sigma)t) \quad \coloneqq \quad 1+\exp(\boldsymbol{\alpha}(\beta,\sigma)t)$$

But due to the difficulty in estimating $\boldsymbol{\alpha}(\beta,\sigma)$, it is preferable to use Theorem 9 to deduce that

$$\forall\ t \geq 0, \qquad \|\mathbb{P}_{t+\ln(1+\frac{1}{\sigma})}^{(\beta,\sigma)}\|_{\mathbf{L}^2(\boldsymbol{v}_{\beta,\sigma})\to\mathbf{L}^{p(\alpha(\beta)t)}(\boldsymbol{v}_{\beta,\sigma})} \quad \leq \quad 1$$

To end this subsection, let us mention two other applications of the interweaving relations of Proposition 13.

• **Approximations**: For any $\beta, \varsigma > 0$, let $X^{(\beta,\varsigma)} := (X_t^{(\beta,\varsigma)})_{t\geq 0}$ (respectively $\widetilde{X}^{(\beta,\varsigma)} := (\widetilde{X}_t^{(\beta,\varsigma)})_{t\geq 0}$) be a Markov process associated to $P^{(\beta,\varsigma)}$ (resp. $\mathbb{P}^{(\beta,\varsigma)}$). As seen in [32], for large $\sigma > 0$ the birth and death process $(\mathbb{X}_t^{(\beta,\sigma\varsigma)})_{t\geq 0}$ provides an isospectral approximation of $(X_t^{(\beta,\varsigma)})_{t\geq 0}$. This is related to the fact that $\bar{\alpha}(\beta)$ should be close to $\alpha(\beta)$, as suggested by (28).

• **Simulations**: For $\sigma > 0$, $x \in \mathbb{R}_+$ and $t \geq 0$, the random variable $Y := X^{(\beta,\varsigma)}_{\ln(1+\frac{1}{\varsigma\sigma})+t}$ can be simulated by first sampling $\widetilde{x}$ under the probability $\Lambda_\sigma(x, d\widetilde{x})$, next by simulating $\widetilde{X}_t^{(\beta,\varsigma\sigma)}$ starting with $\widetilde{X}_0^{(\beta,\varsigma\sigma)} = \widetilde{x}$ (comprehensively, this amounts to simulate $\widetilde{X}_t^{(\beta,\varsigma\sigma)}$ with the initial distribution $\Lambda_\sigma(x,\cdot)$) and finally by sampling Y under the probability $\widetilde{\Lambda}_{\varsigma^{-1}+\sigma}(\widetilde{X}_t^{(\beta,\varsigma\sigma)},\cdot)$.

2.4 Degenerate hypoelliptic Ornstein-Uhlenbeck processes

We now describe a refined version of a interweaving relation between degenerate and non-degenerate hypoelliptic Ornstein-Uhlenbeck semigroups on $\mathbb{R}^d, d \geq 1$, that was identified in [40]. In that paper, the authors exploit the interweaving relations to obtain the hypocoercive estimate with explicit constants for the convergence to equilibrium in the weighted Hilbert space of the degenerate hypoelliptic Ornstein-Uhlenbeck semigroups which are non-normal. Therein, we provide further applications of these interweaving relations to these degenerate semigroups including entropy and hypercontractivity estimates and the cut-off phenomena. To define these semigroups, we let B and Γ be $d \times d$-matrices with $\sigma(B) \subseteq \{z \in \mathbb{C};\ \Re(z) > 0\}$ and Γ being positive semi-definite such that $\det \Gamma_t > 0$ for all $t > 0$ where

$$\Gamma_t = \int_0^t e^{-sB}\Gamma e^{-sB^*} ds,$$

and the matrix B^* stands for the adjoint of B. In particular, this holds when Γ is invertible, which we call the non-degenerate case, although it can happen that $\det \Gamma_t > 0$, for all $t > 0$, with $\det \Gamma = 0$, which we call the degenerate case. An equivalent condition to $\det \Gamma_t > 0$ for all $t > 0$ is that $\ker \Gamma$, the kernel of Γ, does not contain any invariant subspace of B^*. Under these assumptions on (Γ, B), the hypoelliptic Ornstein-Uhlenbeck semigroup P admits an unique invariant measure which is the following gaussian distribution

$$\rho_{\Gamma_\infty}(d\mathbf{x}) = \frac{e^{-\langle \Gamma_\infty^{-1}\mathbf{x},\mathbf{x}\rangle/2}}{\sqrt{(2\pi)^d \det \Gamma_\infty}} d\mathbf{x},\ \mathbf{x} \in \mathbb{R}^d,$$

with $\Gamma_\infty = \int_0^\infty e^{-tB}\Gamma e^{-tB^*} ds$ and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^d$. P extends to a contraction semigroup on the weighted Hilbert space $\mathbf{L}^2(\rho_\infty)$. We also recall that the generator of the Ornstein-Uhlenbeck semigroup $P = (e^{-t\mathbf{A}})_{t\geq 0}$ acts on suitable functions f via

$$\mathbf{A}[f](\mathbf{x}) = \frac{1}{2}\sum_{i,j=1}^d \gamma_{ij}\partial_i\partial_j f(\mathbf{x}) - \sum_{i,j=1}^d b_{ij}x_j\partial_i f(\mathbf{x}) = \frac{1}{2}\mathrm{tr}(\Gamma\nabla^2)f(\mathbf{x}) - \langle B\mathbf{x}, \nabla\rangle f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

and the condition $\det \Gamma_t > 0$, for all $t > 0$, is equivalent to the hypoellipticity of $\frac{\partial}{\partial t} + \mathbf{A}$ in the $d+1$ variables $(t, x_1, \ldots, x_d)$, hence the terminology. In Metafune, Pallara and Priola [28, Theorem 3.1] (see also Bogatchev [12] and Aleman and Viola [3]) it was shown that the spectrum of $\mathbf{A}$ in $\mathbf{L}^2(\rho_{\Gamma_\infty})$ is entirely determined by the one of the matrix B, specifically that, writing $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\sigma(\mathbf{A}) = \{\sum_{i=1}^r k_i b_i;\ k_i \in \mathbb{N}\}$, where $b_1, \ldots, b_r$ are the distinct eigenvalues of B. Hence, in particular, the spectral gap of $\mathbf{A}$ is $\lambda_1 = b_\wedge$ as the smallest eigenvalue of $\frac{1}{2}(B + B^*)$.

Next, we denote by $\kappa(V)$ the condition number of any invertible matrix V, and note that if V is positive-definite then $\kappa(V) = v_\vee / v_\wedge$, where $v_\vee, v_\wedge > 0$ are the largest and smallest eigenvalues of V, respectively. In the following we write, for a vector $\boldsymbol{\alpha} \in \mathbb{R}^d$, $D_{\boldsymbol{\alpha}}$ for the diagonal matrix with diagonal entries given by $\boldsymbol{\alpha}$.

Proposition 17 *Let P be a (possibly) degenerate hypoelliptic Ornstein-Uhlenbeck semigroup associated to (Γ, B), that is $\ker \Gamma$ does not contain any invariant subspace of B^*. Suppose that B is diagonalizable with similarity matrix V, and that $\sigma(B) \subseteq (0, \infty)$, that is $VBV^{-1} = D_{\boldsymbol{b}}$, where $\boldsymbol{b} \in \mathbb{R}^d$ is the vector of eigenvalues of B with $b_i > 0$ for all $i \in \{1, \ldots, d\}$ and we set*

$$\alpha_i = \gamma_{\wedge,\infty} e^{\frac{b_i}{b_\wedge} \log \kappa(V\Gamma_\infty V^*)} \quad \text{and} \quad \delta_i = \gamma_\infty$$

where $\gamma_{\wedge,\infty}$ (resp. $b_\wedge$) is the smallest eigenvalues of $V\Gamma_\infty V^*$ (resp. B). Then, there exists a non-degenerate hypoelliptic Ornstein-Uhlenbeck semigroup $\tilde{P}$ associated to $(D_{\boldsymbol{\alpha}+2\boldsymbol{b}}, D_{\boldsymbol{b}})$, self-adjoint on $\mathbf{L}^2(\tilde{\rho}_{D_{\boldsymbol{\alpha}}})$, such that

$$P \overset{\mathbf{t}}{\leftrightsquigarrow} \tilde{P}$$

where $\mathbf{t} = \frac{1}{b_\wedge} \log \kappa(V\Gamma_\infty V^*)$, $\Lambda : \mathbf{L}^2(\rho_{D_{\boldsymbol{\alpha}}}) \to \mathbf{L}^2(\rho_{\Gamma_\infty})$ and $\tilde{\Lambda} : \mathbf{L}^2(\rho_{\Gamma_\infty}) \to \mathbf{L}^2(\rho_{D_{\boldsymbol{\alpha}}})$ are bounded and one-to-one Markov operators defined respectively by

$$\Lambda f(\mathbf{x}) = f * \rho_{D^{(\boldsymbol{\alpha})}}(V\mathbf{x}) \text{ and } \tilde{\Lambda} f(\mathbf{x}) = \frac{1}{\rho_{D^{(\boldsymbol{\alpha})}}(\mathbf{x})}((f_V * \rho_{D^{(\boldsymbol{\delta})}})\rho_{D^{(\boldsymbol{\delta})}}) * \rho_{D_{\boldsymbol{\alpha}-\boldsymbol{\delta}}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{29}$$

where $*$ denotes the additive convolution operator, for $\boldsymbol{a} \in \mathbb{R}^d$, $D^{(\boldsymbol{a})} = D_{\boldsymbol{a}} - V\Gamma V^*$ and $f_V(\mathbf{x}) = f(V^{-1}\mathbf{x})$.

Proof: First note that the change of coordinates map $\Phi_V f(\mathbf{x}) = f(V^{-1}\mathbf{x})$ is a unitary operator from $\mathbf{L}^2(\rho_{\Gamma_\infty})$ to $\mathbf{L}^2(\rho_{\Gamma_\infty}^{\Phi_V})$, where $\rho_\infty^{\Phi_V}$ denotes the image density of ρ_∞ under Φ_V, i.e. for $\mathbf{x} \in \mathbb{R}^d$, $\rho_\infty^{\Phi_V}(\mathbf{x}) = \frac{1}{|\det V|}\rho_\infty(V^{-1}\mathbf{x})$. Next, since B is diagonalizable with similarity matrix V we have that $VBV^{-1} = D_{\mathbf{b}}$, where $\mathbf{b} \in \mathbb{R}^d$ is the vector of eigenvalues of B with $b_i > 0$ for all $i = 1, \ldots, d$. Under this change of coordinates, (Γ, B) gets mapped to $(V\Gamma V^*, D_{\mathbf{b}})$ and a simple calculation shows that Γ_∞ then gets mapped to $V\Gamma_\infty V^*$. Hence if we prove the desired result for the Ornstein-Uhlenbeck semigroup $\overline{P}$ associated to $(V\Gamma V^*, D_{\mathbf{b}})$ then, since $P_t = {\Phi_V}^{-1}\overline{P}_t\Phi_V$ we get, by Theorem 7 and the unitary property of Φ_V, that the claims hold for the Ornstein-Uhlenbeck semigroup P associated to (Γ, B). From [40, Proposition 4.2], we know that $\widetilde{P} \overset{\mathbf{t}}{\hookleftarrow} \overline{P}$ where $\widetilde{P}$ is the Ornstein-Uhlenbeck semigroup associated to $(D_{\boldsymbol{\alpha}+2\mathbf{b}}, D_{\mathbf{b}})$ which is self-adjoint on $\mathbf{L}^2(\rho_{D_{\boldsymbol{\alpha}}})$, hence non-degenerate and the operators Λ and $\tilde{\Lambda}$ are quasi-affinities on the appropriate weighted $\mathbf{L}^2$ spaces. In particular, they are both one-to-one and hence the interweaving relation is symmetric by Theorem 5 which completes the proof with another application of Theorem 7.

■

We proceed by providing some by-products of this interweaving relation. First, we recall that in [40, Theorem 3.1], the following hypocoercive estimate was given, for any $f \in \mathbf{L}^2(\rho_{\Gamma_\infty})$,

$$\forall\, t \geq 0, \qquad \operatorname{Var}_{\rho_{\Gamma_\infty}}(P_t f) \leq \kappa(V\Gamma_\infty V^*) \exp(-2b_\wedge t) \operatorname{Var}_{\rho_{\Gamma_\infty}}(f)$$

where $\operatorname{Var}_{\rho_{\Gamma_\infty}}(f) = \int_{\mathbb{R}^d} (f(\mathbf{x}) - \rho_{\Gamma_\infty} f)^2 \rho_{\Gamma_\infty}(\mathbf{x}) d\mathbf{x}$. We carry on by recalling that, in the one-dimensional case $d = 1$, it is well known that the self-adjoint Ornstein-Uhlenbeck semigroup $\tilde{P}^{(i)}, i = 1, \ldots, d$, associated to (α_i, b_i) and whose generator is given by

$$\tilde{\mathbf{A}}_{(i)}[f](x) = -\frac{(\alpha_i + 2b_i)^2}{2} f''(x) - b_i x f'(x), \quad x \in \mathbb{R},$$

satisfies the so-called curvature dimension $CD(b_i, \infty)$ which is equivalent to the strict log-Sobolev inequality with constant b_i, see [9, Section 2.7.1]. Then observing that $\tilde{P}$, defined in Proposition 17, is the product of the $\tilde{P}^{(i)}$'s, that is $\tilde{P} = \bigotimes_{i=1}^d \tilde{P}^{(i)}$, we get from the stability of the log-Sobolev inequality under products, see [9, Proposition 5.2.7], that $\tilde{P}$ satisfies the strict log-Sobolev inequality with constant $b_\wedge$ the minimum of the log-Sobolev constants. This yields the following estimate for the convergence in entropy

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0\tilde{P}_t|\tilde{\rho}_\infty) \;\; \leq \;\; \exp(-b_\wedge t)\mathrm{Ent}(m_0|\tilde{\rho}_\infty)$$

valid for any initial distribution m_0 on $\mathbb{R}^d$. Moreover, resorting again to the famous equivalence between the log-Sobolev inequality and the hypercontractivity property due to Gross [25], we get, writing

$$\forall\, t \geq 0, \qquad \tilde{p}(t) \;\; \coloneqq \;\; 1 + \exp(b_\wedge t)$$

that

$$\forall\, t \geq 0, \qquad \|\tilde{P}_t\|_{\mathbf{L}^2(\rho_\infty)\to\mathbf{L}^{\tilde{p}(t)}(\rho_\infty)} \;\; \leq \;\; 1$$

We emphasize that the extension of such estimates to degenerate hypoelliptic Ornstein-Uhlenbeck semigroup P have met with resistance so far due to the fact that P is non-self-adjoint (even non-normal) on $\mathbf{L}^2(\rho_\infty)$, see [34, Lemma 3.3]. However, the interweaving relation described in Proposition 17 combined with the theorems 8 and 9 enable us to obtain the following.

Corollary 18 *Let P be the degenerate hypoelliptic Ornstein-Uhlenbeck semigroup as defined in Proposition 17. Then, for any initial distribution m_0 on $\mathbb{R}^d$, we have*

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 P_t|\rho_{\Gamma_\infty}) \;\; \leq \;\; \kappa(V\Gamma_\infty V^*)\exp(-b_\wedge t)\mathrm{Ent}(m_0|\rho_{\Gamma_\infty}) \tag{30}$$

and

$$\forall\, t \geq 0, \qquad \|P_{t+\mathbf{t}}\|_{\mathbf{L}^2(\rho_{\Gamma_\infty})\to\mathbf{L}^{\tilde{p}(t)}(\rho_{\Gamma_\infty})} \;\; \leq \;\; 1$$

We mention that Arnold and Erb [6] have obtained hypocoercivity estimate of the form (30), under our assumptions, with exponential rate given by the spectral gap $b_\wedge$ and that Arnold et al. [5] and Monmarché [33] have proved hypocoercivity with exponential rate $b_\wedge$ without assuming that B is diagonalizable. However, in contrast to these existing results, we are able to explicitly identify the constant in front of the exponential, i.e. $\kappa(V\Gamma_\infty V^*)$, in terms of the initial data Γ and B. Note that, in particular, if B is symmetric then V is unitary and $\kappa(V\Gamma_\infty V^*) = \kappa(\Gamma_\infty)$. However we are not aware of results regarding the hypercontractivity estimates.

We now turn to another application of interweaving which allows to identify the cut-off phenomena for degenerate hypoelliptic Ornstein-Uhlenbeck semigroups. To this end, let $\alpha \coloneqq (\alpha_1, \alpha_2, ..., \alpha_d)$ and $b \coloneqq (b_1, b_2, ..., b_d)$ be vectors from $(0, +\infty)^d$. Denote $b_\wedge \coloneqq \min(b_l, l \in [\![d]\!])$ and for any $n \in \mathbb{N}$, $\boldsymbol{\alpha}^{(n)} \coloneqq (\alpha, \dots, \alpha) \in \mathbb{R}^{dn}$ and $\boldsymbol{b}^{(n)} \coloneqq (b, \dots, b) \in \mathbb{R}^{dn}$. Consider the family of semigroups $(\tilde{P}^{(n)})_{n\in\mathbb{Z}_+}$ associated for each $n \in \mathbb{Z}_+$ to $(D_{\boldsymbol{\alpha}^{(n)}+2\boldsymbol{b}^{(n)}}, D_{\boldsymbol{b}^{(n)}})$. Lachaud [27] has shown that this family has a cut-off at the time

$$t^{(n)} \;\; \coloneqq \;\; \frac{\log n}{2b_\wedge} \tag{31}$$

(more precisely, Lachaud [27] has only considered the case $d = 1$, but her arguments extend to any $d \in \mathbb{N}$, see also Barrera, Lachaud and Ycart [11]).

We have the following generalization.

Corollary 19 *For any $n \in \mathbb{N}$, let $P^{(n)}$ be the degenerate hypoelliptic Ornstein-Uhlenbeck semigroup in $\mathbb{R}^{dn}$ as defined in Proposition 17 and associated to some $(\Gamma^{(n)}, B^{(n)})$, with $\kappa(\Gamma_\infty^{(n)})$ satisfying*

$$\lim_{n\to\infty} \frac{\log(n)}{\kappa(\Gamma_\infty^{(n)})} = +\infty$$

and $\boldsymbol{\alpha}^{(n)}$ as above $\boldsymbol{b}^{(n)}$ as above. Then, the family $(P^{(n)})_{n\in\mathbb{N}}$ has a cut-off at the times $(t^{(n)})_{n\in\mathbb{N}}$ defined in (31).

Proof: Under the conditions of the claim, we easily check that Proposition 17 entails that for each $n \in \mathbb{N}$, $P^{(n)} \overset{\mathbf{t}^{(n)}}{\hookleftarrow} \widetilde{P}^{(n)}$ where $\mathbf{t}^{(n)} = \frac{1}{b_\wedge} \log \kappa(\Gamma_\infty^{(n)})$ and $\tilde{P}^{(n)}$ is the semigroup of the self-adjoint Ornstein-Uhlenbeck process defined before the corollary. We conclude the proof by invoking the result of Lachaud [27] recalled before the corollary and Theorem 10.

■

To finish this section, let us give a concrete example.

Consider the matrices

$$\Gamma := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad B := \begin{pmatrix} 0 & -1 \\ 1/2 & 2 \end{pmatrix}$$

The corresponding Ornstein-Uhlenbeck is a simple example of a kinetic model: the first coordinate corresponds to the position in $\mathbb{R}$ of a particle in the quadratic potential $\mathbb{R} \ni x \mapsto x^2/4$, and the second coordinate is the speed, on which is acting a Brownian motion. It is a typical instance of a hypoelliptic system. To see it admits an invariant probability and the existence of Γ_∞, it is sufficient to check that the eigenvalues b_1, b_2 of B are positive. They are indeed the solutions of the second order equation $X^2 - 2X + 1/2 = 0$ and we get

$$\begin{aligned} b_1 &= 1 - 1/\sqrt{2} \\ b_2 &= 1 + 1/\sqrt{2} \end{aligned}$$

Let Γ_∞ and α_1, $\alpha_2 > 0$ be as in Proposition 17. Denote $b := (b_1, b_2)$ and $\alpha := (\alpha_1, \alpha_2)$.

For any $n \in \mathbb{N}$, introduce the tensorizations

$$\Gamma^{(n)} := \begin{pmatrix} \Gamma & 0 & 0 & \cdots & 0 \\ 0 & \Gamma & 0 & \cdots & 0 \\ 0 & 0 & \Gamma & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \Gamma \end{pmatrix} \qquad B^{(n)} := \begin{pmatrix} B & 0 & 0 & \cdots & 0 \\ 0 & B & 0 & \cdots & 0 \\ 0 & 0 & B & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & B \end{pmatrix}$$

This block structure implies that for any $n \in \mathbb{N}$, the Ornstein-Uhlenbeck semigroup $P^{(n)}$ associated to $(\Gamma^{(n)}, B^{(n)})$ is hypoelliptic and we have

$$\Gamma_\infty^{(n)} := \begin{pmatrix} \Gamma_\infty & 0 & 0 & \cdots & 0 \\ 0 & \Gamma_\infty & 0 & \cdots & 0 \\ 0 & 0 & \Gamma_\infty & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \Gamma_\infty \end{pmatrix}$$

In particular $\kappa(\Gamma_\infty^{(n)})$ does not depend on $n \in \mathbb{N}$ and $\boldsymbol{\alpha}^{(n)} = (\alpha, \ldots, \alpha) \in \mathbb{R}^{dn}$ and $\boldsymbol{b}^{(n)} = (b, \ldots, b) \in \mathbb{R}^{dn}$.

It follows from Corollary 19 that the family $(P^{(n)})_{n\in\mathbb{N}}$ has a cut-off at the times $(\log(n)/(2b_1))_{n\in\mathbb{N}}$.

3 Random warm-up time examples

In this section, we present several examples of interweaving relations for which the warm-up time is a positive random variable. This includes the family of Laguerre and Jacobi processes and examples of Subsection 2.3 that are extended in various directions either by playing with the underlying parameters or by pertubating their generator by a non-local component, that is by adding jumps in their dynamics. We also describe several interesting applications of interweaving relations in these contexts.

3.1 Diffusive Laguerre operators

The classical Laguerre generators $L_{\beta,\sigma}$, for $\beta, \sigma > 0$, were recalled in Subsection 2.3. Here we will drop the second parameter $\sigma > 0$, since we are more interested in the parameter $\beta > 0$: we would like to counter the bad behavior of the logarithmic Sobolev constant for small $\beta > 0$ via interweaving relations, in the spirit of what we have done for the two-point state space in Subsection 2.1. Namely we are looking for interweaving relations between Laguerre semigroups with different parameters $\beta > 0$.

For any $\beta > 0$, we write simply $L_\beta := L_{\beta,1}$, $\nu_\beta := \nu_{\beta,1}$ and $P^{(\beta)} := P^{(\beta,1)}$, with the notations of Subsection 2.3. For any $\beta, \varepsilon > 0$, consider the Markov kernel $\Lambda_{\beta_\varepsilon}$ from $\mathbb{R}_+$ to $\mathbb{R}_+$ corresponding to the multiplication by a Beta random variable of parameters ε and β, namely for any $f \in \mathbf{B}(\mathbb{R}_+)$, the set of bounded measurable mappings on $\mathbb{R}_+$,

$$\forall\, x \in \mathbb{R}_+, \qquad \Lambda_{\beta_\varepsilon}[f](x) \;\; := \;\; \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)} \int_0^1 f(rx) r^{\varepsilon-1}(1-r)^{\beta-1}\, dr$$

Its interest for us, is that according to Patie and Savov [36] we have the intertwining relation

$$\forall\, \beta > \varepsilon > 0, \forall\, t \geq 0, \qquad P_t^{(\beta+\varepsilon)} \Lambda_{\beta_\varepsilon} \;\; = \;\; \Lambda_{\beta_\varepsilon} P_t^{(\varepsilon)}$$

where the products are understood as the compositions of Markov kernels. They can also be seen as compositions of operators acting on $\mathbf{L}^2$-spaces and we have the following commuting diagram for any $\beta > \varepsilon > 0$ and $t \geq 0$:

$$\begin{array}{ccc} \mathbf{L}^2(\nu_{\beta+\varepsilon}) & \xrightarrow{P_t^{(\beta+\varepsilon)}} & \mathbf{L}^2(\nu_{\beta+\varepsilon}) \\ {\scriptstyle \Lambda_{\beta_\varepsilon}}\downarrow & & \downarrow{\scriptstyle \Lambda_{\beta_\varepsilon}} \\ \mathbf{L}^2(\nu_\varepsilon) & \xrightarrow{P_t^{(\varepsilon)}} & \mathbf{L}^2(\nu_\varepsilon) \end{array}$$

Figure 4: Intertwining relation between $P_t^{(\beta+\varepsilon)}$ and $P_t^{(\varepsilon)}$

To get an intertwining relation in the reverse direction, we pass to the adjoint relations, taking into account that $P_t^{(\beta+\varepsilon)}$ and $P_t^{(\varepsilon)}$ are self-adjoint in $\mathbf{L}^2(\nu_{\beta+\varepsilon})$ and $\mathbf{L}^2(\nu_\varepsilon)$ respectively:

$$\begin{array}{ccc} \mathbf{L}^2(\nu_\varepsilon) & \xrightarrow{P_t^{(\varepsilon)}} & \mathbf{L}^2(\nu_\varepsilon) \\ {\scriptstyle \Lambda^*_{\beta_\varepsilon}}\downarrow & & \downarrow{\scriptstyle \Lambda^*_{\beta_\varepsilon}} \\ \mathbf{L}^2(\nu_{\beta+\varepsilon}) & \xrightarrow{P_t^{(\beta+\varepsilon)}} & \mathbf{L}^2(\nu_{\beta+\varepsilon}) \end{array}$$

Figure 5: Intertwining relation between $P_t^{(\varepsilon)}$ and $P_t^{(\beta+\varepsilon)}$

where $\Lambda_{\beta_\varepsilon}^* : \mathbf{L}^2(\nu_{\beta+\varepsilon}) \to \mathbf{L}^2(\nu_\varepsilon)$ is the adjoint operator of $\Lambda_{\beta_\varepsilon} : \mathbf{L}^2(\nu_\varepsilon) \to \mathbf{L}^2(\nu_{\beta+\varepsilon})$.

Since $\nu_{\beta+\varepsilon}$ and ν_ε are both probability measures, it is known a priori that $\Lambda_{\beta_\varepsilon}^*$ corresponds to a Markov kernel. Let us compute it more precisely:

Lemma 20 *We have for any $\beta, \varepsilon > 0$ and any $g \in \mathbf{B}(\mathbb{R}_+)$,*

$$\forall\ x \in \mathbb{R}_+, \qquad \Lambda_{\beta_\varepsilon}^*[g](x) \;=\; \frac{x^\beta}{\Gamma(\beta)} \int_0^{+\infty} g((1+s)x)\, s^\beta \exp(-sx)\, ds$$

Proof: For any $f, g \in \mathbf{B}(\mathbb{R}_+)$, we compute

$$\begin{aligned}
\nu_{\beta+\varepsilon}[g\Lambda_{\beta_\varepsilon}[f]] &= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)} \int_{\mathbb{R}_+} g(x) \left(\int_0^1 f(rx) r^{\varepsilon-1}(1-r)^{\beta-1}\, dr \right) \frac{x^{\beta+\varepsilon-1}\exp(-x)}{\Gamma(\beta+\varepsilon)}\, dx \\
&= \int_0^1 \left(\int_{\mathbb{R}_+} g(x) f(rx) x^{\beta+\varepsilon-1} \exp(-x)\, dx \right) r^{\varepsilon-1}(1-r)^{\beta-1}\, dr \\
&= \int_0^1 \left(r^{-(\beta+\varepsilon)} \int_{\mathbb{R}_+} g(x/r) f(x) x^{\beta+\varepsilon-1} \exp(-x/r)\, dx \right) r^{\varepsilon-1}(1-r)^{\beta-1}\, dr \\
&= \int_{\mathbb{R}_+} f(x) \left(x^\beta \int_0^1 g(x/r)\, r^{\varepsilon-\beta+\varepsilon-1}(1-r)^{\beta-1} \exp(-x(1/r-1))\, dr \right) x^{\varepsilon-1}\exp(-x)\, dx
\end{aligned}$$

Since the last expression must be equal to $\Gamma(\beta)\Gamma(\varepsilon)\nu_\varepsilon[f\Lambda_{\beta+\varepsilon,\varepsilon}^*[g]]$, for any $f \in \mathbf{B}(\mathbb{R}_+)$, we obtain

$$\forall\ x \in \mathbb{R}_+, \qquad \Lambda_{\beta_\varepsilon}^*[g](x) \;=\; \frac{x^\beta}{\Gamma(\beta)} \int_0^1 g(x/r)\, \frac{1}{r^2} \left(\frac{1-r}{r} \right)^{\beta-1} \exp(-x(1-r)/r)\, dr$$

and we deduce the announced result via the change of variable $s = (1-r)/r$.

■

To get a c.mi.r., let us compute the Markov kernel $\Lambda_{\beta_\varepsilon}\Lambda_{\beta_\varepsilon}^*$. Following the argumentation of Remark 1(f), we know a priori that $\Lambda_{\beta_\varepsilon}\Lambda_{\beta_\varepsilon}^*$ commutes with the $P_t^{(\beta+\varepsilon)}$, for all $t \geq 0$. Since $L_{\beta+\varepsilon}$ is diagonalizable in $\mathbf{L}^2(\nu_{\beta+\varepsilon})$ and all its eigenvalues are non-positive and simple, it follows from functional calculus that $\Lambda_{\beta_\varepsilon}\Lambda_{\beta_\varepsilon}^*$ is of the form $F(-L_{\beta+\varepsilon})$, where $F : \mathbb{R}_+ \to \mathbb{R}$ is a measurable mapping. Here is its explicit formula:

Proposition 21 *For any $\beta, \varepsilon > 0$, we have*

$$\Lambda_{\beta_\varepsilon}\Lambda_{\beta_\varepsilon}^* \;=\; F_{\beta_\varepsilon}(-L_{\beta+\varepsilon}) \tag{32}$$

with

$$\forall\ u \in \mathbb{R}_+, \qquad F_{\beta_\varepsilon}(u) \;\coloneqq\; \int_0^\infty e^{-us}\mathbb{P}(\tau^{(\beta_\varepsilon)} \in ds) = \frac{\Gamma(\beta+\varepsilon)\Gamma(u+\varepsilon)}{\Gamma(\varepsilon)\Gamma(u+\beta+\varepsilon)}$$

and

$$\forall\ s \geq 0, \qquad \mathbb{P}(\tau^{(\beta_\varepsilon)} \in ds) \;\coloneqq\; \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)} \exp(-\varepsilon s)(1-\exp(-s))^{\beta-1}\, ds \tag{33}$$

Similarly, we have, still for $\beta, \varepsilon > 0$,

$$\Lambda_{\beta_\varepsilon}^*\Lambda_{\beta_\varepsilon} \;=\; F_{\beta_\varepsilon}(-L_\varepsilon)$$

Proof: It is well-known (see e.g. the book of Szegö [47]) that the spectrum of $-L_{\beta+\varepsilon}$ is $\mathbb{Z}_+$ and for each eigenvalue $n \in \mathbb{Z}_+$, an associated eigenvector is the Laguerre polynomial $\mathcal{L}_n^{(\beta+\varepsilon)}$ of degree n. It follows that to prove (32), it is sufficient to show that for any $n \in \mathbb{Z}_+$, we have

$$\Lambda_{\beta_\varepsilon}\Lambda^*_{\beta_\varepsilon}[\mathcal{L}_n^{(\beta+\varepsilon)}] \;=\; F_{\beta_\varepsilon}(n)[\mathcal{L}_n^{(\beta+\varepsilon)}]$$

From the commutation of $\Lambda_{\beta_\varepsilon}\Lambda^*_{\beta_\varepsilon}$ with the $P_t^{(\beta+\varepsilon)}$ for all $t \geq 0$, we know a priori that the l.h.s. is proportional to $\mathcal{L}_n^{(\beta+\varepsilon)}$. Thus, denoting $p_n : \mathbb{R}_+ \ni x \mapsto x^n$, the monomial of degree n, it is sufficient to check that $\Lambda_{\beta_\varepsilon}\Lambda^*_{\beta_\varepsilon}[p_n]$ is equal to $F_{\beta_\varepsilon}(n)p_n$, up to a polynomial of degree $n-1$. This operation can be decomposed into two similar sub-tasks. Indeed from Figure 4 we deduce that for any $t \geq 0$,

$$P_t^{(\beta+\varepsilon)}\Lambda_{\beta_\varepsilon}[\mathcal{L}_n^{(\varepsilon)}] \;=\; \Lambda_{\beta_\varepsilon}P_t^{(\varepsilon)}[\mathcal{L}_n^{(\varepsilon)}] \;=\; \exp(-nt)\Lambda_{\beta_\varepsilon}[\mathcal{L}_n^{(\varepsilon)}]$$

namely $\Lambda_{\beta_\varepsilon}[\mathcal{L}_n^{(\varepsilon)}]$ is proportional to $\mathcal{L}_n^{(\beta+\varepsilon)}$. So let $\widetilde{F}_{\beta_\varepsilon}(n) \in \mathbb{R}$ be such that $\Lambda_{\beta_\varepsilon}[p_n]$ is equal to $\widetilde{F}_{\beta_\varepsilon}(n)p_n$, up to a polynomial of degree $n-1$. Similarly, taking into account Figure 5, there exists $\widehat{F}_{\beta_\varepsilon}(n) \in \mathbb{R}$ such that $\Lambda^*_{\beta_\varepsilon}[p_n]$ is equal to $\widehat{F}_{\beta_\varepsilon}(n)p_n$, up to a polynomial of degree $n-1$. It follows that $F_{\beta_\varepsilon}(n) = \widetilde{F}_{\beta_\varepsilon}(n)\widehat{F}_{\beta_\varepsilon}(n)$ and we just need to compute $\widetilde{F}_{\beta_\varepsilon}(n)$ and $\widehat{F}_{\beta_\varepsilon}(n)$. Let us start with $\widetilde{F}_{\beta_\varepsilon}(n)$. We have for any $x \in \mathbb{R}_+$,

$$\begin{aligned}
\Lambda_{\beta_\varepsilon}[p_n](x) &= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)}\int_0^1 (rx)^n r^{\varepsilon-1}(1-r)^{\beta-1}\,dr \\
&= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)}\left(\int_0^1 r^{n+\varepsilon-1}(1-r)^{\beta-1}\,dr\right)x^n \\
&= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)}\frac{\Gamma(n+\varepsilon)\Gamma(\beta)}{\Gamma(n+\beta+\varepsilon)}p_n(x)
\end{aligned} \tag{34}$$

and thus

$$\widetilde{F}_{\beta_\varepsilon}(n) \;=\; \frac{\Gamma(\beta+\varepsilon)\Gamma(n+\varepsilon)}{\Gamma(\varepsilon)\Gamma(n+\beta+\varepsilon)} \;=\; \frac{(n+\varepsilon-1)(n+\varepsilon-2)\cdots\varepsilon}{(n+\beta+\varepsilon-1)(n+\beta+\varepsilon-2)\cdots\beta+\varepsilon}. \tag{35}$$

On the other hand, for $\widehat{F}_{\beta_\varepsilon}(n)$, we have for any $x \in \mathbb{R}_+$,

$$\begin{aligned}
\Lambda^*_{\beta_\varepsilon}[p_n](x) &= \frac{x^\beta}{\Gamma(\beta)}\int_0^{+\infty}((1+s)x)^n\, s^{\beta-1}\exp(-sx)\,ds \\
&= \frac{x^\beta}{\Gamma(\beta)}\left(\int_0^{+\infty}(1+s)^n\, s^{\beta-1}\exp(-sx)\,ds\right)x^n \\
&= \frac{1}{\Gamma(\beta)}\left(\int_0^{+\infty}(1+s/x)^n\, s^{\beta-1}\exp(-s)\,ds\right)x^n \\
&= \sum_{m=0}^n \binom{n}{m}\frac{1}{\Gamma(\beta)}\left(\int_0^{+\infty}s^m\, s^{\beta-1}\exp(-s)\,ds\right)x^{n-m}.
\end{aligned}$$

It follows that

$$\widehat{F}_{\beta_\varepsilon}(n) \;=\; \binom{n}{0}\frac{1}{\Gamma(\beta)}\int_0^{+\infty} s^{\beta-1}\exp(-s)\,ds = \frac{1}{\Gamma(\beta)}\Gamma(\beta) = 1$$

Thus we get that for all $n \in \mathbb{Z}_+$, $F_{\beta_\varepsilon}(n) = \widetilde{F}_{\beta_\varepsilon}(n)$. Coming back to (34), it appears that for any $n \in \mathbb{Z}_+$,

$$\begin{aligned} F_{\beta_\varepsilon}(n) &= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)} \int_0^1 r^{n+\varepsilon-1}(1-r)^{\beta-1}\, dr \\ &= \frac{\Gamma(\beta+\varepsilon)}{\Gamma(\beta)\Gamma(\varepsilon)} \int_0^{+\infty} \exp(-ns)\exp(-\varepsilon s)(1-\exp(-s))^{\beta-1}\, ds \end{aligned}$$

where we considered the change of variable $r = \exp(-s)$. It justifies (32).

The last assertion of the proposition is proven similarly, or by applying the $\mathbf{L}^2$-version of Proposition 4: $\Lambda_{\beta_\varepsilon}$ is one-to-one, since it transforms the orthogonal basis $(\mathcal{L}_n^{(\varepsilon)})_{n\in\mathbb{Z}_+}$ of $\mathbf{L}^2(\nu_\varepsilon)$ into an orthogonal basis of $\mathbf{L}^2(\nu_{\beta+\varepsilon})$:

$$\forall\, n \in \mathbb{Z}_+, \qquad \Lambda_{\beta_\varepsilon}[\mathcal{L}_n^{(\varepsilon)}] \;=\; \widetilde{F}_{\beta_\varepsilon}(n)\mathcal{L}_n^{(\beta+\varepsilon)}$$

where $\widetilde{F}_{\beta_\varepsilon}(n) > 0$ is given in (35).

■

Thus we have shown the symmetric c.m.i.r. between $P_t^{(\beta+\varepsilon)}$ and $P_t^{(\varepsilon)}$ described in the following Figure 6, for any $\beta+\varepsilon > \varepsilon > 0$ and $t \geq 0$:

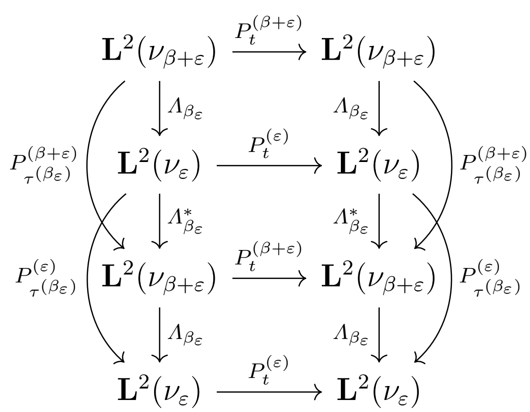

Figure 6: interweaving relations between $P_t^{(\beta+\varepsilon)}$ and $P_t^{(\varepsilon)}$

Since we are interested in the behavior for small shape parameter, let us denote for $\beta+\varepsilon \in (0,1/2)$, $\tau_{\beta+\varepsilon} \coloneqq \pi_{1/2,\beta+\varepsilon}$. We deduce the following bound from Theorem 8 and from the fact that $\alpha(1/2) = 1$:

Corollary 22 *For any $\varepsilon \in (0,1/2)$ and any $m_0 \in \mathcal{P}((0,+\infty))$,*

$$\forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 P^{(\varepsilon)}_{t+\tau(\beta_\varepsilon)}|\nu_\varepsilon) \;\leq\; \exp(-t)\mathrm{Ent}(m_0|\nu_\varepsilon) \tag{36}$$

Recall the estimate directly obtained by applying the logarithmic Sobolev inequality satisfied by the generator L_ε for any $\varepsilon \in (0,1/2)$:

$$\forall\, m_0 \in \mathcal{P}((0,+\infty)),\, \forall\, t \geq 0, \qquad \mathrm{Ent}(m_0 P_t^{(\varepsilon)}|\nu_\varepsilon) \;\leq\; \exp(-\alpha(\varepsilon)t)\mathrm{Ent}(m_0|\nu_\varepsilon) \tag{37}$$

(where $\alpha(\varepsilon)$ is defined in (14)). The bounds (36) and (37) are not directly comparable, since they concern different distributions, namely $m_0 P^{(\varepsilon)}_{t+\tau(\beta_\varepsilon)}$ and $m_0 P_t^{(\varepsilon)}$ and the former is not just a deterministic time translate through P of the latter. Nevertheless, to highlight the potential advantage of (36), let us make the following observation. Let $X^{(\varepsilon)} := (X_t^{(\varepsilon)})_{t\geq 0}$ be a diffusion process associated to the Markov semigroup $P^{(\varepsilon)}$, with small $\varepsilon \in (0,1/2)$, starting with $X_0^{(\varepsilon)}$ uniformly distributed over $[0,1]$. We want to use this trajectory to sample according to ν_ε, with an accuracy given by $\delta > 0$ in the entropy sense. Relying on (37), we consider the position $X_{T_1}^{(\varepsilon)}$ at the time $T_1 \geq 0$ such that

$$\exp(-\alpha(\varepsilon)T_1)\mathrm{Ent}(m_0|\nu_\varepsilon) \quad \leq \quad \delta$$

for some $\delta > 0$. Letting ε going to 0_+ and recalling that $\alpha(\varepsilon) \sim 4/\ln(1/\varepsilon)$, we easily compute that

$$\begin{aligned}
\mathrm{Ent}(m_0|\nu_\varepsilon) &= \int_0^1 \ln(\Gamma(\varepsilon)x^{1-\varepsilon}\exp(x))\,dx = \ln(\Gamma(\varepsilon)) + (1-\varepsilon)\int_0^1 \ln(x)\,dx + \int_0^1 x\,dx \\
&= \ln(\Gamma(\varepsilon)) + \varepsilon - 1/2 \sim \ln(\Gamma(\varepsilon)) \\
&\sim \ln(1/\varepsilon)
\end{aligned}$$

So we get that

$$T_1 \quad \simeq \quad \ln(1/\varepsilon)\ln(\ln(1/\varepsilon)/\delta)/4 = \ln(1/\varepsilon)(\ln(\ln(1/\varepsilon) + \ln(1/\delta))/4$$

Relying on (36), we consider the position $X_{T_2+T_3}^{(\varepsilon)}$, where T_2 is independent from $X^{(\varepsilon)}$ and has the same law than τ_ε and $T_3 \geq 0$ is such that

$$\exp(-T_3)\mathrm{Ent}(m_0|\nu_\varepsilon) \quad \leq \quad \delta$$

namely

$$T_3 \quad \simeq \quad \ln(\ln(1/\varepsilon)/\delta)$$

To get a rough idea of T_2, let us compute its expectation, as ε goes to zero:

$$\begin{aligned}
\mathbb{E}[T_2] &= \int_0^{+\infty} s\,\mathbb{P}(\tau^{(\beta_\varepsilon)} \in ds) = \frac{\Gamma(1/2)}{\Gamma(1/2-\varepsilon)\Gamma(\varepsilon)}\int_0^{+\infty} s\exp(-\varepsilon s)(1-\exp(-s))^{-\varepsilon-1/2}\,ds \\
&= -\frac{\Gamma(1/2)}{\Gamma(1/2-\varepsilon)\Gamma(\varepsilon)}\int_0^1 \ln(r)r^{\varepsilon-1}(1-r)^{-\varepsilon-1/2}\,dr \\
&\sim -\frac{1}{\Gamma(\varepsilon)}\int_0^1 \ln(r)r^{\varepsilon-1}\,dr = \frac{1}{\Gamma(\varepsilon+1)}\int_0^1 r^{\varepsilon-1}\,dr = \frac{1}{\varepsilon\Gamma(\varepsilon+1)} \\
&\sim \frac{1}{\varepsilon}
\end{aligned}$$

(where an integration by parts was used for the fourth equality), and thus

$$\mathbb{E}[T_2] + T_3 \quad \simeq \quad \frac{1}{\varepsilon} + \ln(1/\delta)$$

When $\delta > 0$ is very small, e.g. of order $\exp(-1/\varepsilon)$, the quantity $\mathbb{E}[T_2]+T_3$ is much smaller than T_1, suggesting that the approach based on (36) is a more effcient sampling procedure.

Similar observations are also valid for hyperboundedness, as we deduce from Theorem 9:

Corollary 23 *For any $\varepsilon \in (0,1/2)$, we have*

$$\forall\, t \geq 0, \qquad \|P_{t+\tau^{(\beta_\varepsilon)}}^{(\varepsilon)}\|_{\mathbf{L}^2(\nu_\varepsilon)\to\mathbf{L}^{p(t)}(\nu_\varepsilon)} \quad \leq \quad 1 \tag{38}$$

where

$$\forall\, t \geq 0, \qquad p(t) \quad := \quad 1 + \exp(t)$$

Note that for small $\varepsilon > 0$ and large $t \geq 0$, the exponent $p(t)$ is much larger than $1 + \exp(\alpha(\varepsilon)t)$, the quantity one gets via the traditional application of the logarithmic Sobolev associated to L_ε. Thus up to waiting a warm-up time variable $\tau^{(\beta_\varepsilon)}$), the hyperboundedness estimate (38) is more interesting than the usual hypercontractive bound.

3.2 The Jacobi processes

We proceed with another important and classical example in the theory of diffusions which is the Jacobi semigroup $\mathrm{J}^{(\beta)} = (\mathrm{J}_t^{(\beta)})_{t\geq 0}$. Its infinitesimal generator is defined for a function $f \in \mathrm{C}^2(V)$, the space of twice continuously differentiable functions on $V = [0, 1]$, by

$$J_\beta[f](x) = x(1-x)f''(x) + (\lambda_1 - \beta - \lambda_1 x)f'(x), \quad x \in [0,1], \tag{39}$$

where $\lambda_1 \geq 2\beta > 2$ and refer here and below to [16, Section 5] for a thorough review of the Jacobi semigroup.

It admits as unique invariant measure ν_β, the distribution of a beta $B(\lambda_1, \beta)$ random variable, defined on $(0, 1)$ as

$$\nu_\beta(dx) = \frac{\Gamma(\lambda_1)}{\Gamma(\lambda_1-\beta)\Gamma(\beta)} x^{\beta-1}(1-x)^{\lambda_1-\beta-1}dx, \quad 0 < x < 1.$$

As a by-product, the Hölder inequality yields that $\mathrm{J}^{(\beta)}$ extends to a contraction semigroup from the Hilbert space $\mathbf{L}^2(\nu_\beta)$ into itself. We recall that for any $n \in \mathbb{N}$,

$$\int_0^\infty x^n \nu_\beta(dx) = \frac{\Gamma(\lambda_1)}{\Gamma(\beta)} \frac{\Gamma(n+\beta)}{\Gamma(n+\lambda_1)}.$$

We say that the Jacobi operator is symmetric when $\lambda_1 = 2\beta$ and, in this case, we write $\widetilde{\mathrm{J}} = (\widetilde{\mathrm{J}}_t)_{t\geq 0}$ for the symmetric Jacobi semigroup whose infinitesimal generator is $\widetilde{J} = J_{\frac{\lambda_1}{2}}$ that is

$$\widetilde{J}[f](x) = x(1-x)f''(x) + \frac{\lambda_1}{2}(1-2x)f'(x), \quad 0 < x < 1.$$

We remark that, when $\frac{\lambda_1}{2} = n \in \mathbb{N}$, there exists a homeomorphism between J_β and the radial part of the Laplace-Beltrami operator on the n-sphere, which leads to the curvature-dimension condition $CD(\lambda_1 - 1, \lambda_1)$, see [9] for the definition. We deduce from [16, Proposition 3.6], choosing in the notation thereout $\mu = \frac{\lambda_1}{2}$ and $\hbar \equiv 0$, the following interweaving relation between the symmetric and other Jacobi semigroups.

Proposition 24 *For any $\lambda_1 > 2\beta > 1$, we have*

$$\widetilde{\mathrm{J}} \overset{\tau^{(\lambda_1,\beta)}}{\leftrightsquigarrow} \mathrm{J}^{(\beta)}$$

with

$$\forall\, u \in \mathbb{R}_+, \qquad \int_0^\infty e^{-us}\mathbb{P}(\tau_\phi^{(\lambda_1,\beta)} \in ds) = \frac{\Gamma(\lambda_1-\beta)\Gamma(\rho(u)+\frac{\lambda_1}{2})}{\Gamma(\rho(u)+\lambda_1-\beta)\Gamma(\frac{\lambda_1}{2})}$$

where $\rho(u) = \sqrt{u + \frac{(\lambda_1-1)^2}{4}} - \frac{\lambda_1-1}{2}$.

As a self-adjoint operator J_β has nice spectral properties: its spectrum is discrete with simple eigenvalues given by the set $(-n(n-1) - \lambda_1 n)_{n\geq 0}$. Moreover, it satisfies certain functional inequalities which give some quantitative rates of convergence to the equilibrium measure ν_β. For instance, from the Poincaré inequality for J_β, see [9, Chapter 4.2], one gets the following variance decay estimate, valid for any $f \in \mathbf{L}^2(\nu_\beta)$ and $t \geq 0$,

$$\mathrm{Var}_{\nu_\beta}(\mathrm{J}_t^{(\beta)}[f]) \leq e^{-2\lambda_1 t}\mathrm{Var}_{\nu_\beta}(f),$$

where for a measure ν, we have set $\mathrm{Var}_\nu(f) = ||f - \nu f||^2_{\mathbf{L}^2(\nu)}$. Next, note, writing

$$\bar{J}_\beta[f](x) = (1 - x^2)f''(x) + (\lambda_1 - 2\beta - \lambda_1 x)f'(x) \tag{40}$$

and $g(x) = \frac{x+1}{2}$, that

$$\bar{J}_\beta[f \circ g](g^{-1}(x)) \;=\; x(1-x)f''(x) + (\lambda_1 - \beta - \lambda_1 x)f'(x) = J_\beta[f](x).$$

Then, the log-Sobolev constant being invariant by homeomorphism, one gets, from Saloff-Coste [44], see also Fontenas [24], that the log-Sobolev constant $\alpha(\lambda_1, \beta)$ of the Jacobi operator J_β is such that

$$\alpha\left(\lambda_1, \frac{\lambda_1}{2}\right) = \frac{\lambda_1}{2} \tag{41}$$

for the symmetric Jacobi and otherwise, $\alpha(\lambda_1, \beta) < \frac{\lambda_1}{2}$ for $\lambda_1 > 2\beta$, with for any fixed β and large λ_1, $\alpha(\lambda_1, \beta) \sim \frac{\lambda_1}{4}$. Since always $\alpha(\lambda_1, \beta) \leq 2\lambda_1$, we thus get, from (41), that the symmetric Jacobi semigroup attains the optimal entropic decay and hypercontractivity rate. We point out that the explicit expression of the log-Sobolev constant for the symmetric case goes back to Barky in [8]. Although the log-Sobolev constant is not attainable in the other cases, the interweaving relation described above combined with theorems 8 and 9 enable us to provide the following information regarding the non-symmetric Jacobi semigroups.

Proposition 25 *For any $\lambda_1 > 2\beta > 1$, $m_0 \in \mathcal{P}((0,1))$ and $t \geq 0$, we have*

$$\mathrm{Ent}(m_0 \mathrm{J}^{(\beta)}_{t+\tau(\lambda_1,\beta)} | \nu_\beta) \;\leq\; e^{-\frac{\lambda_1}{2}t}\, \mathrm{Ent}(m_0|\nu_\beta)$$

and

$$\|\mathrm{J}^{(\beta)}_{t+\tau(\lambda_1,\beta)}\|_{\mathbf{L}^2(\nu)\to\mathbf{L}^{p(t)}} \leq 1 \text{ where } p(t) = 1 + e^{\frac{\lambda_1}{2}t}. \tag{42}$$

We close this example by mentioning that in [16] interweaving relations are established between the symmetric Jacobi semigroup and a class of non-local and non-self-adjoint Markov semigroups on the unit interval $[0,1]$.

3.3 The non-self-adjoint generalized Laguerre semigroups

In this part, we illustrate that the concept of interweaving relation is also useful in the context of non-reversible and non-local Markov semigroups. More specifically, let $P = (P_t)_{t\geq 0}$ be the generalized Laguerre semigroup as introduced and thoroughly studied in [36]. We also refer to this paper for further details about the objects that will be introduced in this part. It can be characterized through its infinitesimal generator which takes the form, for a function f smooth,

$$L_\phi[f](x) = xf''(x) + (m + 1 - x)\, f'(x) + \int_0^\infty \left(f(e^{-y}x) - f(x)\right) \mathit{\Pi}(x, dy), \quad x > 0,$$

where $m \geq 0$ and $\mathit{\Pi}(x, dy) = \frac{\mathit{\Pi}(dy)}{x}$ with $\mathit{\Pi}$ a finite non-negative Radon measure on $\mathbb{R}^+$ with a finite first moment, that is $\overline{\Pi} = \int_0^\infty y\mathit{\Pi}(dy) < \infty$. Observe that, writing $p_n(x) = x^n, x > 0$, $n \in \mathbb{N}$, an integration by parts yields

$$L_\phi[p_n](x) = n\phi(n)p_{n-1}(x) - np_n(x),$$

where, for $u \geq 0$, we have set

$$\phi(u) = u + m + \int_0^\infty (e^{-uy} - 1)\mathit{\Pi}(u, dy). \tag{43}$$

Note that ϕ is a Bernstein function and it is in fact the Laplace exponent of the descending ladder height process $\xi = (\xi_t)_{t\geq 0}$ of the spectrally negative Lévy process with Laplace exponent $u\phi(u)$, see e.g. [26, Sec. 6.5.2]. P admits an unique invariant measure which is an absolutely continuous probability measure with a density denoted by ν. Its law is determined by its integer moments which are given, for any $n \in \mathbb{N}$, by

$$\int_0^\infty x^n \nu_\phi(x)dx = W_\phi(n+1)$$

where $W_\phi(1) = 1$ and $W_\phi(n+1) = \prod_{k=1}^n \phi(k)$. P extends to a non-self-adjoint strongly continuous contraction semigroup on $\mathrm{L}^2(\nu_\phi)$. Next, let $\widetilde{P}^{(\beta)} = (\widetilde{P}_t^{(\beta)})_{t\geq 0}$ denotes the semigroup of the classical Laguerre process of index $\beta \geq 0$ (or dimension $\beta+1$) and recall from Section 2.3 that its generator is the differential operator

$$L_{\beta+1}[f](x) = xf''(x) + (\beta + 1 - x)\, f'(x), \quad x > 0.$$

$\widetilde{P}^{(\beta)}$ is a self-adjoint operator on $\mathrm{L}^2(\nu_\beta)$ where here, for sake of simplicity, we write $\nu_\beta(dx) = \frac{x^\beta}{\Gamma(\beta+1)}e^{-x}dx, x > 0$. We disregard the parameter β when it is 0, that is we simply write $\widetilde{P} = \widetilde{P}^{(0)}$ and $\nu = \nu_0$.

Now, according to [36], there exists a multiplicative Markov kernel I_ϕ defined by

$$\mathrm{I}_\phi[f](x) = \mathbb{E}\left[f(xI_\phi)\right], \quad x > 0, \tag{44}$$

where $I_\phi = \int_0^\infty e^{-\xi_t}dt$ with ξ the subordinator with Laplace exponent the Bernstein function ϕ and, for any $n \in \mathbb{N}$,

$$\mathrm{I}_\phi[p_n](x) = \frac{\Gamma(n+1)}{W_\phi(n+1)}p_n(x), \quad x > 0. \tag{45}$$

We also introduce for any $\beta > 0$, the Markov kernel B^*_β, acting on any bounded Borelian function f via

$$\mathrm{B}^*_\beta[f](x) = \frac{x^\beta}{\Gamma(\beta)}\int_0^\infty f((1+y)x)y^{\beta-1}e^{-yx}dy, \quad x > 0. \tag{46}$$

We are ready to state and proof the following.

Proposition 26 *For any $\beta > \overline{\Pi} + m$, we have*

$$P \overset{\tau^{(\beta)}}{\leftrightsquigarrow} \widetilde{P}^{(\beta)}$$

where $\tau^{(\beta)}$ is an infinitely divisible variable characterized by

$$\int_0^\infty e^{-us}\mathbb{P}(\tau^{(\beta)} \in ds) = \left(\frac{\Gamma(1+\beta)\Gamma(u+1)}{\Gamma(u+\beta+1)}\right) = e^{-\phi_\beta(u)t}, \qquad u > 0. \tag{47}$$

In particular, $\mathbb{P}(\tau^{(\beta)} \in ds) = (1+\beta)(1+\log s)^\beta ds, s \in (1/e, 1)$. Moreover, for any such β, we have

$$\Lambda = \mathrm{I}_\phi \mathrm{B}^*_\beta \text{ and } \widetilde{\Lambda} = \mathrm{V}_\beta \tag{48}$$

where V_β is a Markov kernel associated to the variable Y_β whose distribution is determined by its moments given by, for any $n \in \mathbb{N}$,

$$\mathrm{V}_\beta[p_n](x) = \mathbb{E}\left[p_n(xY_\beta)\right] = \Gamma(1+\beta)\frac{W_\phi(n+1)}{\Gamma(n+1+\beta)}p_n(x), \quad x > 0. \tag{49}$$

Finally, we have for any $t \geq 0$ and $m_0 \in \mathcal{P}((0,+\infty))$,

$$\mathrm{Ent}(m_0 P_{t+\tau^{(\beta)}}|\nu_\phi) \leq e^{-t}\mathrm{Ent}(m_0|\nu_\phi), \tag{50}$$

and

$$\|P_{t+\tau^{(\beta)}}\|_{\mathbf{L}^2(\nu_\phi)\to\mathbf{L}^{p(t)}(\nu_\phi)} \leq 1 \text{ where } p(t) = 1 + e^t. \tag{51}$$

Proof: First, we recall from [36, Theorem 7.1] that the following intertwining relationship

$$P_t \mathrm{I}_\phi = \mathrm{I}_\phi \widetilde{P}_t, \quad t \geq 0, \tag{52}$$

holds in $\mathrm{L}^2(\nu)$. Next, [36, Proposition 4.4] entails that, for any $\beta > \overline{\Pi} + m$, $\phi_\beta(u) = \frac{\phi(u)}{u+\beta}$ is a Bernstein function and there exists a Markov kernel V_β associated to the positive random variable Y_β whose moments are given by (49) and determined its law. Moreover, from Lemma 10.2 of the aforementioned paper, we have, in $\mathrm{L}^2(\nu_\phi)$, the following identity

$$\widetilde{P}_t^{(\beta)} \mathrm{V}_\beta = \mathrm{V}_\beta P_t, \quad t \geq 0. \tag{53}$$

Then, invoking either [15, Identity (1.c)] or again [36, Theorem 7.1], we have in $\mathrm{L}^2(\varepsilon)$

$$\widetilde{P}_t^{(\beta)} \mathrm{B}_\beta = \mathrm{B}_\beta \widetilde{P}_t.$$

Taking the adjoint, in the weighted Hilbert space, intertwining identity and using the fact that $\widetilde{P}$ (resp. $\widetilde{P}_t^{(\beta)}$) is self-adjoint in $\mathrm{L}^2(\nu)$ (resp. $\mathrm{L}^2(\nu_\beta)$) yields in $\mathrm{L}^2(\nu_\beta)$

$$\widetilde{P}_t \mathrm{B}_\beta^* = \mathrm{B}_\beta^* \widetilde{P}_t^{(\beta)} \tag{54}$$

Combining this with (52) entails that in $\mathrm{L}^2(\nu_\beta)$

$$P_t \mathrm{I}_\phi \mathrm{B}_\beta^* = \mathrm{I}_\phi \widetilde{P} \mathrm{B}_\beta^* = \mathrm{I}_\phi \mathrm{B}_\beta^* \widetilde{P}_t^{(\beta)}.$$

Finally, this combines with the intertwining relationship (53) yields the identity in $\mathrm{L}^2(\nu_\phi)$

$$P_t \mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta = \mathrm{I}_{\phi_\mathrm{d}} \mathrm{B}_\beta^* \widetilde{P}_t^{(\beta)} \mathrm{V}_\beta = \mathrm{I}_{\phi_d} \mathrm{B}_\beta^* \mathrm{V}_\beta P_t \tag{55}$$

and

$$\widetilde{P}_t^{(\beta)} \mathrm{V}_\beta \mathrm{I}_\phi \mathrm{B}_\beta^* = \mathrm{V}_\beta \mathrm{I}_\phi \mathrm{B}_\beta^* \widetilde{P}_t^{(\beta)}. \tag{56}$$

Since from [36, Theorem 7.1(2) and Lemma 8.16], we have that I_ϕ and B_β^* are one-to-one in $\mathrm{L}^2(\nu)$ and $\mathrm{L}^2(\nu_\beta)$ respectively, we get that their composition $\mathrm{I}_\phi \mathrm{B}_\beta^*$ is also one-to-one in $\mathrm{L}^2(\nu_\beta)$. Thus, it remains to show that $P_\tau = \mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta$ or, by Theorem 3, equivalently $\widetilde{P}_\tau^{(\beta)} = \mathrm{V}_\beta \mathrm{I}_\phi \mathrm{B}_\beta^*$. To justify the latter identity, we proceed as in the proof of Proposition 21, we have from [36, Theorem 1.22(c)], that, for any $t > 0$, the spectrum of P_t in $\mathrm{L}^2(\nu_\phi)$ is discrete and given by $e^{-t\mathbb{N}}$ and each eigenvalue is simple with for all $n \in \mathbb{N}$,

$$P_t[\mathcal{P}_n](x) = e^{-nt} \mathcal{P}_n(x)$$

where the polynomials $\mathcal{P}_n$ are defined via the identity $\mathcal{P}_n(x) = \mathrm{I}_\phi[\mathcal{L}_n](x)$, $(\mathcal{L}_n)_{n \geq 0}$ being the orthonormal sequence of Laguerre polynomials. Thus, we deduce from (55) that

$$P_t \mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta[\mathcal{P}_n] = \mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta P_t[\mathcal{P}_n] = e^{-nt} \mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta[\mathcal{P}_n],$$

that is $\mathrm{I}_\phi \mathrm{B}_\beta^* \mathrm{V}_\beta[\mathcal{P}_n]$ is proportional to $\mathcal{P}_n$. More specifically, recalling that for any $n \in \mathbb{N}$,

$$\mathrm{B}_\beta^*[p_n](x) = p_n(x) + P_{n-1}(x)$$

where here and below $P_{n-1}(x)$ stands for a generic polynomial of order $n-1$, we deduce from (45) and (49) that

$$\mathrm{V}_\beta \mathrm{I}_\phi \mathrm{B}_\beta^*[p_n](x) = \frac{\Gamma(n+1+\mathrm{d})}{\Gamma(1+\mathrm{d}) W_\phi(n+1)} \Gamma(1+\beta) \frac{W_\phi(n+1)}{\Gamma(n+1+\beta)} p_n(x) + P_{n-1}(x) \tag{57}$$

$$= \frac{\Gamma(1+\beta)\Gamma(n+1+\mathrm{d})}{\Gamma(1+\mathrm{d})\Gamma(n+1+\beta)} p_n(x) + P_{n-1}(x) \tag{58}$$

and hence

$$\mathrm{I}_\phi \mathrm{B}^*_\beta \mathrm{V}_\beta[\mathcal{P}_n](x) \quad = \quad \frac{\Gamma(1+\beta)\Gamma(n+1+\mathrm{d})}{\Gamma(1+\mathrm{d})\Gamma(n+1+\beta)}\mathcal{P}_n(x). \tag{59}$$

On the other hand, it is well known that $\phi_\beta(u) = -\log \frac{\Gamma(1+\beta)\Gamma(u+1+\mathrm{d})}{\Gamma(1+\mathrm{d})\Gamma(n+1+\beta)}$ is a Bernstein function which corresponds to the Laplace exponent of the positive infinitely divisible variable $\tau^{(\beta)} = -\log B_\beta$, where B_β is a beta variable of parameter $\beta > 0$. Finally, since $\beta > \overline{\Pi} + m > 0$, one gets that the log-Sobolev constant of the classical Laguerre $P^{(\beta)}$ is 1, see Remark 15. We complete the proof by invoking theorems 8 and 9.

3.3.1 Subordinate generalized Laguerre semigroups

It is well-known, see e.g. [9], that, for any $t > 0$, $\widetilde{P}^{(\beta)}_t$ is an Hilbert-Schmidt operator in $\mathrm{L}^2(\nu_\beta)$ that admits, for any $f \in \mathrm{L}^2(\nu_\beta)$, the diagonalization

$$\widetilde{P}^{(\beta)}_t[f] = \sum_{n=0}^{\infty} e^{-nt}\mathfrak{c}_n(\beta)\langle f, \mathcal{L}^{(\beta)}_n\rangle_{\nu_\beta}\, \mathcal{L}^{(\beta)}_n \tag{60}$$

where the sequence of Laguerre polynomials $(\sqrt{\mathfrak{c}_n(\beta)}\mathcal{L}^{(\beta)}_n)_{n\geq 0}$ forms an orthonormal basis of $\mathrm{L}^2(\nu_\beta)$ and we recall that

$$\mathcal{L}^{(\beta)}_n(x) = \sum_{r=0}^{n}(-1)^r\binom{n+\beta}{n-r}\frac{x^r}{r!}$$

and $\mathfrak{c}_n(\beta) = \frac{\Gamma(n+1)\Gamma(\beta+1)}{\Gamma(n+\beta+1)}$. Moreover, a classical argument based on the spectral theory of reversible compact Markov semigoups yields, for any $t \geq 0$ and $f \in \mathrm{L}^2(\nu_\beta)$, the spectral gap estimate

$$\mathrm{Var}_{\nu_\beta}\left(\widetilde{P}^{(\beta)}_t[f]\right) \leq e^{-t}\,\mathrm{Var}_{\nu_\beta}(f) \tag{61}$$

where, we recall that for a measure ν, we have set $\mathrm{Var}_\nu(f) = ||f-\nu[f]||^2_{\mathbf{L}^2(\nu)}$. Let us denote by $\widetilde{P}^{\tau^{(\beta)}} = (\widetilde{P}^{\tau^{(\beta)}}_t)_{t\geq 0}$ the Bochner subordination of $\widetilde{P}^{(\beta)}$ by the subordinator $(\tau^{(\beta)}_t)_{t\geq 0}$ where $\tau^{(\beta)}_1$ has the same law than the positive infinitely divisible variable $\tau^{(\beta)}$ defined in Proposition 26 and use the same notation for the subordinated semigroup $P^{\tau^{(\beta)}}$.

Corollary 27 *For any $\beta > 0$, $t > 0$, $\widetilde{P}^{\tau^{(\beta)}}_t$ is a self-adjoint Hilbert-Schmidt operator in $\mathbf{L}^2(\nu_\beta)$ that admits, for any $f \in \mathbf{L}^2(\nu_\beta)$, the diagonalization*

$$\widetilde{P}^{\tau^{(\beta)}}_t[f] = \sum_{n=0}^{\infty} \mathfrak{c}^{t+1}_n(\beta)\,\langle f, \mathcal{L}^{(\beta)}_n\rangle_{\nu_\beta}\, \mathcal{L}^{(\beta)}_n \tag{62}$$

and

$$\mathrm{Var}_{\nu_\beta}\left(\widetilde{P}^{\tau^{(\beta)}}_t[f]\right) \leq (1+\beta)^{-t}\,\mathrm{Var}_{\nu_\beta}(f) \tag{63}$$

Moreover, for any $\beta > \overline{\Pi} + m$, $P^{\tau^{(\beta)}} \overset{1}{\leftrightsquigarrow} \widetilde{P}^{\tau^{(\beta)}}$ and for any $f \in \mathrm{L}^2(\nu)$ and $t > 1$, we have in $\mathrm{L}^2(\nu)$

$$P^{\tau^{(\beta)}}_t[f] = \sum_{n=0}^{\infty} \mathfrak{c}^{t}_n(\beta)\,\langle f, \mathcal{V}_n\rangle_{\nu}\, \mathcal{P}_n \tag{64}$$

and for any $t \geq 0$

$$\mathrm{Var}_{\nu}\left(P^{\tau^{(\beta)}}_t[f]\right) \leq (1+\beta)^{1-t}\,\mathrm{Var}_{\nu}(f) \tag{65}$$

Proof: The fact that $\widetilde{P}^{\tau^{(\beta)}}$ is self-adjoint in $\mathbf{L}^2(\nu_\beta)$ can easily be checked by means of Fubini theorem as, for any non-negative $f,g\in \mathrm{L}^2(\nu_\beta)$ and $t\geq 0$,

$$\begin{aligned}\langle \widetilde{P}^{\tau^{(\beta)}}_t[f],g\rangle_{\nu_\beta} &= \int_0^\infty \langle \widetilde{P}^{(\beta)}_s[f],g\rangle_{\nu_\beta}\mathbb{P}(\tau^{(\beta)}_t\in ds) = \int_0^\infty \langle f,\widetilde{P}^{(\beta)}_s[g]\rangle_{\nu_\beta}\mathbb{P}(\tau^{(\beta)}_t\in ds)\\ &= \langle f,\widetilde{P}^{\tau^{(\beta)}}_t[g]\rangle_{\nu_\beta}\end{aligned}$$

where we used that $\widetilde{P}^{(\beta)}$ is self-adjoint in $\mathbf{L}^2(\nu_\beta)$. Next, one has that for any $f\in\mathbf{L}^2(\nu_\beta)$, the diagonalization

$$\begin{aligned}\widetilde{P}^{\tau^{(\beta)}}_t[f] &= \int_0^\infty \mathbb{P}(\tau^{(\beta)}_t\in ds)\widetilde{P}^{(\beta)}_s[f]\\ &= \int_0^\infty \mathbb{P}(\tau^{(\beta)}_t\in ds)\sum_{n=0}^\infty e^{-sn}\mathfrak{c}_n(\beta)\,\langle f,\mathcal{L}^{(\beta)}_n\rangle_{\nu_\beta}\,\mathcal{L}^{(\beta)}_n\\ &= \sum_{n=0}^\infty\left(\frac{\Gamma(1+\beta)\Gamma(n+1)}{\Gamma(n+\beta+1)}\right)^t\mathfrak{c}_n(\beta)\,\langle f,\mathcal{L}^{(\beta)}_n\rangle_{\nu_\beta}\,\mathcal{L}^{(\beta)}_n\end{aligned}$$

where we used (60) in the second equality and to conclude we combined the identity (47), the Stirling formula that yields that for n large enough

$$\mathfrak{c}_n(\beta)=\frac{\Gamma(n+1)\Gamma(\beta+1)}{\Gamma(n+\beta+1)}\sim\Gamma(\beta+1)n^{-\beta} \tag{66}$$

with the fact that $\widetilde{P}^{(\beta)}_t$ is closed as an Hilbert-Schmidt operator. Next, using the interweaving relation described in Proposition 26 combined with Theorem 3 since $\tau^{(\beta)}$ is infinitely divisible, we get that for any $\beta>\overline{\Pi}+m$, $P^{\tau^{(\beta)}}\overset{1}{\leftsquigarrow\!\!\rightsquigarrow}\widetilde{P}^{\tau^{(\beta)}}$. From this relation, we deduce that, for any $f\in\mathbf{L}^2(\nu)$ and $t>0$,

$$\begin{aligned}P^{\tau^{(\beta)}}_{t+1}[f] &= P^{\tau^{(\beta)}}_t\Lambda_\beta\mathrm{V}_\beta[f]=\Lambda_\beta\widetilde{P}^{\tau^{(\beta)}}_t\mathrm{V}_\beta[f] &(67)\\ &= \Lambda_\beta\sum_{n=0}^\infty\left(\frac{\Gamma(1+\beta)\Gamma(n+1)}{\Gamma(n+\beta+1)}\right)^t\mathfrak{c}_n(\beta)\,\langle \mathrm{V}_\beta[f],\mathcal{L}^{(\beta)}_n\rangle_{\nu_\beta}\,\mathcal{L}^{(\beta)}_n &(68)\\ &= \sum_{n=0}^\infty\mathfrak{c}^t_n(\beta)\mathfrak{c}_n(\beta)\,\langle f,\mathcal{V}_n\rangle_\nu\,\mathcal{P}_n &(69)\end{aligned}$$

where $\mathcal{V}_n=\mathrm{V}^*_\beta\mathcal{L}^{(\beta)}_n$ and $\Lambda_\beta\mathcal{L}^{(\beta)}_n=\mathrm{I}_\phi\mathrm{B}^*_\beta\mathcal{L}^{(\beta)}_n=\mathrm{I}_\phi\mathcal{L}_n=\mathcal{P}_n(x)$, which completes the proof of the spectral expansion of $P^{\tau^{(\beta)}}_t f$ for $t>1$. The last claim follows from the interweaving relation with warm-up time 1 and an application of Theorem 28 below by choosing $\varphi(x)=x^2-1$

4 Proofs of the main results

In the following subsections, we prove the main results about interweaving relations announced in the introduction.

4.1 Proof of the results from section 1.1

4.1.1 Proof of Theorem 3

Here we consider warm-up distributions which are infinitely divisible distributions and we construct via subordination other interweaved Markov semigroups which brought us back to the situation of deterministic warm-up times, thus showing Theorem 3. More precisely, assume that $\boldsymbol{\tau}$ is infinitely divisible. Then there exists a unique convolution semigroup on $\mathbb{R}^+$ which determines the transition kernel of the subordinator $(\tau_t)_{t\geq 0}$ where $\tau_1 \overset{(d)}{=} \boldsymbol{\tau}$. Given a Markov semigroup P, define the family of Markov operators $Q \coloneqq (Q_t)_{t\geq 0}$ via

$$\forall\, t \geq 0, \qquad Q_t \ \coloneqq\ P_{\tau_t} = \int_{\mathbb{R}_+} P_s\, \tau_t(ds)$$

Q is the subordination of P in the sense of Bochner and it is also a Markov semigroup, see e.g. [46, Chap. 12]. Similarly, given another Markov semigroup $\widetilde{P}$, define the Markov semigroup $\widetilde{Q} \coloneqq (\widetilde{Q}_t)_{t\geq 0} \coloneqq (\widetilde{P}_{\tau_t})_{t\geq 0}$.

As in Theorem 3, assume an interweaving relation holds between the semigroups P and $\widetilde{P}$ with warm-up distribution $\boldsymbol{\tau}$, that is $P \overset{\boldsymbol{\tau}}{\hookleftarrow} \widetilde{P}$. Denote by Λ and $\widetilde{\Lambda}$ the corresponding Markov kernels between the underlying state spaces V and $\widetilde{V}$ Then Figure 1 leads to the following diagram for all $t \geq 0$.

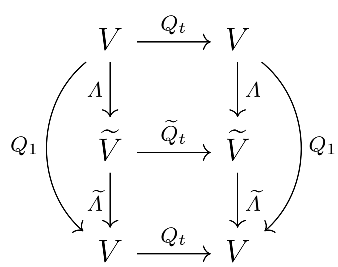

Figure 7: Intertwining relations for Q and $\widetilde{Q}$

Indeed, by definition, we have $\Lambda\widetilde{\Lambda} = P_{\boldsymbol{\tau}} = Q_1$ and for any $t \geq 0$, we get

$$\begin{aligned} Q_t\Lambda &= \int_{\mathbb{R}_+} P_s\, \tau_t(ds)\Lambda \\ &= \int_{\mathbb{R}_+} P_s\Lambda\, \tau_t(ds) \\ &= \int_{\mathbb{R}_+} \Lambda\widetilde{P}_s\, \tau_t(ds) \\ &= \Lambda\widetilde{Q}_t \end{aligned}$$

Similarly, we have

$$\forall\, t \geq 0, \qquad \widetilde{Q}_t\widetilde{\Lambda} \ =\ \widetilde{\Lambda}Q_t$$

and this ends the proof of Theorem 3.

4.1.2 Proof of Theorem 5

The first claim is obvious. Next, if $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ with $P \overset{\Lambda}{\curvearrowright} \widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$, then, clearly $\widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P \overset{\Lambda}{\curvearrowright} \widetilde{P}$. Moreover, since Markovian intertwining relationship is stable by mixture with a positive measure, we get that that $P^\tau \overset{\Lambda}{\curvearrowright} Q^\tau$ and as $P^\tau = \Lambda\widetilde{\Lambda}$, we get

$$\Lambda\,(\mathrm{VI} - Q^\tau) = 0$$

which concludes the proof of (ii) by an injectivity argument. Next, if $P \overset{\Lambda}{\curvearrowright} \widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$ and $\widetilde{P} \overset{\mathbb{V}}{\curvearrowright} \overline{P} \overset{\overline{\Lambda}}{\curvearrowright} \widetilde{P}$ then $P \overset{\Lambda\mathbb{V}}{\curvearrowright} \overline{P} \overset{\overline{\Lambda}\widetilde{\Lambda}}{\curvearrowright} P$. Moreover, we have

$$\Lambda\mathbb{V}\,\overline{\Lambda}\widetilde{\Lambda} = \Lambda Q^{\overline{\tau}}\widetilde{\Lambda} = \Lambda\widetilde{\Lambda}P^{\overline{\tau}} = P^{\tau}P^{\overline{\tau}} = F(L)\overline{F}(L)$$

where we used successively that $\widetilde{P} \overset{\overline{\tau}}{\hookleftarrow} \overline{P}$, $\widetilde{P}^{\tau} \overset{\widetilde{\Lambda}}{\curvearrowright} P^{\tau}$ which itself follows as above from $\widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$, $P \overset{\tau}{\hookleftarrow} \widetilde{P}$ and the last identity sets a notation. To complete the proof we observe that the product $F\overline{F}$ is the Laplace transform of the sum of the independent random variables $\tau + \overline{\tau}$.

4.1.3 Proof of Theorem 7

First, by since $P \overset{\Lambda}{\curvearrowright} \widetilde{P} \overset{\widetilde{\Lambda}}{\curvearrowright} P$ and $P \overset{V}{\curvearrowright} P^V$ and $P^V \overset{V^{-1}}{\curvearrowright} P$, we easily deduce that $P^V \overset{V\Lambda}{\curvearrowright} \widetilde{P} \overset{\widetilde{\Lambda}V^{-1}}{\curvearrowright} P^V$ and we conclude the proof of the first item by observing that $V\Lambda\widetilde{\Lambda}V^{-1} = VP_{\tau}V^{-1} = P^V_{\tau}$. Next, the identities (8), (9) and the second gateway in (3) yield

$$\Lambda\mathbb{P}\mathbb{I} = P\Lambda\mathbb{I} = PI\Lambda = \mathrm{I}Q\Lambda = \mathrm{I}\Lambda\widetilde{\mathbb{P}} = \Lambda\mathbb{I}\widetilde{\mathbb{P}}$$

and the injectivity of Λ gives that $\mathbb{P} \overset{\mathbb{I}}{\curvearrowright} \widetilde{\mathbb{P}}$. On can interchange the role of $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ in the previous sequence of identities to conclude that $\mathbb{P} \overset{\mathbb{I}}{\curvearrowright} \widetilde{\mathbb{P}} \overset{\mathbb{V}}{\curvearrowright} \mathbb{P}$. Next, as above, by stability of intertwining relation by mixture, we get that $P^{\tau} \overset{\Lambda}{\curvearrowright} \mathbb{P}$ and hence $\Lambda\mathbb{P}^{\tau} = P^{\tau}\Lambda = \Lambda\widetilde{\Lambda}\Lambda = \Lambda\mathbb{I}\mathbb{V}$ which concludes the proof by invoking the injectivity of Λ.

4.2 Extensions and proofs of the results from Section 1.2

4.2.1 Proof of Theorem 8

Here we extend the statement of Theorem 8 by considering (relative) φ-entropies.

Let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function such that $\varphi(1) = 0$. The (relative) **φ-entropy** of two probability measures m and ν defined on the same state space is given by

$$\mathrm{Ent}_{\varphi}(m|\nu) \;\; := \;\; \int \varphi\left(\frac{dm}{d\nu}\right) d\nu + \left(1 - \int \frac{dm}{d\nu}\, d\nu\right) \lim_{x\to+\infty} \frac{\varphi(x)}{x}$$

where $dm/d\nu$ stands for the Radon-Nikodym density of m with respect to ν. In this definition the convention $0 \cdot \infty = 0$ is enforced, namely, when m is absolutely continuous with respect to ν, the second term vanishes. When m is not absolutely continuous with respect to ν, i.e. $\int \frac{dm}{d\nu}\, d\nu < 1$, their **$\varphi$-entropy** is $+\infty$ as soon as $\lim_{x\to+\infty} \frac{\varphi(x)}{x} = +\infty$. The case of the usual entropy $\mathrm{Ent}_{\varphi}(\cdot)$ corresponds to the particular function φ given by

$$\forall\, x \in \mathbb{R}_+, \qquad \varphi(x) \;\; := \;\; x\ln(x) - x + 1 \tag{70}$$

Recall the framework of the introduction: P and $\widetilde{P}$ are two Markov semigroups, respectively on the state spaces V and $\widetilde{V}$. Let Λ and $\widetilde{\Lambda}$ be Markov kernels from V to $\widetilde{V}$ and from $\widetilde{V}$ to V. We assume that P and $\widetilde{P}$ admit invariant probability measures ν and $\widetilde{\nu}$ and that $\nu\Lambda = \widetilde{\nu}$ and $\widetilde{\nu}\widetilde{\Lambda} = \nu$. Estimates in the φ-entropy sense on the speed of convergence to equilibrium for $\widetilde{P}$ can be transferred to P with the help of a c.m.i.r.:

Theorem 28 *Assume that there exists a interweaving relation from P to $\widetilde{P}$ with warm-up distribution τ and that*

$$\forall\, \widetilde{m}_0 \in \mathcal{P}(\widetilde{V}),\, \forall\, t \geq 0, \qquad \mathrm{Ent}_{\varphi}(\widetilde{m}_0\widetilde{P}_t|\widetilde{\nu}) \;\; \leq \;\; \varepsilon(t, \mathrm{Ent}_{\varphi}(\widetilde{m}_0|\widetilde{\nu})) \tag{71}$$

for some function $\varepsilon : \mathbb{R}_+ \times \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$, which is non-decreasing with respect to the second variable. Then we have

$$\forall\ m_0 \in \mathcal{P}(V),\ \forall\ t \geq 0, \qquad \mathrm{Ent}_\varphi(m_0 P_{\theta_t(\tau)}|\nu) \ \leq\ \varepsilon(t, \mathrm{Ent}_\varphi(m_0|\nu)) \tag{72}$$

where θ_t is the translation operator on $\mathbb{R}_+$.

Remark 29 *As in the introduction, for this estimate to be meaningful, one should furthermore require that*

$$\forall\ E \in \mathbb{R}_+, \qquad \lim_{t\to+\infty} \varepsilon(t, E) \ =\ 0$$

Proof of Theorem 28

Consider E and $\widetilde{E}$ two measurable spaces and Ξ a Markov kernel from E to $\widetilde{E}$. Let $\widetilde{m}$ and m be two probability measures on E. As a consequence of Jensen inequality, we have for any convex function φ as above,

$$\mathrm{Ent}_\varphi(\widetilde{m}\Xi|m\Xi) \ \leq\ \mathrm{Ent}_\varphi(\widetilde{m}|m) \tag{73}$$

(see e.g. [17]).

The interweaving relation between P and $\widetilde{P}$ implies that for any $t \geq 0$, we have

$$\Lambda\widetilde{P}_t\widetilde{\Lambda} \ =\ P_{\theta_t(\tau)}$$

It follows that for any $m_0 \in \mathcal{P}(V)$,

$$\begin{aligned} \mathrm{Ent}_\varphi(m_0 P_{\theta_t(\tau)}|\nu) \ &=\ \mathrm{Ent}_\varphi(m_0\Lambda\widetilde{P}_t\widetilde{\Lambda}|\widetilde{\nu}\widetilde{\Lambda}) \\ &\leq\ \mathrm{Ent}_\varphi(m_0\Lambda\widetilde{P}_t|\widetilde{\nu}) \end{aligned}$$

where (73) was applied with $\widetilde{m} \coloneqq m_0\Lambda\widetilde{P}_t$, $m \coloneqq \widetilde{\nu}$ and $\Xi \coloneqq \widetilde{\Lambda}$. Taking into account (71), we get

$$\begin{aligned} \mathrm{Ent}_\varphi(m_0\Lambda\widetilde{P}_t|\widetilde{\nu}) \ &\leq\ \varepsilon(t, \mathrm{Ent}_\varphi(m_0\Lambda|\widetilde{\nu})) \\ &=\ \varepsilon(t, \mathrm{Ent}_\varphi(m_0\Lambda|\nu\Lambda)) \\ &\leq\ \varepsilon(t, \mathrm{Ent}_\varphi(m_0|\nu)) \end{aligned}$$

where we used again (73) with $\widetilde{m} \coloneqq m_0$, $m \coloneqq \nu$ and $\Xi \coloneqq \Lambda$.

■

The traditional way to deduce a bound such as (71) is via φ-Sobolev inequalities. Without entering into the general theory, let us e.g. consider the case where $\widetilde{V}$ is a finite state space and $\widetilde{P}$ is generated by an irreducible Markov generator $\widetilde{L}$. Denote $\widetilde{\mathcal{A}}$ the set of positive functions defined on $\widetilde{V}$ with $\widetilde{\nu}[f] = 1$ and assume that φ is differentiable on $(0, +\infty)$ (in particular $\varphi'(1) = 0$). Consider the energy

$$\forall\ f \in \widetilde{\mathcal{A}}, \qquad \widetilde{\mathcal{E}}_\varphi(f, \varphi'(f)) \ \coloneqq\ -\widetilde{\nu}[f\widetilde{L}[\varphi'(f)]]$$

(the r.h.s. is always non-negative) and denote

$$\widetilde{\alpha}_\varphi \ \coloneqq\ \inf_{f\in\widetilde{\mathcal{A}}\setminus\{\widetilde{\mathbb{1}}\}} \frac{\widetilde{\mathcal{E}}_\varphi(f, \varphi'(f))}{\mathrm{Ent}_\varphi(f\cdot\widetilde{\nu}|\widetilde{\nu})}$$

where $\widetilde{\mathbb{1}}$ is the function only taking the value 1 on $\widetilde{V}$ and $f\cdot\widetilde{\nu}$ is the probability on $\widetilde{V}$ admitting the density f w.r.t. $\widetilde{\nu}$. The quantity $\widetilde{\alpha}_\varphi$ is non-negative and is called the φ-Sobolev constant. Then (71) holds with the function ε given by

$$\forall\, t\geq 0,\, \forall\, E\geq 0, \qquad \varepsilon(t,E) \;\coloneqq\; \exp(-\widetilde{\alpha}_\varphi t)E$$

This result is obtained by differentiating the quantity $\mathrm{Ent}_\varphi(\widetilde{m}_0\widetilde{P}_t|\widetilde{\nu})$ with respect to $t>0$, for any fixed $\widetilde{m}_0\in\mathcal{P}(\widetilde{V})$, and by applying Grönwall lemma. The validity of this approach is very general, up to the appropriate definition of the domain $\widetilde{\mathcal{A}}$.

In the classical case (70) and when the finite generator $\widetilde{L}$ is assumed to be furthermore reversible, the energy is given by

$$\forall\, f\in\widetilde{\mathcal{A}}, \qquad \widetilde{\mathcal{E}}(f,\ln(f)) \;=\; \frac{1}{2}\sum_{x,y\in\widetilde{V}}(f(y)-f(x))(\ln(f(y))-\ln(f(x)))\,\widetilde{\nu}(x)\widetilde{L}(x,y)$$

and the corresponding constant $\widetilde{\alpha}$ is called the modified logarithmic Sobolev constant. It is bounded below by the usual logarithmic Sobolev constant, obtained by replacing $\widetilde{\mathcal{E}}(f,\ln(f))$ by

$$4\widetilde{\mathcal{E}}(\sqrt{f},\sqrt{f}) \;=\; 2\sum_{x,y\in\widetilde{V}}(\sqrt{f}(y)-\sqrt{f}(x))^2\,\widetilde{\nu}(x)\widetilde{L}(x,y)$$

in the above definitions. In the diffusion framework, the modified and usual logarithmic Sobolev constant coincide (for the previous functional analysis assertions, see for instance the book of Ané et al. [4]).

Let us consider the situation of a deterministic warm-up time: there exists $t_0\geq 0$ such that $\tau=\delta_{t_0}$, as in Section 2. Assume that $(\widetilde{P},\widetilde{\nu})$ satisfies a modified logarithmic Sobolev inequality with constant $\widetilde{\alpha}>0$, so that for any initial distribution $\widetilde{m}_0\in\mathcal{P}(\widetilde{V})$, we have

$$\forall\, t\geq 0, \qquad \mathrm{Ent}(\widetilde{m}_t|\widetilde{\nu}) \;\leq\; \exp(-\widetilde{\alpha}t)\mathrm{Ent}(\widetilde{m}_0|\widetilde{\nu})$$

Theorem 28 enables to get for (P,ν) that for any initial distribution $m_0\in\mathcal{P}(V)$, we have

$$\forall\, t\geq 0, \qquad \mathrm{Ent}_\varphi(m_{t_0+t}|\nu) \;\leq\; \exp(-\widetilde{\alpha}t)\mathrm{Ent}_\varphi(m_0|\nu)$$

Alternatively, taking into account that the relative entropy of the time marginal laws of a Markov process with respect to its invariant measure is always non-increasing with respect to time (see e.g. [17]), we get

$$\forall\, t\geq 0, \qquad \mathrm{Ent}_\varphi(m_t|\nu) \;\leq\; \exp(-\widetilde{\alpha}(t-t_0)_+)\mathrm{Ent}_\varphi(m_0|\nu) \tag{74}$$

In this bound, the time t_0 clearly appears as a warm-up period. The fact that no contractive estimate of $\mathrm{Ent}_\varphi(m_t|\nu)$ can be deduced for $t\in[0,t_0]$ relates (74) to hypocoercive bounds (see e.g. Villani [48]).

These considerations were illustrated by the classical and discrete examples of Subsection 2.3. In Subsection 3.3, we presented a interweaving relation with a random warm-up time between jump Laguerre processes and classical Laguerre processes. It enables to get estimates on convergence to equilibrium in entropy sense for non-reversible jump processes without the a priori knowledge of corresponding modified logarithmic Sobolev inequalities. It shows the applicative potential of c.m.i.r.

Remark 30 In general, it is not possible to deduce from a bound such as (72) an estimate on $\mathrm{Ent}_\varphi(m_0P_t|\nu)$ for given large $t\geq 0$, except in the case of a deterministic warm-up time. Indeed, consider P the deterministic semigroup generated on the circle $\mathbb{T} \coloneqq \mathbb{R}/(2\tau\mathbb{Z})$ by the usual derivation ∂. Starting from $x_0 \in \mathbb{T}$, the position at time $t \geq 0$ of an associated Markov process is $x_0 + t\ [2\tau]$. The associated invariant measure ν is the uniform distribution over $\mathbb{T}$. Let τ be the uniform distribution over $[0, 2\tau]$. For any $t \geq 0$, we have $\mathrm{Ent}_\varphi(m_0 P_{\theta_t(\tau)}|\nu) = 0$ for any initial distribution m_0, while $\mathrm{Ent}_\varphi(m_0 P_t|\nu) = +\infty$ when m_0 is a Dirac mass.

Remark 31 Another approach to convergence to equilibrium is based on strong stationary times, see Aldous and Diaconis [2] and Diaconis and Fill [19] for seminal works about this alternative point of view. It is more probabilistic in spirit, since it constructs stopping times τ such that the position of the underlying Markov process is at equilibrium and independent from τ. Furthermore, it is an important motivation for the investigation of intertwining relations. Thus it is natural to wonder if interweaving relations enable the transfer of strong stationary times. Unfortunately we did not find a satisfactory procedure, especially when the warm-up distribution is not a Dirac mass. Nevertheless, strong stationary times are often used due to their close relation to the convergence to equilibrium in the separation sense (see e.g. Diaconis and Fill [19]), and interweaving relations enable to directly transfer corresponding estimates.

Recall that the separation discrepancy $\mathfrak{s}(m, \nu)$ between two probability measures m and ν on the same state space is defined as

$$\mathfrak{s}(m,\nu) \quad \coloneqq \quad \operatorname*{ess\,sup}_{\nu} 1 - \frac{dm}{d\nu}$$

The separation discrepancy is in fact a limit case of φ-entropies. More precisely, for $p \geq 1$, consider the convex mapping

$$\forall\, x \in \mathbb{R}_+, \qquad \varphi_p(x) \quad \coloneqq \quad (1-x)_+^p$$

where $(\cdot)_+$ stands for the non-negative part. It is not difficult to show that for any probability measures m and ν on the same state space, we have

$$\lim_{p\to+\infty} \left(\mathrm{Ent}_{\varphi_p}(m,\nu)\right)^{1/p} \quad = \quad \mathfrak{s}(m,\nu)$$

This result in conjunction with Theorem 8 show that we can transfer separation estimates through c.m.i.r. More precisely, assume that we have a interweaving relation with warm-up distribution τ between the ergodic semigroups P and $\widetilde{P}$, with invariant probability ν and $\widetilde{\nu}$. Let m_0 be an initial distribution on V and denote $\widetilde{m}_0 \coloneqq m_0\Lambda$. Assume that we have a function $\widetilde{\varepsilon} : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\forall\, t \geq 0, \qquad \mathfrak{s}(\widetilde{m}_0\widetilde{P}_t, \widetilde{\nu}) \quad \leq \quad \widetilde{\varepsilon}(t)$$

Since we have for any $p \geq 1$ and any probability measure $\widetilde{m}$ on $\widetilde{V}$,

$$\mathrm{Ent}_{\varphi_p}(\widetilde{m}, \widetilde{\nu}) \quad \leq \quad \mathfrak{s}(\widetilde{m}, \widetilde{\nu})^p$$

Theorem 8 implies that

$$\forall\, p \geq 1, \forall\, t \geq 0, \qquad \mathrm{Ent}_{\varphi_p}(m_0 P_{\theta_t(\tau)}, \nu) \quad \leq \quad \widetilde{\varepsilon}(t)^p$$

It remains to take the power $1/p$ and to let p go to infinity to get

$$\forall\, t \geq 0, \qquad \mathfrak{s}(m_0 P_{\theta_t(\tau)}, \nu) \quad \leq \quad \widetilde{\varepsilon}(t)$$

which corresponds to the wanted separation estimate transfer.

4.3 Hyperboundedness

As in the previous subsection, the underlying principle for the transfer of hyperboundedness via interweaving relations is convexity, so that the Orlicz spaces are the natural framework here, not only the $\mathbf{L}^p$ spaces, for $p \geq 2$, as stated in Theorem 9.

Let us recall the notion of Orlicz spaces (for a general introduction, see for instance the book of Rao and Ren [43]). Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be a Young function: it is a even convex function $\varphi \neq 0$ satisfying $\varphi(0) = 0$. When E is a measurable space endowed with a probability measure m, the **Orlicz space** $\mathbf{L}^\varphi(m)$ is the vector space of measurable functions $f : E \to \mathbb{R}$ such that

$$\|f\|_{\mathbf{L}^\varphi(m)} \;\; \coloneqq \;\; \inf\{r > 0 : \int \varphi(f/r)\, dm \leq 1\}$$

is finite. The quantity $\|\cdot\|_{\mathbf{L}^\varphi(m)}$ defines a norm on $\mathbf{L}^\varphi(m)$, when the functions are identified up to a m-negligible set. The key property of Orlicz spaces we will need is:

Lemma 32 *Consider Λ a Markov kernel from E to another measurable space $\widetilde{E}$. Let $\widetilde{m}$ be the image of the probability measure m on E by Λ. For any measurable function $f : \widetilde{E} \to \mathbb{R}$, we have*

$$\|\Lambda[f]\|_{\mathbf{L}^\varphi(m)} \;\; \leq \;\; \|f\|_{\mathbf{L}^\varphi(\widetilde{m})}$$

Proof: This is an immediate consequence of convexity. Indeed, by Jensen's inequality, we have m-a.s. and for any $r \geq 0$,

$$\varphi(\Lambda[f/r]) \;\; \leq \;\; \Lambda[\varphi(f/r)]$$

Integrating with respect to m, we get

$$\begin{aligned} \int \varphi(\Lambda[f/r])\, dm \;\; &\leq \;\; \int \Lambda[\varphi(f/r)]\, dm \\ &= \;\; \int \varphi(f/r)\, d\widetilde{m} \end{aligned}$$

and it remains to take the infimum of the $r > 0$ such that $\int \varphi(f/r)\, d\widetilde{m} \leq 1$ to get the announced result.

■

As in the introduction, let be given P a Markov semigroup from V to V and $\widetilde{P}$ a Markov semigroup from $\widetilde{V}$ to $\widetilde{V}$. Assume that ν and $\widetilde{\nu}$ are respectively invariant probability measures for P and $\widetilde{P}$ and that an interweaving relation holds, as described in Figure 1, with Markov kernels Λ from V to $\widetilde{V}$ and $\widetilde{\Lambda}$ from $\widetilde{V}$ to V, as well as warm-up distribution τ. As usual, $\nu\Lambda$ and $\widetilde{\nu}\widetilde{\Lambda}$ are respectively invariant for $\widetilde{P}$ and P. In case of non-uniqueness of these invariant probability measures, we furthermore assume that $\widetilde{\nu} = \nu\Lambda$ and $\nu = \widetilde{\nu}\widetilde{\Lambda}$. Here is an extension of Theorem 9:

Theorem 33 *Assume that for some time $T \geq 0$ and some Young function φ, we have in the operator norm*

$$\|\widetilde{P}_T\|_{\mathbf{L}^2(\widetilde{\nu}) \to \mathbf{L}^\varphi(\widetilde{\nu})} \;\; \leq \;\; 1 \tag{75}$$

Then we get

$$\|P_{T+\tau}\|_{\mathbf{L}^2(\nu) \to \mathbf{L}^\varphi(\nu)} \;\; \leq \;\; 1 \tag{76}$$

Proof: As in the proof of Theorem 28, the starting point is

$$\Lambda\widetilde{P}_T\widetilde{\Lambda} \;=\; P_{T+\tau}$$

It follows that

$$\begin{aligned}\|P_{T+\tau}\|_{\mathbf{L}^2(\nu)\to\mathbf{L}^\varphi(\nu)} &\leq \|\Lambda\|_{\mathbf{L}^\varphi(\widetilde{\nu})\to\mathbf{L}^\varphi(\nu)}\|\widetilde{P}_T\|_{\mathbf{L}^2(\widetilde{\nu})\to\mathbf{L}^\varphi(\widetilde{\nu})}\|\widetilde{\Lambda}\|_{\mathbf{L}^2(\nu)\to\mathbf{L}^2(\widetilde{\nu})}\\ &\leq \|\Lambda\|_{\mathbf{L}^\varphi(\widetilde{\nu})\to\mathbf{L}^\varphi(\nu)}\|\widetilde{\Lambda}\|_{\mathbf{L}^2(\nu)\to\mathbf{L}^2(\widetilde{\nu})}\end{aligned}$$

Lemma 32 applied with $m=\nu$ and $\widetilde{m}=\widetilde{\nu}$ (recall that $\nu\Lambda=\widetilde{\nu}$) implies that

$$\|\Lambda\|_{\mathbf{L}^\varphi(\widetilde{\nu})\to\mathbf{L}^\varphi(\nu)} \;=\; 1$$

Considering the Young function $\mathbb{R}\ni x\mapsto x^2$, Lemma 32 applied with $m=\widetilde{\nu}$ and $m=\nu$ (recall that $\widetilde{\nu}\widetilde{\Lambda}=\nu$) implies that

$$\|\widetilde{\Lambda}\|_{\mathbf{L}^2(\nu)\to\mathbf{L}^2(\widetilde{\nu})} \;=\; 1$$

concluding the proof of the wanted bound.

■

Theorem 9 is a consequence of Theorem 28, applied, for fixed $t\geq 0$, with $T=t$ and

$$\varphi\,:\,\mathbb{R}\ni x \;\mapsto\; x^{p(\widetilde{\alpha}t)}$$

Note that due to the warm-up distribution, it is not possible to deduce from the conclusion of Theorem 9 that the semigroup P satisfies a logarithmic Sobolev inequality (for the classical links between the latter inequality and hypercontractivity, again see e.g. Ané et al. [4]).

4.4 Proof of Theorem 10

Assume first that a cut-off phenomenon occurs for the family $(P^{(n)})_{n\in\mathbb{Z}_+}$, with cut-off times $(t^{(n)})_{n\in\mathbb{Z}_+}$, and let us show the same is true for $(\widetilde{P}^{(n)})_{n\in\mathbb{Z}_+}$.

Consider the Young function $\mathbb{R}\ni x\mapsto|x-1|$. The associated entropy between the probability measures m and ν is just twice the total variation

$$2\,\|m-\nu\|_{\mathrm{tv}} \;=\; \int\left|\frac{dm}{d\nu}-1\right|d\nu+1-\int\frac{dm}{d\nu}\,d\nu$$

The proof of Theorem 28 with this particular Young function shows that for any $n\in\mathbb{Z}_+$,

$$\forall\, t\geq 0,\,\forall\,\widetilde{m}_0\in\mathcal{P}(\widetilde{V}^{(n)}),\qquad \begin{aligned}\left\|\widetilde{m}_0\widetilde{P}^{(n)}_{t_0^{(n)}+t}-\widetilde{\nu}^{(n)}\right\|_{\mathrm{tv}} &\leq \left\|\widetilde{m}_0\widetilde{\Lambda}P^{(n)}_t-\nu^{(n)}\right\|_{\mathrm{tv}}\\ &\leq \mathfrak{d}^{(n)}(t)\end{aligned}$$

where $\mathfrak{d}^{(n)}$ is given in (14). Considering a similar definition of $\widetilde{\mathfrak{d}}^{(n)}$ for the semigroup $\widetilde{P}^{(n)}$, we obtain

$$\forall\, t\geq 0,\qquad \widetilde{\mathfrak{d}}^{(n)}(t_0^{(n)}+t) \;\leq\; \mathfrak{d}^{(n)}(t)$$

Taking into account that for any $n \in \mathbb{Z}_+$, the function $\widetilde{\mathfrak{d}}^{(n)}$ is non-increasing, we deduce from the cut-off phenomenon for P and from (15) that for any $r > 0$,

$$\begin{aligned}
\overline{\lim_{n\to\infty}}\, \widetilde{\mathfrak{d}}^{(n)}((1+r)t^{(n)}) &\leq \overline{\lim_{n\to\infty}}\, \widetilde{\mathfrak{d}}^{(n)}(t_0^{(n)} + (1+r/2)t^{(n)}) \\
&\leq \lim_{n\to\infty} \mathfrak{d}^{(n)}((1+r/2)t^{(n)}) \\
&= 0
\end{aligned}$$

For the other point in the definition of the cut-off phenomenon, assume by contradiction that for some $r_0 \in (0,1)$, we have

$$\underline{\lim}_{n\to\infty}\, \widetilde{\mathfrak{d}}^{(n)}((1-r_0)t^{(n)}) \quad < \quad 1 \tag{77}$$

By the assumed symmetry of the interweaving relations between the sequence $(P^{(n)})_{n\in\mathbb{Z}_+}$ and $(\widetilde{P}^{(n)})_{n\in\mathbb{Z}_+}$, we show as above that

$$\forall\, t \geq 0, \qquad \mathfrak{d}^{(n)}(t_0^{(n)} + t) \quad \leq \quad \widetilde{\mathfrak{d}}^{(n)}(t)$$

Taking into account that for any $n \in \mathbb{Z}_+$, the function $\mathfrak{d}^{(n)}$ is non-increasing, we deduce from (77) and from (15) that

$$\begin{aligned}
\lim_{n\to\infty} \mathfrak{d}^{(n)}((1-r_0/2)t^{(n)}) &\leq \underline{\lim}_{n\to\infty}\, \mathfrak{d}^{(n)}(t_0^{(n)} + (1-r_0)t^{(n)}) \\
&\leq \underline{\lim}_{n\to\infty}\, \widetilde{\mathfrak{d}}^{(n)}((1-r_0)t^{(n)}) \\
&< 1
\end{aligned}$$

which is in contradiction with the cut-off phenomenon for the family $(P^{(n)})_{n\in\mathbb{Z}_+}$. Thus we get that for any $r \in (0,1)$,

$$\lim_{n\to\infty} \widetilde{\mathfrak{d}}^{(n)}((1-r)t^{(n)}) \quad = \quad 1$$

and this ends the proof that a cut-off phenomenon occurs for the family $(P^{(n)})_{n\in\mathbb{Z}_+}$ with cut-off times $(t^{(n)})_{n\in\mathbb{Z}_+}$.

The remaining claims of Theorem 10 are proven by a similar line of reasoning.

References

[1] N. I. Akhiezer and I. M. Glazman. *Theory of linear operators in Hilbert space. Vol. II*, volume 10 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1981. Translated from the third Russian edition by E. R. Dawson, Translation edited by W. N. Everitt.

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

[3] Alexandru Aleman and Joe Viola. Singular-value decomposition of solution operators to model evolution equations. *Int. Math. Res. Not. IMRN*, (17):8275–8288, 2015.

[4] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[5] A. Arnold, A. Einav, and T. Wöhrer. On the rates of decay to equilibrium in degenerate and defective Fokker-Planck equations. *J. Differential Equations*, 264(11):6843–6872, 2018.

[6] A. Arnold and J. Erb. Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift. *arXiv e-prints*, page arXiv:1409.5425, Sep 2014.

[7] Theodoros Assiotis. On a gateway between the Laguerre process and dynamics on partitions arXiv:1903.01265 [math.PR], 2019.

[8] Dominique Bakry. Remarques sur les semigroupes de Jacobi. *Astérisque*, (236):23–39, 1996. Hommage à P. A. Meyer et J. Neveu.

[9] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[10] G. Barrera. *Abrupt convergence for a family of Ornstein-Uhlenbeck processes*, Braz. J. Probab. Stat., 32(1):188–199, 2018.

[11] Javiera Barrera, Béatrice Lachaud, and Bernard Ycart. Cut-off for n-tuples of exponentially converging processes. *Stochastic Process. Appl.*, 116(10):1433–1446, 2006.

[12] V. I. Bogachev. Ornstein-Uhlenbeck operators and semigroups. *Uspekhi Mat. Nauk*, 73(2(440)):3–74, 2018.

[13] A. Borodin and G. Olshanski. Markov processes on the path space of the Gelfand-Tsetlin graph and on its boundary. Journal of Functional Analysis, Vol. 263, pp 248-303, 2012.

[14] A. Borodin and G. Olshanski. Markov dynamics on the Thoma cone: a model of time-dependent determinantal processes with infinitely many particles Electron. J. Probab. Volume 18 (2013), paper no. 75, 43 pp.

[15] Philippe Carmona, Frédérique Petit, and Marc Yor. Beta-gamma random variables and intertwining relations between certain Markov processes. *Rev. Mat. Iberoamericana*, 14(2):311–367, 1998.

[16] Patrick Cheridito, Pierre Patie, Aditya Vaidyanathan and Anna Srapionyan. On non-local ergodic Jacobi semigroups: spectral theory, convergence-to-equilibrium, and contractivity. *available at arXiv:1905.07832 [math.PR]*, 2019.

[17] Pierre Del Moral, Michel Ledoux, and Laurent Miclo. On contraction properties of Markov kernels. *Probab. Theory Related Fields*, 126(3):395–420, 2003.

[18] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, 93(4):1659–1664, 1996.

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

[20] Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[21] E.B. Dynkin, Markov Processes. Vol. I. Springer, 1965.

[22] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[23] J.A. Fill. On hitting times and fastest strong stationary times for skip-free and more general chains. J. Theor. Probab. 22, 587–600, 2009.

[24] É. Fontenas. Sur les minorations des constantes de Sobolev et de Sobolev logarithmiques pour les opérateurs de Jacobi et de Laguerre. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 14–29. Springer, Berlin, 1998.

[25] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[26] A. E. Kyprianou. *Fluctuations of Lévy processes with applications*. Universitext. Springer, Heidelberg, second edition, 2014. Introductory lectures.

[27] B. Lachaud *Cut-off and hitting times of a sample of Ornstein-Uhlenbeck processes and its average*, J. Appl. Probab., 42(4),1069–1080, 2005.

[28] G. Metafune, D. Pallara, and E. Priola. Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures. *J. Funct. Anal.*, 196(1):40–60, 2002.

[29] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields*, 5(3):319–330, 1999.

[30] Laurent Miclo. Sur l'inégalité de Sobolev logarithmique des opérateurs de Laguerre à petit paramètre. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 222–229. Springer, Berlin, 2003.

[31] Laurent Miclo. On the Markovian similarity. Preprint available at `https://hal.archives-ouvertes.fr/hal-01281029`, March 2016.

[32] Laurent Miclo and Pierre Patie. On a gateway between continuous and discrete Bessel and Laguerre processes. *Annales Henri Lebesgue*, 2, 59-98, 2019.

[33] P. Monmarché. Generalized Γ Calculus and Application to Interacting Particles on a Graph. *Potential Anal.*, 50(3):439–466, 2019.

[34] M. Ottobre, G. A. Pavliotis, and K. Pravda-Starov. Some remarks on degenerate hypoelliptic Ornstein-Uhlenbeck operators. *J. Math. Anal. Appl.*, 429(2):676–712, 2015.

[35] S. Pal and M. Shkolnikov, Intertwining diffusions and wave equations, http://arxiv.org/pdf/1306.0857v3.pdf, 2013.

[36] P. Patie and M. Savov. Spectral expansion of non-self-adjoint generalized Laguerre semigroups. *Mem. Amer. Math. Soc.,* to appear, page 179, 2018.

[37] P. Patie and M. Savov. Bernstein-gamma functions and exponential functionals of Lévy Processes. Electron. J. Probab., (23)75, 101 pp., 2018.

[38] P. Patie, M. Savov, and Y. Zhao. Intertwining, excursion theory and Krein theory of strings for non-self-adjoint Markov semigroups. Ann. Probab., to appear, 51p., 2019.

[39] P. Patie and T. Simon. Intertwining certain fractional derivatives. Potent. Anal., 36:569–587, 2012.

[40] P. Patie and A. Vaidyanathan. A spectral theoretical approach for hypocoercivity applied to some degenerate hypoelliptic, and non-local operators, *arXiv:1905.07042 [math.PR]*, 2019.

[41] P. Patie and Y. Zhao. Spectral decomposition of fractional operators and a reflected stable semigroup. J. Differential Equations, 262(3):1690–1719, 2017.

[42] J.W. Pitman and L.C.G. Rogers. Markov functions. Ann. Probab. 9(4), 573–582, 1981.

[43] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[44] L. Saloff-Coste. Precise estimates on the rate at which certain diffusions tend to equilibrium *Math. Z.*, 217(4), 641–677, 1994.

[45] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999

[46] R. L. Schilling, R. Song, and Z. Vondraček. *Bernstein functions*, volume 37 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2012. Theory and applications.

[47] Gábor Szegö. *Orthogonal polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[48] Cédric Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.

[49] Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.

[50] Kōsaku Yosida. *Functional analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

miclo@math.cnrs.fr

Toulouse School of Economics,
Manufacture des Tabacs, 21, Allée de Brienne
31015 Toulouse cedex 6, France

Institut de Mathématiques de Toulouse
Université Paul Sabatier, 118, route de Narbonne
31062 Toulouse cedex 9, France
miclo@math.univ-toulouse.fr

‡ pp396@cornell.edu

School of Operations Research and Information Engineering
Cornell University
Ithaca, NY 14853
USA