A Dobrushin uniqueness criterion for interacting markovian systems with synchronous updating

Damien THOMINE

January 30, 2012

Abstract

We find, using rather classical spaces of Lipschitz or Hölder functions, a bound on the spectral radius of the transfer operator associated with Markov chains on countably many interacting sites with synchronous updates. This gives us a Dobrushin-like criterion that guarantees the uniqueness of a stationary measure, but also the exponential temporal decay of correlations for nice observables and, under some additional assumptions, a spatial decay of correlations. Then, the stability of the spectral radius is proved for suitable discrete systems or interacting particle systems.

More than forty years ago, R.L. Dobrushin discovered a criterion which ensure that many models of statistical physics have at most one equilibrium measure in the high temperature domain [3]. It says, roughly, that if the state of the system on any given site does not depend too much on the state of the other sites, then the equilibrium measure, provided it exists, is unique. This criterion is remarkably elegant, in that it does not require any translation-invariance or finite range assumption, it is easily computable, and can give rather sharp upper bound on the critical temperature when the system exhibits a phase transition.

In this article, we focus on a dynamical version of this problem: instead of working with Gibbs measures whose specifications are given, we look at a stochastic cellular automaton (i.e. a Markov chain on the space of configurations). Starting from a given configuration, the updating process change all the sites at the same time, independently. The techniques involved will not change much, as we shall basically estimate the norm or the spectral radius of a given transfer operator, as in Dobrushin's original paper [3] or many following works [9], but will nevertheless add a time perspective, introduce some new objects such as Markov chains, as well as change some more minor matters (for example, the diagonal dependence coefficients c_{ii}, where i is a site, may not be null). This time-dependent perspective is already present in an article by Mac Kay [10], but without explicit reference to the Dobrushin coefficients. We also prove that systems on a discrete state space or iterated function systems are stable, under some additional assumptions and provided the perturbations are nice enough.

The criterions we get not only ensure the uniqueness of the stationary measure, but can also be used with great effect to prove an exponential temporal decay of correlation (Corollary 1.15) or, with additional assumptions, a spatial decay of correlations (Theorem 4.4). One could continue along this way to prove, for instance, a central limit theorem, with perturbations of the transfer operator or directly with the exponential decay of correlations.

As the last modifications were done to this article, a recent paper of Mac Kay [10] was brought to my knowledge. It goes along the same line as the article you are now reading, and one of its arguments (the Theorem 1) can doubtlessly be used in our setting to great effect. I chose not to modify the article, since it would need a very thorough overhaul to handle those modifications, but added some comments to point out what one could gain this way (in short : better and more explicit bounds, and a little less computation).

In Section 1 we present the setting we chose, the tools we shall work with, and the equivalent of Dobrushin uniqueness criterion (which is shown to be a upper bound on the norm of the transfer operator). The tools come from the basic arsenal of optimal transport theory: minimizing couplings, Wasserstein distances, and Kantorovich-Rubinstein Theorem. We hope it provides the reader with a simple and almost optimal setting in which Dobrushin uniqueness criterion (Theorem 1.12) is valid - and actually appears with little effort - in Subsection 1.2. In particular, it leaves a lot of freedom in the way one chooses the state spaces, which are not required to be identical, and the distances, which are only required to be lower semi-continuous (and actually only measurable if one looks for some weaker results). Although the techniques involved are in nothing new (but seem to be frequently re-discovered), the setting is slightly different, and in some ways more complete. We also obtain the temporal decay of correlations for free (Corollary 1.15).

The Section 2 presents much more involved arguments. When one looks at dynamics where all the sites behave independently, they can find trivial examples where Dobrushin criterion fails, but where the transfer operator has exactly one eigenvalue of modulus 1: a product of mixing Markov chains with a convenient transition matrix is enough. In the independent case, one can get past this problem since the transfer operator can be trivially iterated. However, we want to know what happens when one perturbs such a system: is there still at most one equilibrium measure? We answer this question for some systems and nice perturbations, but only for discrete state space in Theorem 3.1 or iterated function systems (for which we show a Lasota-Yorke inequality - Theorem 3.5) in Theorem 3.6. This section concentrates most of the new results of this article.

Section 4 is devoted to the proof of a spatial decay of correlations for a unique equilibrium measure, much in the same spirit as in [6], but intrinsically limited by the fact that we do not work with lattices, and that we do not require the sites to be identical, so that it makes no sense speaking of the translate of a function. In exchange, the main result we get - Theorem 4.4 - is much more general.

Section 5 presents some variations on the Ising model, as examples on which our diverse results apply.

1 Setting and tools

We give ourselves a countable index set V, typically infinite (our results are much less interesting if one works with a finite index set). For some applications, it may be useful to see V either as the vertices of a graph, or as the set of nonnegative integers $\mathbb{N}$. Each element of V will be called a site. For each i in V, let E_i be a Polish space, and d_i a lower semi-continuous distance on E_i. We put $\Omega := \prod_{i \in V} E_i$, and we endow it with the product topology and its Borel sets. Any element of Ω will be called a configuration, and by induction a point of E_i may sometimes be called a configuration (or a state) at site i. If $\omega \in \Omega$ is a configuration, we denote by $\omega^i \in E_i$ its value at site i, and by $\overline{\omega}^i \in \prod_{\substack{j \in V \\ j \neq i}} E_j = \overline{\Omega}^i$ the configuration induced outside of site i. Sometimes, we also use notations such as $\omega^{>n}$, which denote a configuration on all sites whose label is strictly greater than n (when the index set is $\mathbb{N}$), and other similar objects.

The advantage of using lower-semi continuous distances on Polish spaces is that it encompasses both the "Polish space with discrete distance" (i.e. total variation norm on measures) and the "compact metric" settings.

We assume that the distances d_i are uniformly bounded, or that in other words there exists a $\Delta > 0$ such that, for all i in V, we have $d_i \leq \Delta$.

Definition 1.1 (Uniform continuity at infinity).
We say that a function f from Ω to a metric space (M, D) is uniformly continuous at infinity (or shorter is UCAI) if, for every $\varepsilon > 0$, there exists an integer N such that, for every couple of configurations ω_1 and ω_2 which are identical on the first N sites, $D(f(\omega_1), f(\omega_2)) < \varepsilon$.

For each i in V, let $\omega \mapsto p_{i,\omega}$ be a measurable application from Ω to $\mathcal{P}(E_i)$. We define a Markov chain on Ω in the following way. First, we give ourselves an initial distribution over Ω; then, if $(X_n^i)_{i\in V}$ is the configuration at time n, the law of $(X_{n+1}^i)_{i\in V}$ knowing $(X_n^i)_{i\in V}$ is $\bigotimes_{i\in V} p_{i,(X_n^j)_{j\in V}}$. In other words, we choose the new state of all sites independently, but with a law which depends on the previous state of all sites.

We define a transfer operator $\mathcal{L}^*$ on the space of probability measures (and also on the space of sign measures) by:

$$\mathcal{L}^*\mu := \int_\Omega \bigotimes_{i\in V} p_{i,\omega}\mu(\,\mathrm{d}\omega).$$

Equivalently, if μ is the law of $(X_0^i)_{i\in V}$, then $\mathcal{L}^*\mu$ is the law of $(X_1^i)_{i\in V}$. We consider its dual operator, which acts naturally on the space on continuous bounded functions. For any such function, we have:

$$\mathcal{L}f(\omega) := \int_\Omega f(\omega') \bigotimes_{i\in V} p_{i,\omega}(\,\mathrm{d}\omega').$$

1.1 Basic tools

We use in this article an approach oriented towards functional analysis. One of the main point is a duality between Lipschitz functions and some finite measures, coming from Kantorovich-Rubinstein Theorem. We need to define all the functions and measures space we work with. Since the norm we shall use on measures spaces do not look very natural (at least at first sight) but are easily explained by duality, we shall start by the functions spaces. Naturally, they are spaces of functions which are Lipschitz in some sense; they were used in [9].

Definition 1.2 (Hölder functions).
Let i in V, and α in $[0, 1]$. The local Hölder spaces of index α are defined by:

$$\widetilde{\mathrm{Lip}}_\alpha(E_i) := \left\{ f : E_i \to \mathbb{R}, \text{measurable}, \sup_{\omega_a^i,\omega_b^i\in E_i} \frac{|f(\omega_a^i) - f(\omega_b^i)|}{d_i^\alpha(\omega_a^i, \omega_b^i)} < +\infty \right\}, \tag{1.1}$$

$$\|f\|_{\widetilde{\mathrm{Lip}}_\alpha(E_i)} := \sup_{\omega_a^i,\omega_b^i\in E_i} \frac{|f(\omega_a^i) - f(\omega_b^i)|}{d_i^\alpha(\omega_a^i, \omega_b^i)}.$$

Now, we can go on an define global functions which are locally Hölder of index α:

$$\widetilde{\mathbb{L}}_\alpha(E_i) := \left\{ f : \Omega \to \mathbb{R}, \text{measurable}, \text{UCAI}, \sup_{\overline{\omega}^i\in\overline{\Omega}^i} \left\| f(\overline{\omega}^i, \cdot) \right\|_{\widetilde{\mathrm{Lip}}_\alpha(E_i)} < +\infty \right\}, \tag{1.2}$$

$$\|f\|_{\widetilde{\mathbb{L}}_\alpha(E_i)} := \sup_{\overline{\omega}^i\in\overline{\Omega}^i} \left\| f(\overline{\omega}^i, \cdot) \right\|_{\widetilde{\mathrm{Lip}}_\alpha(E_i)}.$$

At last, we can define the globally Hölder functions:

$$\widetilde{\mathbb{L}}_\alpha(\Omega) := \left\{ f \in \bigcap_{i\in V} \widetilde{\mathbb{L}}_\alpha(E_i) : \sum_{i\in V} \|f\|_{\widetilde{\mathbb{L}}_\alpha(E_i)} < +\infty \right\}, \tag{1.3}$$

$$\|f\|_{\widetilde{\mathbb{L}}_\alpha(\Omega)} := \sum_{i\in V} \|f\|_{\widetilde{\mathbb{L}}_\alpha(E_i)}.$$

We did not define norms, but semi-norms: if two functions differ only by a constant, then the Hölder semi-norm of their difference is zero. Here, the interest in Lipschitz functions lies only in their variation: we shall quotient our spaces, saying that two functions are identical if they differ only by a constant. This kind of quotient is used in many articles, although with some changes (one may consider, for instance, Lipschitz functions whose value at some reference point is zero). One advantage of our method is that we will not care about constants. The drawback is that, once the quotient is done, it will forbid us to speak of the value of the function in any given point; however, we will still be able to define the difference of the function evaluated in two points, since such differences are the same for any two functions which differ only by a constant.

Definition 1.3.

Let i in V. For any two functions f and g from E_i (resp. Ω) to $\mathbb{R}$, we write $f \sim g$ if there exists some c in $\mathbb{R}$ such that $f = g + c$. The relation $\sim$ is obviously an equivalence relation on the set of functions from E_i (resp. Ω) to $\mathbb{R}$.

For any i in V, we define $\mathrm{Lip}_\alpha(E_i)$, $\|\cdot\|_{\mathrm{Lip}_\alpha(E_i)}$, $\mathbb{L}_\alpha(E_i)$, $\|\cdot\|_{\mathbb{L}_\alpha(E_i)}$, $\mathbb{L}_\alpha(\Omega)$ and $\|\cdot\|_{\mathbb{L}_\alpha(\Omega)}$ as the quotients of respectively $\widetilde{\mathrm{Lip}}_\alpha(E_i)$, $\|\cdot\|_{\widetilde{\mathrm{Lip}}_\alpha(E_i)}$, $\widetilde{\mathbb{L}}_\alpha(E_i)$, $\|\cdot\|_{\widetilde{\mathbb{L}}_\alpha(E_i)}$, $\widetilde{\mathbb{L}}_\alpha(\Omega)$ and $\|\cdot\|_{\widetilde{\mathbb{L}}_\alpha(\Omega)}$ by the relation $\sim$.

Lemma 1.4.

$\|\cdot\|_{\mathrm{Lip}_\alpha(E_i)}$ is a norm on $\mathrm{Lip}_\alpha(E_i)$ for every i in V and every α in $[0,1]$.

$\|\cdot\|_{\mathbb{L}_\alpha(\Omega)}$ is a norm on $\mathbb{L}_\alpha(\Omega)$.

Functions in $\widetilde{\mathbb{L}}_\alpha(\Omega)$ are bounded.

Proof.

We first focus on the two first statements. It is enough to prove that $\|\cdot\|_{\mathbb{L}_1(\Omega)}$ is a norm on $\mathbb{L}_1(\Omega)$, since the former statement is a particular case of this one (one only needs first to take d_i^α instead of d_i to get the result for all α in $[0,1]$, and then to take $V = \{i\}$ to get the result for $\mathrm{Lip}_\alpha(E_i)$). Showing that $\|\cdot\|_{\mathbb{L}_1(\Omega)}$ is a semi-norm is trivial; all we need to prove is that, if $\|f-g\|_{\mathbb{L}_1(\Omega)} = 0$ for f and g in $\mathbb{L}_1(\Omega)$, then $f = g$. This is the same as proving that, if $\left\|\tilde{f}-\tilde{g}\right\|_{\widetilde{\mathbb{L}}_1(\Omega)} = 0$ for $\tilde{f}$ and $\tilde{g}$ in $\widetilde{\mathbb{L}}_1(\Omega)$, then $\tilde{f}$ and $\tilde{g}$ differ only by a constant. Let $\tilde{f}$ and $\tilde{g}$ in $\widetilde{\mathbb{L}}_1(\Omega)$ be such that $\left\|\tilde{f}-\tilde{g}\right\|_{\widetilde{\mathbb{L}}_1(\Omega)} = 0$. Let us fix some ω_0 in Ω. Let ω in Ω. For any nonnegative integer n, we have:

$$\tilde{f}(\omega) - \tilde{f}(\omega_0) = \sum_{k=0}^{n} (\tilde{f}(\omega_0^{<k}, \omega^{\geq k}) - \tilde{f}(\omega_0^{\leq k}, \omega^{>k})) + \tilde{f}(\omega_0^{\leq n}, \omega^{>n}) - \tilde{f}(\omega_0).$$

Since $\tilde{f}$ is continuous at infinity, we get by making n go to infinity:

$$\tilde{f}(\omega) - \tilde{f}(\omega_0) = \sum_{n=0}^{+\infty} \tilde{f}(\omega_0^{<n}, \omega^{\geq n}) - \tilde{f}(\omega_0^{\leq n}, \omega^{>n}). \tag{1.4}$$

Since the same holds for $\tilde{g}$, we have:

$$\tilde{f}(\omega) - \tilde{g}(\omega) = \tilde{f}(\omega_0) - \tilde{g}(\omega_0) + \sum_{n=0}^{+\infty} (\tilde{f} - \tilde{g})(\omega_0^{<n}, \omega^{\geq n}) - (\tilde{f} - \tilde{g})(\omega_0^{\leq n}, \omega^{>n}).$$

For any integer n, by the definition of the Lipschitz norms:

$$|(\tilde{f} - \tilde{g})(\omega_0^{<n}, \omega^{\geq n}) - (\tilde{f} - \tilde{g})(\omega_0^{\leq n}, \omega^{>n})| \leq \left\| \tilde{f} - \tilde{g} \right\|_{\widetilde{\mathbb{L}}_1(E_n)} d_n(\omega^n, \omega_0^n) = 0.$$

Hence, $\tilde{f}(\omega) - \tilde{g}(\omega) = \tilde{f}(\omega_0) - \tilde{g}(\omega_0)$ for any configuration ω in Ω, and $\tilde{f}$ and $\tilde{g}$ differ only by a constant. All is left to prove is the third statement. Starting from Equality 1.4, it is straightforward:

$$\begin{aligned} \tilde{f}(\omega) &= \tilde{f}(\omega_0) + \sum_{n=0}^{+\infty} \tilde{f}(\omega_0^{<n}, \omega^{\geq n}) - \tilde{f}(\omega_0^{\leq n}, \omega^{>n}) \\ &\leq \tilde{f}(\omega_0) + \sum_{n=0}^{+\infty} \left\| \tilde{f} \right\|_{\widetilde{\mathbb{L}}_1(E_n)} d_n(\omega^n, \omega_0^n) \\ &\leq \tilde{f}(\omega_0) + \Delta \left\| \tilde{f} \right\|_{\widetilde{\mathbb{L}}_1(\Omega)}. \end{aligned}$$

□

Now that we have defined everything we need (at least for the current section) on the side of observables, we turn to the measures side. While defining the functions spaces was made at the same time we defined the norm on those spaces, here the measures spaces will be defined directly, but the distances and norms (Wasserstein distances and Wasserstein-induced norms) we put on them will require us to introduce a few tools, including Kantorovich-Rubinstein Theorem. The use of such distances in the wording of Dobrushin uniqueness criterion dates back to a follow-up of Dobrushin's original article [3] a couple of years later [4], but the freedom one has in the choice of the cost function is seldom noticed.

In the following, we denote by S any Polish space. We shall use occasionally the bra-ket notation. Let us recall that, for any finite measure μ on S and any bounded, measurable function f from S to $\mathbb{R}$, one may write:

$$\langle \mu, f \rangle := \int_S f \, \mathrm{d}\mu.$$

We denote by $\mathcal{P}(S)$ the space of probability measures on S, and by $\mathcal{M}_0(S)$ the space of finite measure μ on S such that $\langle \mu, 1 \rangle = 0$.

If μ belongs to $\mathcal{M}_0(\Omega)$ and $\tilde{f}$ to $\widetilde{\mathbb{L}}_1(\Omega)$, then $\langle \mu, \tilde{f} \rangle = \langle \mu, \tilde{f} + c \rangle$ for any given c in $\mathbb{R}$. Hence, $\langle \mu, f \rangle$ is well-defined for f in $\mathbb{L}_1(\Omega)$.

Definition 1.5.

Let μ and ν be two probability measures on S. A coupling of μ and ν is a probability measure π on $S \times S$ such that its first marginal is μ, and its second marginal ν.

We say that a measurable function c from $S \times S$ to $\mathbb{R}_+$ is a cost function if it is bounded, symmetric, and if it is null on the diagonal. Then, we can define the optimal transport cost between μ and ν with respect to c by:

$$W_c(\mu, \nu) := \inf_{\substack{\text{coupling } \pi \\ \text{between } \mu \text{ and } \nu}} \int_{S \times S} c \, \mathrm{d}\pi. \tag{1.5}$$

A classical interpretation of the quantity $W_c(\mu,\nu)$ is the following. We can see the two probability measures μ and ν as two different way to spread a unity mass on the space S. Then, as is suggested by its name, $W_c(\mu,\nu)$ is the lowest cost needed to transport the mass from its repartition μ to its repartition ν, given that the cost for transporting a unity mass from point x to point y is $c(x,y)$. If $c(x,y) = 1_{x\neq y}$, the distance W_c between two probability measures is, up to a factor 2, the total variation between these two measures.

We expose in the next lemma some of the well-known properties of the Wasserstein distance (see for instance [11], Chapters 4 and 6).

Lemma 1.6.

If the cost function c is a distance on S, then the corresponding functional $W_c(\cdot,\cdot)$ is a distance on the space $\mathcal{P}(S)$.

If the cost function c is lower semi-continuous, then for any two given probability measures μ and ν on S, there exists an optimal coupling which realizes the infimum in Equation 1.5.

If $c(x,y) = 1_{x\neq y}$, then W_c metrizes the strong topology on $\mathcal{P}(S)$.

If (S,d) is a Polish metric space, then W_d metrizes the weak topology on $\mathcal{P}(S)$. The distance W_d is then called the Wasserstein distance on $\mathcal{P}(S)$.

All those definition and properties can be extended by taking nonnegative finite measures of the same mass instead of probability measures. Then, they induce norms on the spaces $\mathcal{M}_0$:

Definition 1.7.

Let c be a distance on S. Let μ be in $\mathcal{M}_0(S)$. The measure μ can be decomposed as a difference of two nonnegative, finite measures: $\mu = \mu_+ - \mu_-$. We define:

$$\|\mu\|_c := W_c(\mu_+,\mu_-).$$

Then, $\|\cdot\|_c$ is a norm on $\mathcal{M}_0(S)$.

One of the essential tools of this article is the link between the measures and the observables. It is indirectly provided by Kantorovich-Rubinstein Theorem:

Theorem 1.8 (Kantorovich-Rubinstein)**.**

Let c be a lower semi-continuous metric on S, which is also used here to define the Lipschitz functions space on S. Let μ be in $\mathcal{M}_0(S)$. Then:

$$\|\mu\|_c = \sup_{\substack{f\in\mathbb{L}_1(S)\\ \|f\|_{\mathbb{L}_1(S)}\leq 1}} \langle \mu, f\rangle. \tag{1.6}$$

We now move towards applying this theorem to our setting. For this purpose, we define $\ell^1(V,\mathbb{R}_+)$ as the space of nonnegative, summable sequences on V; for any $(u_i)_{i\in V}$ in $\ell^1(V,\mathbb{R}_+)$, we put $\|u\|_{\ell^1} = \sum_{i\in V} u_i$. We define also, for any $(u_i)_{i\in V}$ in $\ell^1(V,\mathbb{R}_+)$, a pseudo-distance $\sum_{i\in V} u_i d_i$ on Ω:

$$\left(\sum_{i\in V} u_i d_i\right)(\omega_a,\omega_b) := \sum_{i\in V} u_i d_i(\omega_a^i,\omega_b^i). \tag{1.7}$$

Lemma 1.9.

Let $(u_i)_{i\in V}$ in $\ell^1(V,\mathbb{R}_+)$. Then, $\sum_{i\in V} u_i d_i$ is a lower semi-continuous pseudo-distance on Ω.

Proof.

Let i in V. The pseudo distance defined on Ω by $d_i(\omega_a, \omega_b) := d_i(\omega_a^i, \omega_b^i)$ is lower semi-continuous. Hence, for every positive integer N, the pseudo-distance $\sum_{i \le N} u_i d_i$ is also lower semi-continuous.

Moreover, since $\sum_{i \in V} \|u_i d_i\|_\infty \le \Delta \|u\|_{\ell^1}$, the sequence of functions $\left(\sum_{i \le N} u_i d_i\right)_{N \in \mathbb{N}}$ is nondecreasing and converges uniformly. As a consequence:

$$\sum_{i \in V} u_i d_i = \sup_{N \in \mathbb{N}} \sum_{i \le N} u_i d_i.$$

As a pointwise supremum of lower semi-continuous functions, the function $\sum_{i \in V} u_i d_i$ is also lower semi-continuous. The fact that it is a pseudo-distance is obvious. □

At last, we define a convenient norm on $\mathcal{M}_0(\Omega)$, by putting for any μ in $\mathcal{M}_0(\Omega)$:

$$\|\mu\|_{(d)} := \sup_{\substack{u \in \ell^1(V, \mathbb{R}_+) \\ \|u\|_{\ell^1(V,\mathbb{R}_+)} \le 1}} W_{\sum_{i \in V} u_i d_i}(\mu_+, \mu_-). \tag{1.8}$$

Lemma 1.10.

$\|\cdot\|_{(d)}$ *is a norm on* $\mathcal{M}_0(\Omega)$, *and:*

$$\|\mu\|_{(d)} = \sup_{\substack{f \in \mathbb{L}_1(\Omega) \\ \|f\|_{\mathbb{L}_1(\Omega)} \le 1}} \langle \mu, f \rangle. \tag{1.9}$$

Proof.

Let μ be in $\mathcal{M}_0(\Omega)$.

Let u in $\ell^1(V, \mathbb{R}_+)$. First, we apply Kantorovich-Rubinstein Theorem 1.8 to the function $\sum_{i \in V} u_i d_i$, which is a lower semi-continuous pseudo-distance on $\Omega \times \Omega$ by Lemma 1.9.

$$W_{\sum_{i \in V} u_i d_i}(\mu_+, \mu_-) = \sup_{\substack{f \in \mathrm{Lip}_1(\Omega, \sum_{i \in V} u_i d_i) \\ \|f\|_{\mathrm{Lip}_1(\Omega, \sum_{i \in V} u_i d_i)} \le 1}} \langle \mu, f \rangle. \tag{1.10}$$

The main issue here is to identify the space of Lipschitz functions with respect to the pseudo-metric $\sum_{i \in V} u_i d_i$. Let f be 1-Lipschitz with respect to $\sum_{i \in V} u_i d_i$, where u belongs to $\ell^1(V, \mathbb{R}_+)$.

Obviously, for all i in V, we have $\|f\|_{\widetilde{\mathrm{Lip}}_1(E_i)} \le u_i$. Moreover, let $N \in \mathbb{N}$ and ω_a, ω_b in Ω be such that $\omega_a^i = \omega_b^i$ for all $i \le N$. Then,

$$\left(\sum_{i \in V} u_i d_i\right)(\omega_a, \omega_b) \le \Delta \sum_{i > N} u_i.$$

The sequence $(u_i)_{i \in V}$ being summable, this converges to 0 as N goes to infinity. Hence, f belongs to $\widetilde{\mathbb{L}}_1(\Omega)$, with $\|f\|_{\widetilde{\mathbb{L}}_1(\Omega)} \le \sum_{i \in V} u_i$ and $\|f\|_{\widetilde{\mathrm{Lip}}_1(E_i)} \le u_i$ for all i.

Conversely, if f belongs to $\widetilde{\mathbb{L}}_1(\Omega)$, as a consequence of Equation (1.4), for all ω_a and ω_b in Ω:

$$|f(\omega_a) - f(\omega_b)| \leq \sum_{i \in V} \|f\|_{\widetilde{\mathrm{Lip}_1}(E_i)} \, d_i(\omega_a^i, \omega_b^i) = \left(\sum_{i \in V} \|f\|_{\widetilde{\mathrm{Lip}_1}(E_i)} \, d_i \right) (\omega_a, \omega_b).$$

This proves that the set of functions which are 1-Lipschitz with respect to the pseudo-metric $\left(\sum_{i \in V} u_i d_i \right)$ for some $u \in \ell^1(V, \mathbb{R}_+)$ such that $\|u\|_{\ell^1(V,\mathbb{R}_+)} \leq 1$ is exactly the set of functions which are of $\widetilde{\mathbb{L}}_1(\Omega)$-norm at most 1. Hence:

$$\begin{aligned} \sup_{\substack{f \in \mathbb{L}_1(\Omega) \\ \|f\|_{\mathbb{L}_1(\Omega)} \leq 1}} \langle \mu, f \rangle &= \sup_{\substack{u \in \ell^1(V,\mathbb{R}_+) \\ \|u\|_{\ell^1(V,\mathbb{R}_+)} \leq 1}} \; \sup_{\substack{f \in \mathrm{Lip}_1(\Omega, \sum_{i \in V} u_i d_i) \\ \|f\|_{\mathrm{Lip}_1(\Omega, \sum_{i \in V} u_i d_i)} \leq 1}} \langle \mu, f \rangle \\ &= \sup_{\substack{u \in \ell^1(V,\mathbb{R}_+) \\ \|u\|_{\ell^1(V,\mathbb{R}_+)} \leq 1}} W_{\sum_{i \in V} u_i d_i}(\mu_+, \mu_-) \\ &= \|\mu\|_{(d)} . \end{aligned}$$

We get that $\|\cdot\|_{(d)}$ is a pseudo-norm on $\mathcal{M}_0(\Omega)$. Let u be a positive summable sequence on V, with $\|u\|_{\ell^1(V,\mathbb{R}_+)} \leq 1$. Then $\sum_{i \in V} u_i d_i$ is a distance on Ω, so that $W_{\sum_{i \in V} u_i d_i}(\mu_+, \mu_-)$ is non-zero for any non-zero measure μ in $\mathcal{M}_0(\Omega)$ and $\|\cdot\|_{(d)}$ is a norm on $\mathcal{M}_0(\Omega)$. □

1.2 Results

Now that the setting is clear, we can go on and prove a theorem akin to Dobrushin's [3] or Klein's [9]. So as to complete this goal, we introduce the interdependence coefficients c_{ij}, which quantify the influence of the configuration of a site i at time t upon the configuration at site j at time $t+1$.

Definition 1.11 (Interdependence coefficients)**.** *Let $i, j \in V$ be two sites. We define the coefficient c_{ij} by:*

$$c_{ij} := \sup_{\overline{\omega}^i \in \overline{\Omega}^i} \sup_{\substack{\omega_a^i, \omega_b^i \in E_i \\ \omega_a^i \neq \omega_b^i}} \frac{W_{d_j}(p_{j,\overline{\omega}^i,\omega_a^i}, p_{j,\overline{\omega}^i,\omega_b^i})}{d_i(\omega_a^i, \omega_b^i)}. \tag{1.11}$$

We shall call $M := (c_{ij})_{i,j \in V}$ the dependence matrix (associated to the dynamics). Let $n \in \mathbb{N}^*$. This matrix can be iterated, so that if we put :

$$c_{ij}^{(n)} := \sum_{\{i_1, \cdots, i_{n-1}\} \in V^{n-1}} c_{ii_1} c_{i_1 i_2} \cdots c_{i_{n-1} j}, \tag{1.12}$$

then $M^n := (c_{ij}^{(n)})_{i,j \in V}$.

In the worst case, if the configuration at site i at time t is switched from ω_a^i to ω_b^i, the distribution of the configuration at site j at time $t+1$ move at most by a distance of $c_{ij} d_i(\omega_a^i, \omega_b^i)$.

Theorem 1.12.

Assume that $\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij}$ is finite. Then, $\mathcal{L}$ acts continuously on $\mathbb{L}_1(\Omega)$, and:

$$\|\mathcal{L}\|_{\mathbb{L}_1(\Omega) \to \mathbb{L}_1(\Omega)} \leq \sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij}. \tag{1.13}$$

Proof.

First step:

We begin by choosing ω_a^i, ω_b^i in E_i and $\overline{\omega}^i$ in $\overline{\omega}^i$, and writing $\mathcal{L}f(\overline{\omega}^i,\omega_a^i)-\mathcal{L}f(\overline{\omega}^i,\omega_b^i)$ as a telescoping sum. Let n be a nonnegative integer.

$$\begin{aligned}\mathcal{L}f(\overline{\omega}^i,\omega_a^i)-\mathcal{L}f(\overline{\omega}^i,\omega_b^i) &= \int_\Omega f(\omega_1)\bigotimes_{j\in V} p_{j,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^j) - \int_\Omega f(\omega_1)\bigotimes_{j\in V} p_{j,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^j)\\ &= \sum_{j=0}^{n}\int_\Omega f(\omega_1)\,(p_{j,\overline{\omega}^i,\omega_a^i}-p_{j,\overline{\omega}^i,\omega_b^i})(\,\mathrm{d}\omega_1^j)\bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)\\ &\quad+\int_\Omega f(\omega_1)\bigotimes_{k\le n} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\left(\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)-\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\right).\end{aligned}$$

We show that the remainder of this sum converges to 0 as n goes to infinity. Let $\Pi_{n,\overline{\omega}^i,\omega_a^i,\omega_b^i}$ be any coupling between $\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_a^i}$ and $\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_b^i}$.

$$\begin{aligned}&\int_\Omega f(\omega_1)\bigotimes_{k\le n} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\left(\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)-\bigotimes_{k>n} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\right)\\ &\qquad= \int_\Omega f(\omega_1)-f(\omega_1^{\le n},\omega_2^{>n})\bigotimes_{k\le n} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\,\Pi_{n,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_1^{>n},\ \mathrm{d}\omega_2^{>n}).\end{aligned}$$

Since f is continuous at infinity, $|f(\omega_1)-f(\omega_1^{\le n},\omega_2^{>n})|$ decays to 0 when n goes to infinity uniformly in ω_1 and ω_2, which implies that the remainder of the telescoping sum also vanishes at infinity. Hence, we have:

$$\mathcal{L}f(\overline{\omega}^i,\omega_a^i)-\mathcal{L}f(\overline{\omega}^i,\omega_b^i) = \sum_{j=0}^{+\infty}\int_\Omega f(\omega_1)\,(p_{j,\overline{\omega}^i,\omega_a^i}-p_{j,\overline{\omega}^i,\omega_b^i})(\,\mathrm{d}\omega_1^j)\bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k). \tag{1.14}$$

Second step:

Now, we use (1.14) to get a upper bound on the norm of $\mathcal{L}f$. We keep the same notations as in the first step, choose j in V, and an optimal coupling $\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}$ between $p_{j,\overline{\omega}^i,\omega_a^i}$ and $p_{j,\overline{\omega}^i,\omega_b^i}$ (following Theorem 4.1 in [11], such a coupling always exists).

$$\begin{aligned}&\left|\int_\Omega f(\omega_1)\,(p_{j,\overline{\omega}^i,\omega_a^i}-p_{j,\overline{\omega}^i,\omega_b^i})(\,\mathrm{d}\omega_1^j)\bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)\right|\\ &\quad= \left|\int_\Omega f(\overline{\omega}_1^j,\omega_1^j)-f(\overline{\omega}_1^j,\omega_2^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_1^j,\ \mathrm{d}\omega_2^j)\bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)\right|\\ &\quad\le \int_\Omega \|f\|_{\mathbb{L}_1(E_j)}\,d_j(\omega_1^j,\omega_2^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_1^j,\ \mathrm{d}\omega_2^j)\bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^k)\bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^k)\\ &\quad\le \|f\|_{\mathbb{L}_1(E_j)}\,W_{d_j}(p_{j,\overline{\omega}^i,\omega_a^i},p_{j,\overline{\omega}^i,\omega_b^i})\\ &\quad\le c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\,d_i(\omega_a^i,\omega_b^i).\end{aligned}$$

Now, we sum this inequality over all j in V, and get:

$$\left|\mathcal{L}f(\overline{\omega}^i,\omega_a^i)-\mathcal{L}f(\overline{\omega}^i,\omega_b^i)\right|\leq\sum_{j\in V}c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\,d_i(\omega_a^i,\omega_b^i).$$

Since this is true for all $\overline{\omega}^i$ in $\overline{\Omega}^i$ and all ω_a^i and ω_a^i in E_i, obviously,

$$\|f\|_{\mathbb{L}_1(E_i)}\leq\sum_{j\in V}c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\,,$$

and:

$$\|f\|_{\mathbb{L}_1(\Omega)}\leq\sum_{i\in V}\sum_{j\in V}c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\leq\sum_{j\in V}\|f\|_{\mathbb{L}_1(E_j)}\sum_{i\in V}c_{ij}\leq\left(\sup_{j\in V}\sum_{i\in V}c_{ij}\right)\|f\|_{\mathbb{L}_1(\Omega)}\,.$$

However, we do not have proved yet that $\mathcal{L}f$ belongs to $\mathbb{L}_1(\Omega)$.

Third step:

All we have to check now is that $\mathcal{L}f$ is still continuous at infinity; this is where the assumption that all the $p_{i,\omega}$ are continuous at infinity will play its role. Let n be a nonnegative integer, and let ω_a and ω_b in Ω such that $\omega_a^{\leq n}=\omega_b^{\leq n}$ (that is, these two configurations are the same on the first n sites). For each i in V, we choose and optimal coupling $\pi_{i,\omega_a,\omega_b}$ between p_{i,ω_a} and p_{i,ω_b}. Let m be a nonnegative integer.

$$\begin{aligned}
&\mathcal{L}f(\omega_a)-\mathcal{L}f(\omega_b)\\
&\quad=\int_\Omega f(\omega_1)\left(\bigotimes_{k\in V}p_{k,\omega_a}(\,\mathrm{d}\omega_1^k)-\bigotimes_{k\in V}p_{k,\omega_b}(\,\mathrm{d}\omega_1^k)\right)\\
&\quad=\int_\Omega f(\omega_1)-f(\omega_2^)\bigotimes_{k\in V}\pi_{k,\omega_a,\omega_b}(\,\mathrm{d}\omega_1^k,\,\mathrm{d}\omega_2^k)\\
&\quad=\sum_{j=0}^{m}\int_\Omega f(\omega_2^{<j},\omega_1^{\geq j})-f(\omega_2^{\leq j},\omega_1^{>j})\bigotimes_{k\in V}\pi_{k,\omega_a,\omega_b}(\,\mathrm{d}\omega_1^k,\,\mathrm{d}\omega_2^k)\\
&\qquad\qquad+\int_\Omega f(\omega_2^{\leq m},\omega_1^{>m})-f(\omega_2)\bigotimes_{k\in V}\pi_{k,\omega_a,\omega_b}(\,\mathrm{d}\omega_1^k,\,\mathrm{d}\omega_2^k).
\end{aligned}$$

As in the first step of this proof, the remainder vanishes as m goes to infinity; we use the same kind of upper bound as in the second step.

$$|\mathcal{L}f(\omega_a)-\mathcal{L}f(\omega_b)|\leq\sum_{j=0}^{+\infty}\|f\|_{\mathbb{L}_1(E_j)}\,W_{d_j}(p_{j,\omega_a},p_{j,\omega_b}).\tag{1.15}$$

Moreover, we have for all j in V and nonnegative integer m:

$$W_{d_j}(p_{j,\omega_a},p_{j,\omega_b})=\sum_{i=0}^{m}W_{d_j}(p_{j,\omega_b^{\leq n+i},\omega_a^{>n+i}},p_{j,\omega_b^{\leq n+i+1},\omega_a^{>n+i+1}})+W_{d_j}(p_{j,\omega_b^{\leq n+m+1},\omega_a^{>n+m+1}},p_{j,\omega_b}).$$

Since each $p_{j,\omega}$ is continuous at infinity, the remainder vanishes as m goes to infinity. Hence:

$$W_{d_j}(p_{j,\omega_a},p_{j,\omega_b})=\sum_{i=0}^{+\infty}W_{d_j}(p_{j,\omega_b^{\leq n+i},\omega_a^{>n+i}},p_{j,\omega_b^{\leq n+i+1},\omega_a^{>n+i+1}})\leq\sum_{i=n+1}^{+\infty}c_{ij}d_i(\omega_a^i,\omega_b^i)\leq\Delta\sum_{i=n+1}^{+\infty}c_{ij}.\tag{1.16}$$

Now, we apply Inequality (1.16) to Inequality (1.15), and obtain:

$$|\mathcal{L}f(\omega_a) - \mathcal{L}f(\omega_b)| \leq \Delta \sum_{j=0}^{+\infty} \sum_{i=n+1}^{+\infty} c_{ij} \left\| f \right\|_{\mathbb{L}_1(E_j)}.$$

For all nonnegative integer n, we have:

$$\sum_{i=n+1}^{+\infty} c_{ij} \left\| f \right\|_{\mathbb{L}_1(E_j)} \leq \left(\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij} \right) \left\| f \right\|_{\mathbb{L}_1(E_j)},$$

the right hand side being a summable function of j as soon as the hypothesis of Theorem 1.12 is satisfied. Moreover, $\lim_{n \to +\infty} \sum_{i=n+1}^{+\infty} c_{ij} \left\| f \right\|_{\mathbb{L}_1(E_j)} = 0$ for all j. By the dominated convergence theorem,

$$\lim_{n \to +\infty} \Delta \sum_{j=0}^{+\infty} \sum_{i=n+1}^{+\infty} c_{ij} \left\| f \right\|_{\mathbb{L}_1(E_j)} = 0.$$

Since this upper bound also does not depend neither on ω_a nor on ω_b, the function $\mathcal{L}f$ is continuous at infinity and belongs to $\mathbb{L}_1(\Omega)$, which ends this proof. □

The bound 1.13 is worth commenting. If all c_{ij} are finite, the dependence matrix M is an infinite, nonnegative (in that all its coefficients are nonnegative) matrix. The operator norm of M when acting on $\ell^1(V)$ is $\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij}$; in particular, it acts continuously if and only if this expression if finite. One could make M act on different Banach spaces and get different expressions for the operator norm. It is not clear how one could translate such an action to a well-defined function space on Ω: if one makes M act on $\ell^p(V)$ for some $p > 1$, then the corresponding norm on functions on Ω do not control the supremum norm, and by duality we can not get a good control of the norm of $\mathcal{L}^*$ on the whole $\mathcal{M}_0(\Omega)$ space (although some control is still possible on some subspace of $\mathcal{M}_0(\Omega)$). However, one can still hope for good properties of the system such as in [5], where it is enough to control some operator norm of the dependence matrix.

Actually, the conclusion of Theorem 1.12 is valid under weaker assumptions. While we worked here with d_i as lower semi-continuous metrics, the same conclusion holds with costs functions which are merely nonnegative, symmetric, measurable, zero on the diagonal and uniformly bounded by some Δ. In other words, neither the positive definiteness, the triangular inequality nor the lower semi-continuity are necessary. However, those assumptions are needed for Corollary 1.13, and the lower semi-continuity makes for a simpler proof (it ensures the existence of optimal couplings, and spare us the use of couplings close to this optimum).

Corollary 1.13.

Assume that $\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij} < 1$. *Then, there exists at most one stationary measure.*

Proof.

Let μ_1 and μ_2 be two stationary measures. Since they are stationary, $\mathcal{L}^*(\mu_1 - \mu_2) = \mu_1 - \mu_2$.

$$\begin{aligned}
\|\mu_1 - \mu_2\|_{(d)} &= \sup_{\substack{f \in \mathbb{L}_1(\Omega) \\ \|f\|_{\mathbb{L}_1(\Omega)} \leq 1}} \langle \mu_1 - \mu_2, f \rangle \\
&= \sup_{\substack{f \in \mathbb{L}_1(\Omega) \\ \|f\|_{\mathbb{L}_1(\Omega)} \leq 1}} \langle \mu_1 - \mu_2, \mathcal{L} f \rangle \\
&\leq \sup_{\substack{f \in \mathbb{L}_1(\Omega) \\ \|f\|_{\mathbb{L}_1(\Omega)} \leq \|\mathcal{L}\|_{\mathbb{L}_1(\Omega) \to \mathbb{L}_1(\Omega)}}} \langle \mu_1 - \mu_2, f \rangle \\
&\leq \|\mathcal{L}\|_{\mathbb{L}_1(\Omega) \to \mathbb{L}_1(\Omega)} \|\mu_1 - \mu_2\|_{(d)} \\
&\leq \left(\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij} \right) \|\mu_1 - \mu_2\|_{(d)} .
\end{aligned}$$

Hence, if $\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij} < 1$, then $\|\mu_1 - \mu_2\|_{(d)} = 0$ and $\mu_1 = \mu_2$. □

This corollary states that, if the dependence matrix has an operator norm smaller than 1 when acting on $\ell^1(V)$, then there is at most one stationary measure. One can weaken this condition by iterating this matrix: if its spectral radius is smaller than 1, then the conclusion of this corollary holds. In other words, for the uniqueness of the stationary measure to be known it is enough to prove that, for some $n \in \mathbb{N}^*$,

$$\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij}^{(n)} < 1.$$

Remark 1.14.

One could also be interested in conditions which ensure that there exists at least one stationary measure. This is the case if each E_i is compact. Then, Ω is compact, and so is $\mathcal{P}(\Omega)$. Since $\mathcal{L}^*$ act continuously on $\mathcal{P}(\Omega)$, it has a fixed point by Schauder's theorem, and thus there exists at least one stationary measure.

However, on can use alternative conditions in some cases. For instance, assume that $\rho(\mathcal{L})$ is strictly smaller than 1. Then, we take any probability measure μ on Ω and define a sequence (u_n) by:

$$u_n := \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{L}^* \mu.$$

This is a Cauchy sequence, and if it has a limit then this limit is a stationary measure. All we need is a condition to ensure that $(\mathcal{P}(\Omega), \|\cdot - \cdot\|_{(d)})$ is complete. This is true if either each (E_i, d_i) is a metric Polish space, or if for some δ we know, for every site i and every points x and y in E_i, that $d_i(x, y) \geq \delta 1_{x \neq y}$ (then, we may work with a stronger topology on $\mathcal{P}(\Omega)$, which would still be complete).

A bound on the norm or on the spectral radius of the transfer operator can give much more than the mere uniqueness of the stationary measure.

Corollary 1.15 (Temporal decay of correlations)**.**

Assume that the system has a unique stationary probability measure μ. Let f and g be in $\mathbb{L}_1(\Omega)$. Then:

$$|\mathbb{E}_\mu(f(X_0) g(X_n)) - \mathbb{E}_\mu(f)\mathbb{E}_\mu(g)| \leq \|f\|_{\mathbb{L}^1(\Omega,\mu)} \|g\|_{\mathbb{L}_1(\Omega)} \left(\sup_{j \in V} \sum_{i=0}^{+\infty} c_{ij} \right)^n . \tag{1.17}$$

Proof.

By the definition of the transfer operator, we have:

$$\begin{aligned}|\mathbb{E}_\mu(f(X_0)g(X_n)) - \mathbb{E}_\mu(f)\mathbb{E}_\mu(g)| &= \left| \int_\Omega f(\omega)\left(\mathcal{L}^n g(\omega) - \mathbb{E}_\mu(g)\right)\mu(\,\mathrm{d}\omega)\right| \\ &\leq \|f\|_{\mathbb{L}^1(\Omega,\mu)} \|\mathcal{L}^n g - \mathbb{E}_\mu(g)\|_{\mathbb{L}^\infty(\Omega,\mu)} \\ &\leq \|f\|_{\mathbb{L}^1(\Omega,\mu)} \|g\|_{\mathbb{L}_1(\Omega)} \left(\sup_{j\in V}\sum_{i=0}^{+\infty} c_{ij}\right)^n ,\end{aligned}$$

where the last inequality comes from the fact that $\mathcal{L}^n g$ converges exponentially fast to $\mathbb{E}_\mu(g)$, which is the projection of g onto the eigenspace of $\mathcal{L}$ corresponding to the eigenvalue 1. □

2 Iteration of the transfer operator

In order to use the conclusion of Corollary 1.13, the transfer operator must be contracting, or at least the spectral radius of the dependancies matrix must be strictly smaller than one. However, one can easily find examples such that the dynamic has trivially a unique stationary measure and one has exponential convergence towards this measure, while the dependence matrix has a spectral radius greater than 1. For example, let us consider a sequence of independent and identically distributed ergodic Markov chains on a finite state space. Let N be the transition matrix, and π_1 the projection on the eigenspace for the eigenvalue 1 of the matrix N^{T}. Then, one has a unique stationary measure, but Corollary 1.13 can be applied if and only if $\|N^{\mathrm{T}} - \pi_1\|$ is strictly smaller than 1.

In the case of a sequence of independent dynamics, a simple workaround is to iterate the dynamic: thanks to the independence, the iterated dynamic keeps the property that the updates on every sites are independent, and to apply the conclusion of Corollary 1.13 it is enough that $\|(N^n)^{\mathrm{T}} - \pi_1\|$ is strictly smaller than 1 for some n - which is true if N is ergodic. However, this method can not be adapted in the interesting cases, when there is even the slightest dependence between the different sites. Our goal in this section is to iterate the transfer operator so as to get estimates of its spectral radius which behave nicely when one slightly perturbs a sequence of independent dynamics.

We shall begin again with a boring list of miscellaneous definitions, starting with a new function space and some operators. Let i in V and α in $[0,1]$; we put:

$$\mathbb{B}_\alpha(E_i) := \left\{ f : \overline{\Omega}^i \times E_i^2 \to \mathbb{R} :\ \exists K \geq 0, \forall \overline{\omega}^i, \omega_a^i, \omega_b^i \in \overline{\Omega}^i \times E_i^2, |f(\overline{\omega}^i, \omega_a^i, \omega_b^i)| \leq K d_i^\alpha(\omega_a^i, \omega_b^i) \right\} \quad (2.1)$$

$$\|f\|_{\mathbb{B}_\alpha(E_i)} := \sup_{\overline{\omega}^i \in \overline{\Omega}^i} \sup_{\omega_a^i, \omega_b^i \in E_i} \frac{f(\overline{\omega}^i, \omega_a^i, \omega_b^i)}{d_i^\alpha(\omega_a^i, \omega_b^i)}$$

For any i in V, for any function f from Ω to $\mathbb{R}$, we will denote $f_{/i}$ the function from $\overline{\Omega}^i \times E_i^2$ to $\mathbb{R}$ such that, for all $\overline{\omega}^i$ in $\overline{\Omega}^i$, for all ω_a^i and ω_b^i in E_i:

$$f_{/i}(\overline{\omega}^i, \omega_a^i, \omega_b^i) := f(\overline{\omega}^i, \omega_a^i) - f(\overline{\omega}^i, \omega_b^i)$$

Clearly, for any i in V and any f in $\mathbb{L}_\alpha(E_i)$, we have $\|f\|_{\mathbb{L}_\alpha(E_i)} = \left\|f_{/i}\right\|_{\mathbb{B}_\alpha(E_i)}$. We also have, for any f in $\mathbb{L}_\alpha(\Omega)$, the equality $\|f\|_{\mathbb{L}_\alpha(\Omega)} = \sum_{i\in V} \|f\|_{\mathbb{L}_\alpha(E_i)}$.

Let i and j in V. Let f in $\mathbb{B}_\alpha(E_j)$. We write:

$$P_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i} := \bigotimes_{k<j} p_{k,\overline{\omega}^i,\omega_b^i} \otimes \bigotimes_{k>j} p_{k,\overline{\omega}^i,\omega_a^i} \in \mathcal{P}(\overline{\Omega}^j).$$

We recall that $\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}$ denotes an optimal coupling between $p_{j,\overline{\omega}^i,\omega_a^i}$ and $p_{j,\overline{\omega}^i,\omega_b^i}$. Let us define the operators we shall use in the proof of Theorem 2.6:

$$(\mathcal{L}_{ij} f)(\overline{\omega}^i,\omega_a^i,\omega_b^i) = \int_\Omega f(\overline{\omega}_1^j,\omega_2^j,\omega_3^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^j,\,\mathrm{d}\omega_3^j)\,P_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\overline{\omega}_1^j). \tag{2.2}$$

Equation (1.14) becomes, for all i in V and f in $\mathbb{L}(\Omega)$:

$$(\mathcal{L}f)_{/i} = \sum_{j\in V} \mathcal{L}_{ij} f_{/j}. \tag{2.3}$$

We now need some reference configuration, to help us decouple the dynamics. Let $(\omega_{0,i}^i)_{i\in V}$ in Ω be this reference; we put:

$$\Delta_0 := \sup_{i\in V}\sup_{\omega^i\in E_i} d_i(\omega^i,\omega_{0,i}^i) \leq \Delta.$$

Let i in V and f in $\mathbb{B}_\alpha(E_i)$. We define:

$$\left(\widetilde{\mathcal{L}}_{ii} f\right)(\overline{\omega}^i,\omega_a^i,\omega_b^i) := \int_\Omega f(\overline{\omega}_1^i,\omega_2^i,\omega_3^i)\,\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^i,\,\mathrm{d}\omega_3^i)\,P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i), \tag{2.4}$$

and:

$$\left(\widehat{\mathcal{L}}_{ii} f\right)(\overline{\omega}^i,\omega_a^i,\omega_b^i) := \int_\Omega f(\overline{\omega}_1^i,\omega_2^i,\omega_3^i)\,\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^i,\,\mathrm{d}\omega_3^i)\,(P_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i} - P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i})(\,\mathrm{d}\overline{\omega}_1^i). \tag{2.5}$$

Trivially, one has:

$$(\mathcal{L}f)_{/i} = \sum_{\substack{j\in V\\ j\neq i}} \mathcal{L}_{ij} f_{/j} + \widehat{\mathcal{L}}_{ii} f_{/i} + \widetilde{\mathcal{L}}_{ii} f_{/i}. \tag{2.6}$$

We need to decompose further these operators in order to prove Proposition 2.7. Let i in V, and $\overline{\omega}_{0,i}^i$ in $\overline{\Omega}^i$. Let f in $\mathbb{B}_\alpha(E_i)$. We define:

$$\left(\widetilde{\widetilde{\mathcal{L}}}_{ii} f\right)(\overline{\omega}^i,\omega_a^i,\omega_b^i) := \int_\Omega f(\overline{\omega}_1^i,\omega_2^i,\omega_3^i)\,\pi_{i,\overline{\omega}_{0,i}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^i,\,\mathrm{d}\omega_3^i)\,P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i), \tag{2.7}$$

and:

$$\left(\widehat{\widehat{\mathcal{L}}}_{ii} f\right)(\overline{\omega}^i,\omega_a^i,\omega_b^i) := \int_\Omega f(\overline{\omega}_1^i,\omega_2^i,\omega_3^i)\,(\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i} - \pi_{i,\overline{\omega}_{0,i}^i,\omega_a^i,\omega_b^i})(\,\mathrm{d}\omega_2^i,\,\mathrm{d}\omega_3^i)\,P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i). \tag{2.8}$$

Once again, we put:

$$\Delta_0^t = \sup_{i\in V}\sup_{\substack{j\in V\\ j\neq i}}\sup_{\omega^j\in E_j} d_j(\omega^j,\omega_{0,i}^j) \leq \Delta.$$

We also need to introduce some kind of reference dynamic on each site; they will be considered as a dynamic we perturb. We create them by specifying some configurations, and then looking at the dynamic on any site knowing this is the configuration outside of this site. Let i in V and $\overline{\omega}^i$ in $\overline{\Omega}^i$. There is a Markov dynamic on E_i such that, if the configuration at time n is ω^i, the configuration at time $n+1$ is chosen with the probability measure $p_{i,\overline{\omega}^i,\omega^i}$. We call $\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}}$ its transfer operator acting on Lipschitz functions on E_i. For any such function,

$$\left(\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}} f\right)(\omega^i) = \int_{E_i} f(\omega_1^i) p_{i,\overline{\omega}_i^i,\omega^i}(\,\mathrm{d}\omega_1^i).$$

2.1 Operator norms

We now give some estimates of the operator norms for the operators previously defined. The first lemma is the most straightforward, since the computation involved is almost the same as in the proof of Theorem 1.12.

Lemma 2.1.
For all i and j in V, for all α in $[0,1]$,

$$\|\mathcal{L}_{ij}\|_{\mathbb{B}_\alpha(E_j)\to\mathbb{B}_\alpha(E_i)} \le c_{ij}^\alpha, \tag{2.9}$$

and:

$$\left\|\widetilde{\mathcal{L}}_{ii}\right\|_{\mathbb{B}_\alpha(E_i)\to\mathbb{B}_\alpha(E_i)} \le c_{ii}^\alpha, \tag{2.10}$$

Proof.
Let f in $\mathbb{B}_\alpha(E_j)$, and $(\overline{\omega}^i, \omega_a^i, \omega_b^i)$ in $\overline{\Omega}^i \times E_i^2$.

$$\begin{aligned}
\left|(\mathcal{L}_{ij} f)\,(\overline{\omega}^i, \omega_a^i, \omega_b^i)\right| &= \left|\int_\Omega f(\overline{\omega}_1^j, \omega_2^j, \omega_3^j)\, P_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\overline{\omega}_1^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^j,\ \mathrm{d}\omega_3^j)\right| \\
&\le \|f\|_{\mathbb{B}_\alpha(E_j)} \int_{E_j^2} d_j^\alpha(\omega_2^j, \omega_3^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^j,\ \mathrm{d}\omega_3^j) \\
&\le \|f\|_{\mathbb{B}_\alpha(E_j)} \left(\int_{E_j^2} d_j(\omega_2^j, \omega_3^j)\,\pi_{j,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^j,\ \mathrm{d}\omega_3^j)\right)^\alpha \\
&= \|f\|_{\mathbb{B}_\alpha(E_j)}\, W_{d_j}^\alpha(p_{j,\overline{\omega}^i,\omega_a^i}, p_{j,\overline{\omega}^i,\omega_b^i}) \\
&\le c_{ij}^\alpha\, \|f\|_{\mathbb{B}_\alpha(E_j)}\, d_i^\alpha(\omega_a^i, \omega_b^i).
\end{aligned}$$

Since this holds for every $(\overline{\omega}^i, \omega_a^i$ and $\omega_b^i)$ in $\overline{\Omega}^i \times E_i^2$, we have proven Inequality (2.9). The proof for Inequality (2.10) is virtually the same. □

Lemma 2.2 is the crux of the matter: if we can afford to lose some regularity, then the perturbations can be adequately controlled by the non-diagonal terms.

Lemma 2.2.
For all i in V, for all α in $[0,1]$, for all f in $\mathbb{L}_1(\Omega)$,

$$\left\|\widehat{\mathcal{L}}_{ii} f_{/i}\right\|_{\mathbb{B}_\alpha(E_i)} \le 2\Delta_0^{1-\alpha} c_{ii}^\alpha \|f\|_{\mathbb{L}_1(E_i)}^\alpha \left(\sum_{\substack{j\in V\\ j\neq i}} c_{ij}\, \|f\|_{\mathbb{L}_1(E_j)}\right)^{1-\alpha}. \tag{2.11}$$

Proof.
Let i in V, f in $\mathbb{L}_1(E_i)$ and $(\overline{\omega}^i, \omega_a^i, \omega_b^i)$ in $\overline{\Omega}^i \times E_i^2$. Let $\eta > 0$.
If $d_i(\omega_a^i, \omega_b^i) \le \eta$, we compute:

$$\begin{aligned}
\left|\left(\widehat{\mathcal{L}}_{ii} f_{/i}\right)(\overline{\omega}^i, \omega_a^i, \omega_b^i)\right| &\le \left|\int_\Omega f_{/i}(\overline{\omega_1}^i, \omega_2^i, \omega_3^i)\,\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^i,\ \mathrm{d}\omega_3^i)\, P_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\overline{\omega}_1^i)\right| \\
&\quad + \left|\int_\Omega f(\overline{\omega_1}^i, \omega_2^i, \omega_3^i)\,\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_2^i,\ \mathrm{d}\omega_3^i)\, P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i)\right| \\
&\le 2c_{ii} \left\|f_{/i}\right\|_{\mathbb{B}_1(E_i)} d_i(\omega_a^i, \omega_b^i) \\
&\le 2c_{ii} \|f\|_{\mathbb{L}_1(E_i)}\, \eta^{1-\alpha} d_i^\alpha(\omega_a^i, \omega_b^i).
\end{aligned}$$

If $d_i(\omega_a^i, \omega_b^i) \geq \eta$, we compute:

$$\left|\left(\widehat{\mathcal{L}}_{ii} f_{/i}\right)(\overline{\omega}^i, \omega_a^i, \omega_b^i)\right| \leq \left|\int_\Omega f(\overline{\omega}_1^i, \omega_2^i)\,(P_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i} - P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i})(\,\mathrm{d}\overline{\omega}_1^i)\,p_{i,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_2^i)\right|$$
$$+\left|\int_\Omega f(\overline{\omega}_1^i, \omega_2^i)\,(P_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i} - P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i})(\,\mathrm{d}\overline{\omega}_1^i)\,p_{i,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_2^i)\right|$$
$$\leq 2\Delta_0\left(\sum_{\substack{j\in V\\ j\neq i}} c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\right)$$
$$\leq 2\Delta_0\left(\sum_{\substack{j\in V\\ j\neq i}} c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\right)\eta^{-\alpha} d_i^\alpha(\omega_a^i, \omega_b^i).$$

Now, we optimize the parameter η by choosing:

$$\eta = \left(c_{ii}\,\|f\|_{\mathbb{L}_1(E_i)}\right)^{-1}\left(\Delta_0 \sum_{\substack{j\in V\\ j\neq i}} c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\right).$$

Hence, we have :

$$\left\|\widehat{\mathcal{L}}_{ii} f_{/i}\right\|_{\mathbb{B}_\alpha(E_i)} \leq 2\left(c_{ii}\,\|f\|_{\mathbb{L}_1(E_i)}\right)^\alpha\left(\Delta_0 \sum_{\substack{j\in V\\ j\neq i}} c_{ij}\,\|f\|_{\mathbb{L}_1(E_j)}\right)^{1-\alpha}.$$

□

Lemma 2.3 is a mere translation of a natural hypothesis on the operators $\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}}$.

Lemma 2.3.

Assume that there are some nonnegative real numbers K and λ such that, for all n in $\mathbb{N}$, for all i in V and all sequence of configurations $\overline{\omega}_1^i$, ..., $\overline{\omega}_n^i$ in $\overline{\Omega}^i$:

$$\left\|\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}_{\mathrm{n}}}\cdots\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}_{1}}\right\|_{\mathrm{Lip}_1(E_i)\to\mathrm{Lip}_1(E_i)} \leq K\lambda^n. \tag{2.12}$$

Then, for all i in V, all n in $\mathbb{N}$ and all f in $\mathbb{L}_1(E_i)$,

$$\left\|\left(\widetilde{\mathcal{L}}_{ii}\right)^n f_{/i}\right\|_{\mathbb{B}_1(E_i)} \leq K\lambda^n\,\|f\|_{\mathbb{L}_1(E_i)}. \tag{2.13}$$

Proof.

Let i in V and n in $\mathbb{N}^*$. For f in $\mathbb{L}_1(E_i)$ and $(\overline{\omega}^i, \omega_a^i, \omega_b^i)$ in $\overline{\Omega}^i \times E_i^2$,

$$\left(\widetilde{\mathcal{L}}_{ii}\right)^n f_{/i}(\overline{\omega}^i, \omega_a^i, \omega_b^i) = \int\cdots\int f(\overline{\omega}_n^i, \omega_{a,n}^i) - f(\overline{\omega}_n^i, \omega_{b,n}^i)\,\pi_{i,\overline{\omega}_{n-1}^i,\omega_{a,n-1}^i,\omega_{b,n-1}^i}(\,\mathrm{d}\omega_{a,n}^i,\ \mathrm{d}\omega_{b,n}^i)$$
$$P_{i,\overline{\omega}_{n-1}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_n^i)\cdots\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_{a,1}^i,\ \mathrm{d}\omega_{b,1}^i)\,P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i)$$
$$= \int\cdots\int f(\overline{\omega}_n^i, \omega_{a,n}^i) - f(\overline{\omega}_n^i, \omega_{b,n}^i)\,\pi_{i,\overline{\omega}_{n-1}^i,\omega_{a,n-1}^i,\omega_{b,n-1}^i}(\,\mathrm{d}\omega_{a,n}^i,\ \mathrm{d}\omega_{b,n}^i)$$
$$\cdots\pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_{a,1}^i,\ \mathrm{d}\omega_{b,1}^i)P_{i,\overline{\omega}_{n-1}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_n^i)\cdots P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i).$$

Let $(\overline{\omega}_1^i, \cdots, \overline{\omega}_n^i)$ in $(\overline{\Omega}^i)^n$.

$$\begin{aligned}
\int \cdots \int & f(\overline{\omega}_n^i, \omega_{a,n}^i) - f(\overline{\omega}_n^i, \omega_{b,n}^i)\, \pi_{i,\overline{\omega}_{n-1}^i, \omega_{a,n-1}^i, \omega_{b,n-1}^i}(\,\mathrm{d}\omega_{a,n}^i,\, \mathrm{d}\omega_{b,n}^i) \cdots \pi_{i,\overline{\omega}^i,\omega_a^i,\omega_b^i}(\,\mathrm{d}\omega_{a,1}^i,\, \mathrm{d}\omega_{b,1}^i) \\
&= \int_{E_i} f(\overline{\omega}_n^i, \omega_n^i) \left(\int \cdots \int p_{i,\overline{\omega}_{n-1}^i, \omega_{n-1}^i}(\,\mathrm{d}\omega_n^i) \cdots p_{i,\overline{\omega}^i,\omega_a^i}(\,\mathrm{d}\omega_1^i) \right. \\
&\qquad\qquad \left. - \int \cdots \int p_{i,\overline{\omega}_{n-1}^i, \omega_{n-1}^i}(\,\mathrm{d}\omega_n^i) \cdots p_{i,\overline{\omega}^i,\omega_b^i}(\,\mathrm{d}\omega_1^i) \right) \\
&= \left(\mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i_{n-1}}} \cdots \mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i_1}} \mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i}} f(\overline{\omega}_n^i, \cdot) \right) (\omega_a^i) \\
&\qquad\qquad - \left(\mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i_{n-1}}} \cdots \mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i_1}} \mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i}} f(\overline{\omega}_n^i, \cdot) \right) (\omega_b^i).
\end{aligned}$$

At this point, we use our hypothesis, and integrate:

$$\left| \left(\widetilde{\mathcal{L}}_{ii} \right)^n f_{/i}(\overline{\omega}^i, \omega_a^i, \omega_b^i) \right| \leq K \lambda^n \left\| f \right\|_{\mathbb{L}_1(E_i)} d_i(\omega_a^i, \omega_b^i).$$

This is true for all ω_a^i, ω_b^i and $\overline{\omega}^i$, which ends this proof. □

Lemmas 2.4 and 2.5 (as well as their proofs) are very similar to Lemmas 2.2 and 2.3 respectively; they shall allow us to work with a less stringent hypothesis on the operators $\mathcal{L}_{\mathrm{ref,i,\overline{\omega}^i}}$.

Lemma 2.4.

For all i in V, for all α in $[0,1]$, for all f in $\mathbb{L}_1(\Omega)$,

$$\left\| \widehat{\widehat{\mathcal{L}}}_{ii} f_{/i} \right\|_{\mathbb{B}_\alpha(E_i)} \leq 2(\Delta_0^t)^{1-\alpha} c_{ii}^\alpha \left(\sum_{\substack{j \in V \\ j \neq i}} c_{ji} \right)^{1-\alpha} \left\| f \right\|_{\mathbb{L}_1(E_i)}. \tag{2.14}$$

Proof.

This proof is very close to the proof of Lemma 2.2; we highlight the main differences. Let i in V, f in $\mathbb{L}_1(E_i)$ and $(\overline{\omega}^i, \omega_a^i, \omega_b^i)$ in $\overline{\Omega}^i \times E_i^2$. Let $\eta > 0$.

If $d_i(\omega_a^i, \omega_b^i) \leq \eta$, we still have:

$$\left| \left(\widehat{\widehat{\mathcal{L}}}_{ii} f_{/i} \right) (\overline{\omega}^i, \omega_a^i, \omega_b^i) \right| \leq 2 c_{ii} \left\| f \right\|_{\mathbb{L}_1(E_i)} \eta^{1-\alpha} d_i^\alpha(\omega_a^i, \omega_b^i).$$

If $d_i(\omega_a^i, \omega_b^i) \geq \eta$, we compute:

$$\begin{aligned}
\left| \left(\widehat{\widehat{\mathcal{L}}}_{ii} f_{/i} \right) (\overline{\omega}^i, \omega_a^i, \omega_b^i) \right| &\leq \left| \int_\Omega f(\overline{\omega}_1^i, \omega_2^i)\, (p_{i,\overline{\omega}^i,\omega_a^i} - p_{i,\overline{\omega}_{0,i}^i,\omega_a^i})(\,\mathrm{d}\omega_2^i)\, P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i) \right| \\
&\qquad + \left| \int_\Omega f(\overline{\omega}_1^i, \omega_2^i)\, (p_{i,\overline{\omega}^i,\omega_b^i} - p_{i,\overline{\omega}_{0,i}^i,\omega_b^i})(\,\mathrm{d}\omega_2^i)\, P_{i,\overline{\omega}^i,\omega_{0,i}^i,\omega_{0,i}^i}(\,\mathrm{d}\overline{\omega}_1^i) \right| \\
&\leq 2\Delta_0^t \left(\sum_{\substack{j \in V \\ j \neq i}} c_{ji} \left\| f \right\|_{\mathbb{L}_1(E_i)} \right) \\
&\leq 2\Delta_0^t \left(\sum_{\substack{j \in V \\ j \neq i}} c_{ji} \left\| f \right\|_{\mathbb{L}_1(E_i)} \right) \eta^{-\alpha} d_i^\alpha(\omega_a^i, \omega_b^i).
\end{aligned}$$

We get the wanted result by optimizing the parameter η. □

Lemma 2.5.

Assume that there are some nonnegative real numbers K and λ such that, for all i in V there exists a configuration $\overline{\omega}_i^i$ in $\overline{\Omega}^i$ such that, for all n in $\mathbb{N}$:

$$\left\|\mathcal{L}^n_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}}\right\|_{\mathrm{Lip}_1(E_i)\to\mathrm{Lip}_1(E_i)} \leq K\lambda^n. \tag{2.15}$$

Assume furthermore that we use this reference configuration in order to decompose $\widetilde{\mathcal{L}}_{ii}$ into $\widehat{\widetilde{\mathcal{L}}}_{ii} + \widetilde{\widetilde{\mathcal{L}}}_{ii}$, i.e. that $\overline{\omega}_i^i = \overline{\omega}_{0,i}^i$. Then, for all i in V, all n in $\mathbb{N}$ and all f in $\mathbb{L}_1(E_i)$,

$$\left\|\left(\widetilde{\widetilde{\mathcal{L}}}_{ii}\right)^n f_{/i}\right\|_{\mathbb{B}_1(E_i)} \leq K\lambda^n \left\|f\right\|_{\mathbb{L}_1(E_i)}. \tag{2.16}$$

The proof of this lemma is essentially the same as the proof of Lemma 2.3.

2.2 Evaluating the norm of the transfer operator

Since we are interested in the behavior of product systems which are perturbed by some coupling between the dynamics at different site, it is convenient to quantify the perturbation. Let α be in $[0,1]$. We define:

$$\varepsilon(\alpha) := \sup_{j\in V} \sum_{\substack{i\in V\\ i\neq j}} c_{ij}^\alpha, \tag{2.17}$$

and:

$$\varepsilon := \varepsilon(1). \tag{2.18}$$

When working in the setting of Proposition 2.7, we shall need a control not only on the columns of the dependence matrix, but also on its rows. Thus, we put:

$$\varepsilon^t := \sup_{i\in V} \sum_{\substack{j\in V\\ j\neq i}} c_{ij}. \tag{2.19}$$

The next two propositions give, under different assumptions on the restricted dynamics, a bound on the norm of $\mathcal{L}^n$ when mapping $\mathbb{L}_1$ to $\mathbb{L}_\alpha$ which behaves nicely when ε goes to 0. The loss of smoothness can not be avoided in such a general setting, and will be addressed later for two special cases (discrete state spaces and iterated function systems).

Proposition 2.6.

Assume that there are some nonnegative real numbers K and λ such that, for all n in $\mathbb{N}$, for all i in V and all sequence of configurations $\overline{\omega}_1^i$, ..., $\overline{\omega}_n^i$ in $\overline{\Omega}^i$:

$$\left\|\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}_{\mathrm{n}}}\cdots\mathcal{L}_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}_{1}}\right\|_{\mathrm{Lip}_1(E_i)\to\mathrm{Lip}_1(E_i)} \leq K\lambda^n. \tag{2.20}$$

If $K \geq 1$ and $\lambda > 0$, then for all $\alpha \in [0,1]$ and $n \in \mathbb{N}$,

$$\left\|\mathcal{L}^n\right\|_{\mathbb{L}_1(\Omega)\to\mathbb{L}_\alpha(\Omega)} \leq \Delta^{1-\alpha}K\lambda^n + n\Delta^{1-\alpha}\left(\frac{\varepsilon}{\lambda} + 2\left(\frac{\Delta_0\varepsilon}{\Delta K\lambda}\right)^{1-\alpha}\right)\max(K^\alpha\lambda^\alpha + \varepsilon(\alpha), K\lambda+\varepsilon)^n. \tag{2.21}$$

Proof.

We shall evaluate $\|\mathcal{L}^n f\|_{\mathbb{L}_\alpha(\Omega)}$, with n an nonnegative integer, using a decomposition into elementary operators such as (2.6). We iterate this decomposition, stopping whenever we encounter some $\widehat{\mathcal{L}}_{ii}$ or when we can not decompose anything anymore (that is, when we have iterated this operation n times). We then get (in all the following sums, and for all i, one shall read $\widetilde{\mathcal{L}}_{ii}$ instead of $\mathcal{L}_{ii}$):

$$\begin{aligned}
(\mathcal{L}^n f)_{/i_0} &= \sum_{i_1\in V} \mathcal{L}_{i_0 i_1} \left(\mathcal{L}^{n-1} f\right)_{/i_1} + \widehat{\mathcal{L}}_{i_0 i_0} \left(\mathcal{L}^{n-1} f\right)_{/i_0} \\
&= \cdots \\
&= \sum_{(i_1,\cdots,i_n)\in V^n} \mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{n-1} i_n} f_{/i_n} \\
&\qquad + \sum_{k=0}^{n-1} \sum_{(i_1,\cdots,i_k)\in V^k} \mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{k-1} i_k} \widehat{\mathcal{L}}_{i_k i_k} \left(\mathcal{L}^{n-k-1} f\right)_{/i_k}.
\end{aligned}$$

We now take the $\mathbb{L}_\alpha(E_{i_0})$ norm and sum over all i_0:

$$\|\mathcal{L}^n f\|_{\mathbb{L}_\alpha(\Omega)} \leq \Delta^{1-\alpha} \sum_{(i_0,\cdots,i_n)\in V^{n+1}} \left\|\mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{n-1} i_n} f_{/i_n}\right\|_{\mathbb{B}_1(E_{i_0})} \tag{2.22}$$

$$+ \sum_{k=1}^{n} \sum_{(i_0,\cdots,i_{k-1})\in V^k} \left\|\mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{k-2} i_{k-1}} \widehat{\mathcal{L}}_{i_{k-1} i_{k-1}} \left(\mathcal{L}^{n-k} f\right)_{/i_{k-1}}\right\|_{\mathbb{B}_\alpha(E_{i_0})}. \tag{2.23}$$

First step:

We take care of the first part of this decomposition (2.22). Let $(i_0, ...i_n)$ in V^{n+1}. We split the whole sum into different parts, each one corresponding to a certain number of identical operators $\widetilde{\mathcal{L}}_{i_n i_n}$ in a row at the rigt end of the string of operators. If $i_{n-k} = \cdots = i_n$, an application of Lemma 2.1 and Lemma 2.3 gives us:

$$\left\|\mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{n-1} i_n} f_{/i_n}\right\|_{\mathbb{B}_1(E_{i_0})} \leq c_{i_0 i_1} \cdots c_{i_{n-k-1} i_{n-k}} K \lambda^k \|f\|_{\mathbb{L}_1(E_{i_n})}.$$

Hence, we get:

$$\begin{aligned}
&\sum_{(i_0,\cdots,i_n)\in V^{n+1}} \left\|\mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{n-1} i_n} f_{/i_n}\right\|_{\mathbb{B}_1(E_{i_0})} \\
&\quad \leq \sum_{\substack{(i_0,\cdots,i_n)\in V^{n+1} \\ i_{n-1}\neq i_n}} c_{i_0 i_1} \cdots c_{i_{n-1} i_n} \|f\|_{\mathbb{L}_1(E_{i_n})} \\
&\qquad + K \sum_{k=1}^{n} \sum_{\substack{(i_0,\cdots,i_{n-k})\in V^{n-k+1} \\ i_{n-k-1}\neq i_{n-k}}} c_{i_0 i_1} \cdots c_{i_{n-k-1} i_{n-k}} \lambda^k \|f\|_{\mathbb{L}_1(E_{i_{n-k}})}.
\end{aligned}$$

Now, let us fix some i in V and some integer k between 0 and $n-1$. Then:

$$\begin{aligned}
&\sum_{\substack{(i_0,\cdots,i_{n-k})\in V^{n-k+1}\\ i_{n-k-1}\neq i_{n-k}=i}} c_{i_0 i_1}\cdots c_{i_{n-k-1}i_{n-k}} \|f\|_{\mathbb{L}_1(E_i)} \\
&\quad = \left(\sum_{\substack{i_{n-k-1}\in V\\ i_{n-k-1}\neq i}} c_{i_{n-k-1}i} \sum_{i_{n-k-2}\in V} c_{i_{n-k-2}i_{n-k-1}} \cdots \sum_{i_0\in V} c_{i_0 i_1} \right) \|f\|_{\mathbb{L}_1(E_i)} \\
&\quad \leq \left(\sup_{i_1\in V} \sum_{i_0\in V} c_{i_0 i_1} \right) \left(\sum_{\substack{i_{n-k-1}\in V\\ i_{n-k-1}\neq i}} c_{i_{n-k-1}i} \sum_{i_{n-k-2}\in V} c_{i_{n-k-2}i_{n-k-1}} \cdots \sum_{i_1\in V} c_{i_1 i_2} \right) \|f\|_{\mathbb{L}_1(E_i)} \\
&\quad \leq \cdots \\
&\quad \leq \left(\sup_{i_1\in V} \sum_{i_0\in V} c_{i_0 i_1} \right)^{n-k-1} \left(\sup_{i_1\in V} \sum_{\substack{i_0\in V\\ i_0\neq i_1}} c_{i_0 i_1} \right) \|f\|_{\mathbb{L}_1(E_i)} \\
&\quad \leq \varepsilon (K\lambda+\varepsilon)^{n-k-1} \|f\|_{\mathbb{L}_1(E_i)},
\end{aligned}$$

where we use the fact that $\sup_{i\in V} c_{ii} \leq K\lambda$ (this inequality, by far not optimal, let us reduce the number of constants in our formulas). We sum this inequality over all k:

$$\begin{aligned}
&\sum_{\substack{(i_0,\cdots,i_n)\in V^{n+1}\\ i_n=i}} \left\| \mathcal{L}_{i_0 i_1}\cdots \mathcal{L}_{i_{n-1}i_n} f_{/i} \right\|_{\mathbb{B}_1(E_{i_0})} \\
&\quad \leq \left(\varepsilon (K\lambda+\varepsilon)^{n-1} + K\varepsilon \sum_{k=1}^{n-1} \lambda^k (K\lambda+\varepsilon)^{n-k-1} + K\lambda^n \right) \|f\|_{\mathbb{L}_1(E_i)} \\
&\quad \leq \left(n\frac{\varepsilon}{\lambda} (K\lambda+\varepsilon)^n + K\lambda^n \right) \|f\|_{\mathbb{L}_1(E_i)}.
\end{aligned}$$

At last, summing over all i the last inequality gives us:

$$\sum_{(i_0,\cdots,i_n)\in V^{n+1}} \left\| \mathcal{L}_{i_0 i_1}\cdots \mathcal{L}_{i_{n-1}i_n} f_{/i} \right\|_{\mathbb{B}_1(E_{i_0})} \leq \left(n\frac{\varepsilon}{\lambda} (K\lambda+\varepsilon)^n + K\lambda^n \right) \|f\|_{\mathbb{L}_1(\Omega)}. \tag{2.24}$$

Second step:

We now compute estimates for the second part of the decomposition (2.23). For this part, we lose some regularity: hence, we control its α-Hölder norm instead of its Lipschitz norm. Let $1 \leq k \leq n$

and i in V. Let f in $\mathbb{L}_1(\Omega)$. Applying Lemma 2.1 and Lemma 2.2 gives us:

$$
\begin{aligned}
&\sum_{(i_0,\cdots,i_{k-1})\in V^k} \left\| \mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{k-2} i_{k-1}} \widehat{\mathcal{L}}_{i_{k-1} i_{k-1}} f_{/i_{k-1}} \right\|_{\mathbb{B}_\alpha(E_{i_0})} \\
&\quad \leq \sum_{(i_0,\cdots,i_{k-1})\in V^k} c_{i_0 i_1}^\alpha \cdots c_{i_{k-2} i_{k-1}}^\alpha \left\| \widehat{\mathcal{L}}_{i_{k-1} i_{k-1}} f_{/i_{k-1}} \right\|_{\mathbb{B}_\alpha(E_{i_{k-1}})} \\
&\quad \leq 2 \sum_{(i_0,\cdots,i_{k-1})\in V^k} c_{i_0 i_1}^\alpha \cdots c_{i_{k-2} i_{k-1}}^\alpha \left(c_{i_{k-1} i_{k-1}} \|f\|_{\mathbb{L}_1(E_{i_{k-1}})} \right)^\alpha \left(\Delta_0 \sum_{\substack{j\in V \\ j\neq i_{k-1}}} c_{i_{k-1} j} \|f\|_{\mathbb{L}_1(E_j)} \right)^{1-\alpha} \\
&\quad \leq 2 \left(\sum_{(i_0,\cdots,i_{k-1})\in V^k} c_{i_0 i_1}^\alpha \cdots c_{i_{k-2} i_{k-1}}^\alpha c_{i_{k-1} i_{k-1}} \|f\|_{\mathbb{L}_1(E_{i_{k-1}})} \right)^\alpha \\
&\qquad\qquad \times \left(\Delta_0 \sum_{(i_0,\cdots,i_{k-1})\in V^k} c_{i_0 i_1}^\alpha \cdots c_{i_{k-2} i_{k-1}}^\alpha \sum_{\substack{j\in V \\ j\neq i_{k-1}}} c_{i_{k-1} j} \|f\|_{\mathbb{L}_1(E_j)} \right)^{1-\alpha} \\
&\quad \leq 2 \left((K^\alpha \lambda^\alpha + \varepsilon(\alpha))^{k-1} K\lambda \sum_{i\in V} \|f\|_{\mathbb{L}_1(E_i)} \right)^\alpha \left(\Delta_0 \varepsilon \left(K^\alpha \lambda^\alpha + \varepsilon(\alpha)\right)^{k-1} \sum_{j\in V} \|f\|_{\mathbb{L}_1(E_j)} \right)^{1-\alpha} \\
&\quad \leq 2 (K\lambda)^\alpha (\Delta_0 \varepsilon)^{1-\alpha} \left(K^\alpha \lambda^\alpha + \varepsilon(\alpha)\right)^{k-1} \|f\|_{\mathbb{L}_1(\Omega)} .
\end{aligned}
$$

Now, we use this inequality with $\mathcal{L}^{n-k} f$, and sum over all k:

$$
\begin{aligned}
&\sum_{k=1}^{n} \sum_{(i_0,\cdots,i_k)\in V^k} \left\| \mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{k-2} i_{k-1}} \widehat{\mathcal{L}}_{i_{k-1} i_{k-1}} \left(\mathcal{L}^{n-k} f\right)_{/i_{k-1}} \right\|_{\mathbb{B}_\alpha(E_{i_0})} \\
&\qquad \leq 2(K\lambda)^\alpha (\Delta_0 \varepsilon)^{1-\alpha} \sum_{k=0}^{n-1} \left(K^\alpha \lambda^\alpha + \varepsilon(\alpha)\right)^k (K\lambda + \varepsilon)^{n-k-1} \|f\|_{\mathbb{L}_1(\Omega)} \qquad (2.25) \\
&\qquad \leq 2n \left(\frac{\Delta_0 \varepsilon}{K\lambda} \right)^{1-\alpha} \max(K^\alpha \lambda^\alpha + \varepsilon(\alpha), K\lambda + \varepsilon)^n \|f\|_{\mathbb{L}_1(\Omega)} .
\end{aligned}
$$

Finally, we sum the inequalities (2.24) and (2.25):

$$
\|\mathcal{L}^n\|_{\mathbb{L}_1(\Omega)\to\mathbb{L}_\alpha(\Omega)} \leq \Delta^{1-\alpha} K\lambda^n + n\Delta^{1-\alpha} \left(\frac{\varepsilon}{\lambda} + 2 \left(\frac{\Delta_0 \varepsilon}{\Delta K \lambda} \right)^{1-\alpha} \right) \max(K^\alpha \lambda^\alpha + \varepsilon(\alpha), K\lambda + \varepsilon)^n .
$$

□

Now, we provide another version of this proposition under a condition which is less restrictive and easier to check.

Proposition 2.7.
Assume that there are some nonnegative real numbers K and λ such that, for all i in V there exists a configuration $\overline{\omega}_i^i$ in $\overline{\Omega}^i$ such that, for all n in $\mathbb{N}$:

$$
\left\| \mathcal{L}^n_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}} \right\|_{\mathrm{Lip}_1(E_i)\to\mathrm{Lip}_1(E_i)} \leq K\lambda^n. \qquad (2.26)
$$

If $K \geq 1$ and $\lambda > 0$, then for all $\alpha \in [0,1]$ and $n \in \mathbb{N}$,

$$\|\mathcal{L}^n\|_{\mathbb{L}_1(\Omega)\to\mathbb{L}_\alpha(\Omega)} \leq \Delta^{1-\alpha} K\lambda^n + n\Delta^{1-\alpha}\left(\frac{\varepsilon}{\lambda} + 2\left(\frac{\Delta_0\varepsilon + \Delta_0^t\varepsilon^t}{\Delta K\lambda}\right)^{1-\alpha}\right)\max(K^\alpha\lambda^\alpha + \varepsilon(\alpha), K\lambda+\varepsilon)^n. \tag{2.27}$$

Proof.
The proof for this theorem is essentially the same as the proof of Theorem 2.6. The only difference is the introduction of a new member in the decomposition, similar to (2.23). Now, for all i, one shall read $\widetilde{\widetilde{\mathcal{L}}}_{ii}$ instead of $\mathcal{L}_{ii}$:

$$\begin{aligned}
(\mathcal{L}^n f)_{/i_0} &= \sum_{\substack{i_1\in V\\ i_1\neq i_0}} \mathcal{L}_{i_0 i_1}\left(\mathcal{L}^{n-1}f\right)_{/i_1} + \widehat{\mathcal{L}}_{i_0i_0}\left(\mathcal{L}^{n-1}f\right)_{/i_0} + \widehat{\widehat{\mathcal{L}}}_{i_0i_0}\left(\mathcal{L}^{n-1}f\right)_{/i_0} \\
&= \cdots \\
&= \sum_{(i_1,\cdots,i_n)\in V^n} \mathcal{L}_{i_0i_1}\cdots\mathcal{L}_{i_{n-1}i_n} f_{/i_n} \\
&\qquad + \sum_{k=0}^{n-1}\sum_{(i_1,\cdots,i_k)\in V^k} \mathcal{L}_{i_0i_1}\cdots\mathcal{L}_{i_{k-1}i_k}\widehat{\mathcal{L}}_{i_ki_k}\left(\mathcal{L}^{n-k-1}f\right)_{/i_k} \\
&\qquad + \sum_{k=0}^{n-1}\sum_{(i_1,\cdots,i_k)\in V^k} \mathcal{L}_{i_0i_1}\cdots\mathcal{L}_{i_{k-1}i_k}\widehat{\widehat{\mathcal{L}}}_{i_ki_k}\left(\mathcal{L}^{n-k-1}f\right)_{/i_k}.
\end{aligned}$$

Then, we use Lemma 2.5 instead of Lemma 2.3. This new part is treated exactly the same way as the part (2.23), except that we use Lemma 2.4 instead of Lemma 2.2, and that the computation is simpler (we need not use the Hölder inequality). □

The products $K\lambda$ is larger than 1 in the interesting cases; indeed, $\|\mathcal{L}\|_{\mathbb{L}(\Omega)\to\mathbb{L}(\Omega)} \leq K\lambda$, so if $K\lambda$ is smaller than 1 then the spectral radius of the transfer operator stays smaller than 1 under small perturbations. Hence, Propositions 2.6 and 2.7 do not express any meaningful exponential convergence by themselves.

In the next section, we shall focus on some corollaries of Propositions 2.6 and 2.7. Since we have two very close sets of hypotheses, we shall state the following results only in the setting of Proposition 2.7. To get the corresponding result in the setting of Proposition 2.6 is straightforward: just erase the $\Delta_0^t\varepsilon^t$ in all the formulas.

3 Applications

3.1 Discrete metrics

A first interesting case is when the metrics d_i are uniformly bounded from below (and, as such metrize the discrete topology on each set), so that:

$$\delta = \inf_{i\in V}\inf_{\substack{\omega_a^i,\omega_b^i\in E_i\\ \omega_a^i\neq\omega_b^i}} d_i(\omega_a^i,\omega_b^i) > 0.$$

Since we have in such a situation $\delta 1_{x\neq y} \leq d(x,y) \leq \Delta 1_{x\neq y}$, we work in this subsection with avatars of the total variation norm.

When this inequality holds, the Hölder and Lipschitz spaces are the same, and are endowed with equivalent norms. Hence, we do not actually lose any regularity when we apply the method explained in Subsection 2.2. More precisely, the following holds for any Lipschitz function f and any α in $[0,1]$:

$$\delta^{1-\alpha} \|f\|_{\mathbb{L}_1(\Omega)} \leq \|f\|_{\mathbb{L}_\alpha(\Omega)} \leq \Delta^{1-\alpha} \|f\|_{\mathbb{L}_1(\Omega)}. \tag{3.1}$$

This allows us to derive the following theorems:

Theorem 3.1.
Assume that there are some real numbers $K \geq 1$ and $0 \leq \lambda \leq 1$ such that, for all i in V there exists a configuration $\overline{\omega}_i^i$ in $\overline{\Omega}^i$ such that, for all n in $\mathbb{N}$:

$$\left\| \mathcal{L}^n_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}} \right\|_{\mathrm{Lip}_1(E_i)\to\mathrm{Lip}_1(E_i)} \leq K\lambda^n. \tag{3.2}$$

If we put $\rho(\mathcal{L}) := \inf_{n\in\mathbb{N}^*} \|\mathcal{L}^n\|^{\frac{1}{n}}_{\mathbb{L}_1(\Omega)\to\mathbb{L}_1(\Omega)}$, then, for fixed K, λ and α,

$$\rho(\mathcal{L}) \leq \lambda + (1+\varepsilon(\alpha)) O_{K,\lambda,\alpha}\left(\frac{1}{|\ln(\varepsilon+\varepsilon^t)|}\right). \tag{3.3}$$

Actually, our proof will provide explicit (although not simple, and obviously not optimal) bounds on $\rho(\mathcal{L})$.

Proof.
Up to the choice of a higher λ, we can assume that $\lambda > 0$. An application of the norm equivalence (3.1) to the result of Proposition 2.7 gives us:

$$\begin{aligned} \|\mathcal{L}^n\|_{\mathbb{L}_1(\Omega)\to\mathbb{L}_1(\Omega)} \leq &\left(\frac{\Delta}{\delta}\right)^{1-\alpha} K\lambda^n \\ &+ n\left(\frac{\Delta}{\delta}\right)^{1-\alpha}\left(\frac{\varepsilon}{\lambda} + 2\left(\frac{\Delta_0\varepsilon + \Delta_0^t\varepsilon^t}{\Delta K\lambda}\right)^{1-\alpha}\right)\max(K^\alpha\lambda^\alpha + \varepsilon(\alpha), K\lambda+\varepsilon)^n. \end{aligned}$$

Since, for all integer n, the function $x \to x^{\frac{1}{n}}$ is concave, it is smaller than its tangent in any point. Hence:

$$\begin{aligned} \|\mathcal{L}^n\|^{\frac{1}{n}}_{\mathbb{L}_1(\Omega)\to\mathbb{L}_1(\Omega)} &\leq \left(\left(\frac{\Delta}{\delta}\right)^{1-\alpha} K\right)^{\frac{1}{n}}\lambda \\ &\quad + \left(\left(\frac{\Delta}{\delta}\right)^{1-\alpha} K\right)^{\frac{1}{n}}\left(\frac{\varepsilon}{K} + 2\frac{\lambda}{K}\left(\frac{\Delta_0\varepsilon+\Delta_0^t\varepsilon^t}{\Delta K\lambda}\right)^{1-\alpha}\right) \\ &\quad \times \max\left(\frac{K^\alpha\lambda^\alpha+\varepsilon(\alpha)}{\lambda}, K+\frac{\varepsilon}{\lambda}\right)^n \\ &\leq \left(\left(\frac{\Delta}{\delta}\right)^{1-\alpha} K\right)^{\frac{1}{n}}\lambda \\ &\quad + \left(\varepsilon\left(\frac{\Delta}{\delta}\right)^{1-\alpha} + 2\lambda^\alpha\left(\frac{\Delta_0\varepsilon+\Delta_0^t\varepsilon^t}{\delta K}\right)^{1-\alpha}\right) \\ &\quad \times \max\left(1, \frac{K^\alpha\lambda^\alpha+\varepsilon(\alpha)}{\lambda}, K+\frac{\varepsilon}{\lambda}\right)^n. \end{aligned} \tag{3.4}$$

The second inequality above is very crude. It does not affect our computation of the asymptotic bounds and makes the computations easier, but one might want to avoid it when they look for a quantitative bound on $\rho(\mathcal{L})$.

Now, let us take any four nonnegative real numbers A, B, C and D such that $B \geq 1$, $C < 1$ and $D \geq 1$; we shall consider the quantity $\rho := \inf_{n \in \mathbb{N}^*} A \cdot B^{\frac{1}{n}} + C \cdot D^n$. First, since the function $n \mapsto B^{\frac{1}{n}}$ is nonincreasing and the function $n \mapsto D^n$ is nondecreasing, we know that:

$$\rho = \inf_{n \in \mathbb{N}^*} A \cdot B^{\frac{1}{n}} + C \cdot D^n \leq \inf_{n \in \mathbb{N}^*} \inf_{n-1 < x \leq n} A \cdot B^{\frac{1}{x}} + CD \cdot D^x = \inf_{x > 0} A \cdot B^{\frac{1}{x}} + CD \cdot D^x.$$

Now, everything is up to a good choice of x. One should remember that, in our setting, C can be seen as a small parameter (perturbation). As C goes to 0, everything else being constant, the corresponding choice of x should go to infinity so as to minimize $A \cdot B^{\frac{1}{x}}$. We shall try to choose a x small enough that $CD \cdot D^x$ still vanishes as C goes to 0 (for instance, it decays polynomially in terms of C), but not too small so that $A \cdot B^{\frac{1}{x}}$ still converges towards A. Let $R \in (0, 1)$; we pick for instance:

$$x = -\frac{R \ln C}{\ln D}.$$

This particular estimate gives us the following inequality:

$$\rho \leq A \cdot B^{-\frac{\ln D}{R \ln C}} + DC^{1-R}.$$

Now, one can get a quantitative bound on $\rho(\mathcal{L})$ by replacing A, B, C and D by the corresponding parameters. One can go even a small step further when it comes to asymptotics, at the cost of even huger and less intelligible non-asymptotic bounds, by basically removing the parameter R. For this, we pick:

$$x = -\frac{\ln C + 2 \ln |\ln C|}{\ln D},$$

which gives us, for C close to 0 and after the development of the exponential:

$$\rho \leq A + \frac{A \ln B \ln D}{|\ln C|} + o_{A,B,D}\left(\frac{1}{|\ln C|}\right).$$

This is a more precise version of the bound of Theorem 3.1, provided one replaces all the parameters with those of Equation (3.4). □

Remark 3.2.

These bounds could most probably be improved by using Theorem 1 of [10], for a bound which does not involve any α and is of the form :

$$\|\mathcal{L}\|_{\mathbb{L}_1(\Omega) \to \mathbb{L}_1(\Omega)} \leq \lambda + O(\varepsilon + \varepsilon^t).$$

Most of our computations remain necessary, so as to control the distance of the transfer operator $\mathcal{L}$ to the transfer operator for the product of independent systems and prove it is in, say, $O(\varepsilon + \varepsilon^t)$.

A simple consequence of these quantitative bounds is the following qualitative corollary, which addresses the issue of the uniqueness of stationary measures:

Corollary 3.3.

Assume that there are some real numbers $K \geq 1$ and $0 \leq \lambda < 1$ such that, for all i in V there exists a configuration $\overline{\omega}_i^i$ in $\overline{\Omega}^i$ such that, for all n in $\mathbb{N}$:

$$\left\| \mathcal{L}^n_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}} \right\|_{\mathrm{Lip}_1(E_i) \to \mathrm{Lip}_1(E_i)} \leq K \lambda^n. \tag{3.5}$$

Then, if ε, ε^t and $\varepsilon(\alpha)$ are close enough to 0, the system has at most one stationary measure.

The proof is roughly the same as the proof of Corollary 1.13: all we need is to take a higher iterate of $\mathcal{L}$, so that the norm of $\mathcal{L}^n$ is smaller than 1.

Despite what is written in this corollary, ε and $\varepsilon(\alpha)$ play different roles. Let us take a family of systems depending on a parameter, let say a temperature T, such that the strength of the interaction $\max(\varepsilon(T), \varepsilon^t(T))$ goes to 0 as T goes to infinity, and such that all those systems satisfy the assumptions of Corollary 3.3 with the same K and λ. Then, the conclusion of Corollary 3.3 is valid as long as $\varepsilon(\alpha)(T)$ stays bounded for some $\alpha < 1$ and for T large enough. Actually, $\varepsilon(\alpha)$ mays even go to infinity as T goes to infinity, though at a very slow rate. To put it simply, ε and ε^t really need to go to 0, while $\varepsilon(\alpha)$ only needs to be bounded, or at least not to increase too quickly.

3.2 A Lasota-Yorke inequality for coupled iterated function systems

Now we shall study iterated function systems. Instead of working directly on the space of configurations (that is, we choose the next configuration at random), we work on the space of local contractions (that is, at each site we choose a contraction at random, and we apply it to the current configuration). Essentially nothing changes: one can, from an iterated function system, induce a classical system on the space of configurations Ω. However, this additional layer can be useful, here mainly thanks to a Lasota-Yorke (or Döblin-Fortet) inequality.

For each i in V and σ in $[0,1]$, let $\text{Cont}_\sigma(E_i)$ be the set of contraction mappings from E_i to E_i whose Lipschitz constant is at most σ, i.e. the set of functions T from E_i to E_i such that:

$$\sup_{\substack{\omega_a^i, \omega_b^i \in E_i \\ \omega_a^i \neq \omega_b^i}} \frac{d_i(T(\omega_a^i), T(\omega_b^i))}{d_i(\omega_a^i, \omega_b^i)} \leq \sigma.$$

We define for each i in V and σ in $[0,1]$ the uniform distance on $\text{Cont}_\sigma(E_i)$ by:

$$d_i^*(T_a, T_b) = \sup_{\omega^i \in E_i} d_i(T_a(\omega^i), T_b(\omega^i)).$$

For all i in V, let (F_i, d_i^*) be a compact subset of $\text{Cont}_1(E_i)$ with its Borel σ-algebra. We put $\sigma = \sup_{i \in V} \inf\{\lambda \in [0,1] : F_i \subset \text{Cont}_\lambda(E_i)\}$, and $\Omega^* = \prod_{i \in V} F_i$.

Let $\mathcal{P}(F_i)$ be the space of probability measures on F_i. For each i in V, let $\omega \to p_{i,\omega}^*$ be a measurable and uniformly continuous at infinity application from Ω to $\mathcal{P}(F_i)$. The dynamic we study is the following: first, we choose an initial configuration ω_0 in Ω. Then, at time t, if the current configuration is ω, we choose for each site i a transformation T with distribution $p_{i,\omega}^*$ independently for each site, and apply this transformation to ω^i to get the configuration at time $t+1$ at site i. We construct this way a Markov process on Ω.

For every i and j in V, we define:

$$c_{ij}^* = \sup_{\overline{\omega}^i \in \overline{\Omega}^i} \sup_{\substack{\omega_a^i, \omega_b^i \in E_i \\ \omega_a^i \neq \omega_b^i}} \frac{W_{d_j^*}(p_{j,\overline{\omega}^i,\omega_a^i}^*, p_{j,\overline{\omega}^i,\omega_b^i}^*)}{d_i(\omega_a^i, \omega_b^i)},$$

and for very α in $[0,1]$:

$$C^*(\alpha) = \sup_{j \in V} \sum_{i \in V} (c_{ij}^*)^\alpha, \quad \varepsilon^*(\alpha) = \sup_{j \in V} \sum_{\substack{i \in V \\ i \neq j}} (c_{ij}^*)^\alpha, \quad (\varepsilon^t)^*(\alpha) = \sup_{i \in V} \sum_{\substack{j \in V \\ j \neq i}} (c_{ij}^*)^\alpha,$$

We will denote $C^*(1)$ by C^*, then $\varepsilon^*(1)$ by ε^*, and at last $(\varepsilon^t)^*(1)$ by $(\varepsilon^t)^*$. As in Section 1, we will apply the transfer operator $\mathcal{L}$ to the set of globally Lipschitz functions $\mathbb{L}_1(\Omega)$:

$$\mathcal{L}f(\omega) = \int_{\Omega^*} f(T_0\omega^0, T_1\omega^1, ...) \bigotimes_{i\in V} p^*_{i,\omega}(\,\mathrm{d}T_i).$$

Remark 3.4.

As we have said at the beginning of this subsection, we may look at an iterated function system in two different ways: either we choose a transformation randomly, an then apply it, or we choose directly the image of a configuration at random, which is equivalent to projection the probability measures from F_i onto E_i, by:

$$p_{i,\omega} = T(\omega^i)_* p^*_{i,\omega}.$$

Thus, we can go back to the setting we have used since the beginning of this article. Conversely, we can go from the "classical" setting to the iterated function systems setting by putting $T_{\omega^i} := \omega^i$, and:

$$p^*_{i,\omega} := (T_{\omega^i})_* p_{i,\omega}.$$

Note that the fact that the set E_i are Polish, and not necessarily compact, is not an issue: it is straightforward to generalize what we did with compact subsets of contractions to Polish subsets of contractions.

A particularly interesting consequence is that, for every i and j, we have $c_{ij} \leq c^*_{ij}$, where the former correspond to the "classical" system and the later to the iterated function systems. This can be proven easily by projecting an optimal coupling from F_i onto E_i.

Theorem 3.5 (Lasota-Yorke inequality)**.**

Let $\alpha \in [0,1]$. For all f in $\mathbb{L}_1(\Omega)$, we have:

$$\|\mathcal{L}f\|_{\mathbb{L}_1(\Omega)} \leq \sigma \|f\|_{\mathbb{L}_1(\Omega)} + C^*(\alpha) \|f\|_{\mathbb{L}_\alpha(\Omega)}. \tag{3.6}$$

Proof.

Theorem 1.12 ensures that $\mathcal{L}$ acts continuously on $\mathbb{L}_1(\Omega)$ as soon as C is finite. Following Remark 3.4, this is the case whenever C^* is finite. Hence, the continuity at infinity of $\mathcal{L}f$is proved.

Let f be in $\mathbb{L}_1(\Omega)$. Let $\overline{\omega}^i \in \overline{\Omega}^i$, as well as ω^i_a and $\omega^i_b \in E_i$. We put $\omega_a = (\overline{\omega}^i, \omega^i_a)$ and $\omega_b = (\overline{\omega}^i, \omega^i_b)$.

$$\begin{aligned}
\mathcal{L}f(\overline{\omega}^i, \omega^i_a) - &\mathcal{L}f(\overline{\omega}^i, \omega^i_b) \\
&= \int_{\Omega^*} f(T_0\omega^0_a, T_1\omega^1_a, ...) \bigotimes_{k\in V} p^*_{k,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_k) - \int_{\Omega^*} f(T_0\omega^0_b, T_1\omega^1_b, ...) \bigotimes_{k\in V} p^*_{k,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_k) \\
&= \sum_{j=0}^{+\infty} \int_{\Omega^*} f(T_0\omega^0_a, T_1\omega^1_a, ...) \bigotimes_{k<j} p^*_{k,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_k) \\
&\qquad \otimes (p^*_{j,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_j) - p^*_{j,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_j)) \otimes \bigotimes_{k>j} p^*_{k,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_k) \\
&\qquad + \int_{\Omega^*} f(T_0\omega^0_a, T_1\omega^1_a, ...) - f(T_0\omega^0_b, T_1\omega^1_b, ...) \bigotimes_{k\in V} p^*_{k,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_k).
\end{aligned}$$

Let j in V. Let $\pi^*_{j,\overline{\omega}^i,\omega^i_a,\omega^i_b}$ be an optimal coupling between $p^*_{j,\overline{\omega}^i,\omega^i_a}$ and $p^*_{j,\overline{\omega}^i,\omega^i_b}$. Then:

$$
\begin{aligned}
\int_{\Omega^*} f(T_0\omega_a^0, T_1\omega_a^1, ...) \bigotimes_{k<j} p^*_{k,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_k) \otimes (p^*_{j,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_j) - p^*_{j,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_j)) \otimes \bigotimes_{k>j} p^*_{k,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_k) \\
= \int_{\Omega^*} f(T_0\omega_a^0, T_1\omega_a^1, ..., T_j\omega_a^j, ...) - f(T_0\omega_a^0, T_1\omega_a^1, ..., T'_j\omega_a^j, ...) \\
\bigotimes_{k<j} p^*_{k,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_k) \otimes \pi^*_{j,\overline{\omega}^i,\omega^i_a,\omega^i_b}(\,\mathrm{d}T_j,\ \mathrm{d}T'_j) \otimes \bigotimes_{k>j} p^*_{k,\overline{\omega}^i,\omega^i_a}(\,\mathrm{d}T_k) \\
\leq \|f\|_{\mathbb{L}_\alpha(E_i)} \int_{F_j} d_j^\alpha(T_j\omega_a^j, T'_j\omega_a^j)\, \pi^*_{j,\overline{\omega}^i,\omega^i_a,\omega^i_b}(\,\mathrm{d}T_j,\ \mathrm{d}T'_j) \qquad (3.7) \\
\leq \|f\|_{\mathbb{L}_\alpha(E_i)} \int_{F_j} (d_j^*)^\alpha(T_j, T'_j)\, \pi^*_{j,\overline{\omega}^i,\omega^i_a,\omega^i_b}(\,\mathrm{d}T_j,\ \mathrm{d}T'_j) \\
= \|f\|_{\mathbb{L}_\alpha(E_i)} W_{(d_j^*)^\alpha}(p^*_{j,\overline{\omega}^i,\omega^i_a}, p^*_{j,\overline{\omega}^i,\omega^i_b}) \\
\leq \|f\|_{\mathbb{L}_\alpha(E_i)} c^*_{ij}(\alpha) d_i(\omega_a^i, \omega_b^i).
\end{aligned}
$$

On the other hand, we also have:

$$
\int_{\Omega^*} f(T_0\omega_a^0, T_1\omega_a^1, ...) - f(T_0\omega_b^0, T_1\omega_b^1, ...) \bigotimes_{j\in V} p^*_{j,\overline{\omega}^i,\omega^i_b}(\,\mathrm{d}T_j) \leq \sigma \|f\|_{\mathbb{L}_1(E_i)} d_i(\omega_a^i, \omega_b^i). \qquad (3.8)
$$

With (3.7) and (3.8) together, we get:

$$
\|\mathcal{L}f\|_{\mathbb{L}_1(E_i)} \leq \sigma \|f\|_{\mathbb{L}_1(E_i)} + \sum_{j\in V} c^*_{ij}(\alpha) \|f\|_{\mathbb{L}_\alpha(E_i)}. \qquad (3.9)
$$

The sum of (3.9) over all i in V gives us:

$$
\|\mathcal{L}f\|_{\mathbb{L}_1(\Omega)} \leq \sigma \|f\|_{\mathbb{L}(\Omega)} + C^*(\alpha) \|f\|_{\mathbb{L}_\alpha(\Omega)}.
$$

□

Usually, one works with iterated function systems on spaces for which the injection from $\mathbb{L}_1(\Omega)$ into $\mathbb{L}_\alpha(\Omega)$ is compact. Then, the Lasota-Yorke inequality is enough, with standard arguments (such as a theorem by H. Hennion [7]), to prove that the transfer operator acting on $\widetilde{\mathbb{L}}_1(E_i)$ is quasi-compact. Even if it does not ensure that the system has a unique stationary measure, the consequences of this argument are in general more than satisfying. However, we do not benefit here of such features; we shall nevertheless show that this Lasota-Yorke inequality remains useful.

3.3 Application to coupled iterated function systems

We have proved in Section 2 that, given a slightly perturbed product system, for any function f and for n large enough, we have a nice control of the Hölder norm of $\mathcal{L}^n f$ in term of the Lipschitz norm of f. Of course, this is not sufficient in general. In Subsection 3.1 we have shown that, if the Lipschitz and Hölder norm are equivalent, the uniqueness of the stationary measure is preserved by small perturbations of some systems. Now, we shall use the Lasota-Yorke inequality (3.6) to get back, in a different fashion, the regularity we lost in Proposition 2.7. Hence, we shall get results similar to the ones of Subsection 3.1, but for different systems and when discrete distances are not suitable.

Theorem 3.6.

Assume that there are some real numbers $K \geq 1$ and $0 \leq \lambda \leq 1$ such that, for all i in V there exists a configuration $\overline{\omega}_i^i$ in $\overline{\Omega}^i$ such that, for all n in $\mathbb{N}$:

$$\left\| \mathcal{L}^n_{\mathrm{ref,i},\overline{\omega}^{\mathrm{i}}} \right\|_{\mathrm{Lip}_1(E_i) \to \mathrm{Lip}_1(E_i)} \leq K\lambda^n. \tag{3.10}$$

If we put $\rho(\mathcal{L}) := \inf_{n \in \mathbb{N}^*} \|\mathcal{L}^n\|^{\frac{1}{n}}_{\mathbb{L}_1(\Omega) \to \mathbb{L}_1(\Omega)}$, then, for fixed K, λ, σ and α,

$$\rho(\mathcal{L}) \leq \max(\lambda, \sigma) + (1 + \varepsilon^*(\alpha)) O_{K,\lambda,\sigma,\alpha}\left(\frac{1}{|\ln(\varepsilon^* + (\varepsilon^t)^*)|} \right). \tag{3.11}$$

In particular, if ε^*, $(\varepsilon^t)^*$ and $\varepsilon^*(\alpha)$ are close enough to 0, the system has at most one stationary measure.

Proof.

Let us take some integer n. We iterate the Lasota-Yorke type inequality of Theorem 3.5.

$$\begin{aligned} \|\mathcal{L}^n f\|_{\mathbb{L}_1(\Omega)} &\leq \sigma \left\|\mathcal{L}^{n-1} f\right\|_{\mathbb{L}_1(\Omega)} + C^*(\alpha) \left\|\mathcal{L}^{n-1} f\right\|_{\mathbb{L}_\alpha(\Omega)} \\ &\leq \cdots \\ &\leq \sigma^n \|f\|_{\mathbb{L}_1(\Omega)} + C^*(\alpha) \sum_{k=0}^{n-1} \sigma^k \left\|\mathcal{L}^{n-k-1} f\right\|_{\mathbb{L}_\alpha(\Omega)}. \end{aligned}$$

Now, we apply Proposition 2.7:

$$\begin{aligned} \|\mathcal{L}^n f\|_{\mathbb{L}_1(\Omega)} &\leq \sigma^n \|f\|_{\mathbb{L}_1(\Omega)} \\ &\quad + C^*(\alpha) \sum_{k=0}^{n-1} \sigma^k \Delta^{1-\alpha} \Bigg[K\lambda^{n-k-1} + (n-k-1) \left(\frac{\varepsilon}{\lambda} + 2 \left(\frac{\Delta_0 \varepsilon + \Delta_0^t \varepsilon^t}{\Delta K \lambda} \right)^{1-\alpha} \right) \\ &\quad \times \max(K^\alpha \lambda^\alpha + \varepsilon^*(\alpha), K\lambda + \varepsilon)^{n-k-1} \Bigg] \|f\|_{\mathbb{L}_1(\Omega)} \\ &\leq \Bigg[(1 + nK\Delta^{1-\alpha} C^*(\alpha)) \max(\sigma, \lambda)^n + n^2 \Delta^{1-\alpha} \left(\frac{\varepsilon}{\lambda} + 2 \left(\frac{\Delta_0 \varepsilon + \Delta_0^t \varepsilon^t}{\Delta K \lambda} \right)^{1-\alpha} \right) \\ &\quad \times \max(K^\alpha \lambda^\alpha + \varepsilon^*(\alpha), K\lambda + \varepsilon, \sigma) \Bigg] \|f\|_{\mathbb{L}_1(\Omega)}. \end{aligned}$$

Since c_{ij} is smaller than c^*_{ij} for any i and j, we can replace ε and ε^t by ε^* and $(\varepsilon^t)^*$ respectively. This upper bound on the norm of the operator $\mathcal{L}^n$ is slightly different from the one we studied in the proof of Theorem 3.1. We actually have to look at bounds which look like:

$$\rho := \sup_{n \in N^*} A \cdot B^{\frac{1}{n}} + n \cdot C \cdot D^n,$$

However, even if the value of n we choose and the precise estimates we get do change a little, the conclusions of Theorem 3.1 and Corollary 3.3 still hold. □

Remark 3.7.

Some better bounds and less cumbersome computations may be achieved using perturbation theory alongside Lasota-Yorke inequalities. The work by G. Keller and C. Liverani [8] is central, although the hypotheses used in said article can be slightly weakened (mostly to avoid having to control the spectrum of the perturbated operator).

4 Spatial decay of correlations

We have developed in all the previous sections methods to prove that the spectral radius of the transfer operator $\mathcal{L}$ is smaller than 1. In other words, we are able to prove inequalities such that $\|\mathcal{L}^n\|_{\mathbb{L}(\Omega)\to\mathbb{L}(\Omega)} \leq K\lambda^n$, with $0 \leq \lambda < 1$, for suitable systems. We shall now use these results to show a feature of the corresponding systems, the spatial decay of correlations.

We recall that for any i and j in the vertex set V, the coefficient $c_{ij}^{(n)}$ can be interpreted as an evaluation of the influence of the configuration at site i at any time t on the configuration at site j at time $t+n$. We can define the basin of influence of a given site for a given time:

Definition 4.1 (Basin of influence)**.**
Let i be a site, and n a positive integer. The basin of influence of site i at time n is:

$$I(i,n) := \{j \in V \ : \ c_{ji}^{(n)} > 0\}.$$

Now, we start two prove a spatial decay of correlation for the stationary measure, assuming that it exists and that it is attractive at exponential speed (a feature we have now many ways to prove for suitable systems). We shall proceed in two steps. With Lemma 4.2, we prove that if μ is a product measure, then $(\mathcal{L}^*)^n\mu$ show some kind of spatial decay of correlation for any given n. With Lemma 4.3, we estimate the error between $(\mathcal{L}^*)^n\mu$ and the stationary measure.

Lemma 4.2.
Let $\mu = \bigotimes_{i\in V} \mu_i$ be a product measure on Ω. Let $((X_n^i)_{i\in V})_{n\in\mathbb{N}}$ be the Markov process such that the distribution of $(X_0^i)_{i\in V}$ is μ and the transition process is the one we have studied so far. Let us choose two distinct sites i and j in V. Let f be a Lipschitz function on E_i, and g a Lipschitz function on E_j. Then:

$$\begin{aligned} &|\mathbb{E}(f(X_n^i)g(X_n^j)) - \mathbb{E}(f(X_n^i))\mathbb{E}(g(X_n^j))| \\ &\qquad \leq \Delta \|f\|_{\mathrm{Lip}_1(E_i)} \|g\|_\infty \sum_{\substack{k\in I(i,n)\cap I(j,n) \\ c_{ki}^{(n)} \leq c_{kj}^{(n)}}} c_{ki}^{(n)} + \Delta \|g\|_{\mathrm{Lip}_1(E_j)} \|f\|_\infty \sum_{\substack{k\in I(i,n)\cap I(j,n) \\ c_{ki}^{(n)} > c_{kj}^{(n)}}} c_{kj}^{(n)}. \end{aligned} \tag{4.1}$$

Proof.
By recurrence, one can show that the law of X_n^i depends only on those X_0^k such that k belongs to $I(i,n)$. In particular, if i and j are two sites such that $I(i,n)$ and $I(j,n)$ are disjoint, then X_n^i and X_n^j are independent.

Now, let $((\tilde{X}_n^k)_{k\in V})_{n\in\mathbb{N}}$ be a process whose law is the same that $((X_n^k)_{k\in V})_{n\in\mathbb{N}}$, such that $\tilde{X}_0^k$ is independent of $(X_0^l)_{l\in V}$ for all k in $I(i,n)\cap I(j,n)$ with $c_{ki}^{(n)} \leq c_{kj}^{(n)}$, and such that $\tilde{X}_0^k = X_0^k$ otherwise. We define the same way a process $((\hat{X}_n^k)_{k\in V})_{n\in\mathbb{N}}$ whose law is the same that $((X_n^k)_{k\in V})_{n\in\mathbb{N}}$, such that $\hat{X}_0^k$ is independent of $(X_0^l)_{l\in V}$ and $(\tilde{X}_0^l)_{l\in V}$ for all k in $I(i,n)\cap I(j,n)$ with $c_{ki}^{(n)} > c_{kj}^{(n)}$, and such that $\hat{X}_0^k = X_0^k$ otherwise. By construction, we have:

$$\mathbb{E}(f(\tilde{X}_n^i)g(\hat{X}_n^j)) = \mathbb{E}(f(X_n^i))\mathbb{E}(g(X_n^j)). \tag{4.2}$$

We need an evaluation of the error term:

$$|\mathbb{E}((f(\tilde{X}_n^i) - f(X_n^i))g(\hat{X}_n^j))| \leq \|f\|_{\mathrm{Lip}_1(E_i)} \|g\|_\infty \mathbb{E}(d_i(X_n^i, \tilde{X}_n^i)) \leq \Delta \|f\|_{\mathrm{Lip}_1(E_i)} \|g\|_\infty \sum_{\substack{k\in I(i,n)\cap I(j,n) \\ c_{ki}^{(n)} \leq c_{kj}^{(n)}}} c_{ki}^{(n)}, \tag{4.3}$$

since any change in the value of X_0^k changes the law of X_n^i knowing X_0^k of at most $\Delta c_{ki}^{(n)}$ in Wasserstein distance. We get a similar upper bound for $|\mathbb{E}(f(X_n^i)(g(\hat{X}_n^j)-g(X_n^j)))|$. The equations (4.2) and (4.3) imply the lemma. $\square$

Lemma 4.3.
Let $\mu = \bigotimes_{i\in V} \mu_i$ be a product measure on Ω. Let $((X_n^i)_{i\in V})_{n\in\mathbb{N}}$ be the Markov process such that the distribution of $(X_0^i)_{i\in V}$ is μ and the transition process is the one we have studied so far. Let us choose two distinct sites i and j in V. Let f be a Lipschitz function on E_i, and g a Lipschitz function on E_j.

Let K and λ be such that $\|\mathcal{L}^n\|_{\mathbb{L}_1(\Omega)\mapsto\mathbb{L}_1(\Omega)} \leq K\lambda^n$, and assume that the system has a stationary probability measure ν. Then:

$$|\mathbb{E}_\mu(f(X_n^i)g(X_n^j)) - \mathbb{E}_\nu(f(X_n^i)g(X_n^j))| \leq \Delta(\|f\|_{\mathrm{Lip}_1(E_i)}\|g\|_\infty + \|g\|_{\mathrm{Lip}_1(E_j)}\|f\|_\infty)K\lambda^n \tag{4.4}$$

Proof.
First, we point out that, for any x_1, x_2 in E_i and any y_1, y_2 in E_j,

$$|f(x_1)g(x_1) - f(x_2)g(x_2)| \leq \|f\|_{\mathrm{Lip}_1(E_i)}\|g\|_\infty d_j(y_1,y_2) + \|g\|_{\mathrm{Lip}_1(E_j)}\|f\|_\infty d_i(x_1,x_2).$$

The contraction property of the operator $\mathcal{L}$ can be translated, on the measures side, to:

$$\|\nu - (\mathcal{L}^*)^n\mu\|_{(d)} \leq \Delta K\lambda^n.$$

Let π be an optimal coupling between (X_n^i, X_n^j) taken under the initial law μ and the same couple taken under the initial law ν, with respect to the distance $\|f\|_{\mathrm{Lip}(E_i)}\|g\|_\infty d_j + \|g\|_{\mathrm{Lip}(E_j)}\|f\|_\infty d_i$. By Equation (1.8), we know that:

$$\left|\int_{(E_i\times E_j)^2} f(x_1)g(y_1) - f(x_2)g(y_2)\,\mathrm{d}\pi\right| \leq \Delta(\|f\|_{\mathrm{Lip}_1(E_i)}\|g\|_\infty + \|g\|_{\mathrm{Lip}_1(E_j)}\|f\|_\infty)K\lambda^n.$$

$\square$

Now, we can state the main theorem of this section.

Theorem 4.4.
Let K and λ be such that $\|\mathcal{L}^n\|_{\mathbb{L}(\Omega)\mapsto\mathbb{L}(\Omega)} \leq K\lambda^n$ for all positive integer n, and assume that the system has a stationary probability measure ν. Let i and j be two distinct sites. Let f be a Lipschitz function on E_i, and g a Lipschitz function on E_j. Then:

$$|\mathbb{E}_\nu(fg) - \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)| \leq \Delta^2\|f\|_{\mathrm{Lip}_1(E_i)}\|g\|_{\mathrm{Lip}_1(E_j)} \inf_{n\in\mathbb{N}^*}\left(K\lambda^n + \frac{1}{2}\sum_{k\in I(i,n)\cap I(j,n)} \min(c_{ki}^{(n)}, c_{kj}^{(n)})\right). \tag{4.5}$$

Proof.

If we glue Lemma 4.2 and Lemma 4.3 together, we get:

$$|\mathbb{E}_\nu(fg) - \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)| \le \inf_{n\in\mathbb{N}^*} \left[\Delta \|f\|_{\mathrm{Lip}_1(E_i)} \|g\|_\infty \left(K\lambda^n + \sum_{\substack{k\in I(i,n)\cap I(j,n) \\ c_{ki}^{(n)} \le c_{kj}^{(n)}}} c_{ki}^{(n)} \right) \right.$$
$$\left. + \Delta \|g\|_{\mathrm{Lip}_1(E_j)} \|f\|_\infty \left(K\lambda^n + \sum_{\substack{k\in I(i,n)\cap I(j,n) \\ c_{ki}^{(n)} > c_{kj}^{(n)}}} c_{kj}^{(n)} \right) \right].$$

However, adding a constant to f or to g does not change the quantity $\mathbb{E}_\nu(fg) - \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)$. Hence, we might as well assume that f satisfies $\|f\|_\infty \le \Delta \|f\|_{\mathrm{Lip}_1(E_i)}/2$, and that g satisfies a similar inequality. This concludes the proof of this theorem. □

This theorem takes a very interesting form when one deals with systems with finite range interaction. We define a distance D on V by:

$$D(i,j) := \min\{n \in \mathbb{N}^* \ : \ \min(c_{ij}^{(n)}, c_{ij}^{(n)}) > 0\}, \quad i \neq j. \tag{4.6}$$

Let n be a positive integer, and choose two sites i and j. If $D(i,j) \ge 2n+1$, then for each site k, either $D(k,i) > n$ or $D(k,j) > n$. Hence, k can not belong to $I(i,n)$ and to $I(j,n)$ at the same time, which shows that the correlation may decay exponentially. This is what we state in the following corollary.

Corollary 4.5.
Let K and λ be such that $\|\mathcal{L}^n\|_{\mathbb{L}(\Omega)\mapsto\mathbb{L}(\Omega)} \le K\lambda^n$ for all positive integer n, and assume that the system has a stationary probability measure ν. Let i and j be two distinct sites. Let f be a Lipschitz function on E_i, and g a Lipschitz function on E_j. Then:

$$|\mathbb{E}_\nu(fg) - \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)| \le \Delta^2 \|f\|_{\mathrm{Lip}_1(E_i)} \|g\|_{\mathrm{Lip}_1(E_j)} K\lambda^{\frac{D(i,j)-1}{2}}. \tag{4.7}$$

If λ is strictly smaller than 1, we shall refer to the property described in Equation (4.7) as the "exponential spatial decay of correlations".

Remark 4.6.
In the same spirit, one can prove a spatial decay of correlations for functions which depend on a finite number of sites, or even for functions which depend on an infinite numbers of sites but not too much on sites far enough.

5 Examples

As for applications of our results, we present two versions of Ising model with synchronous update; the first one is a direct application of our Dobrushin-like uniqueness criterion, and the second one involves iterated function systems. In both cases, we shall show that for high enough temperatures, there is at most one equilibrium measure.

5.1 Discrete Ising model

Here, we shall work on $\mathbb{Z}^d$, and each site will have a spin of $+1$ or -1. Hence, for all i in $\mathbb{Z}^d$, we take $E_i := \{-1, +1\}$, and $d_i(1, +1) = 1$. Our distance on $\mathbb{Z}^d$, that we will denote by $|i - j|$ for any sites i and j, will be the l^1 distance (or, equivalently, the lentgh of the shortest path between i and j). Let X_n^i be the configuration at site i and at time n. We want the process $((X_n^i)_{i\in\mathbb{Z}^d})_{n\in\mathbb{N}}$ to be a Markov chain on $\Omega := \bigotimes_{i\in\mathbb{Z}^d} E_i$. We put:

$$\mathbb{P}(X_{n+1}^i = \sigma | (X_n^j)_{j\in\mathbb{Z}^d}) = \frac{e^{-\frac{E(X_n,\sigma)}{RT}}}{e^{-\frac{E(X_n,\sigma)}{RT}} + e^{\frac{E(X_n,\sigma)}{RT}}},$$

where $E(X_n, \sigma) := -\sum_{j\in\mathbb{Z}^d} J_{ji}\sigma X_n^j$, with $\sup_{j\in\mathbb{Z}^d} \sum_{i\in\mathbb{Z}^d} |J_{ij}|$ being finite. The parameter T is the temperature. Neither do we impose that J_{ii} be 0, so that the configuration of a given site at time $n+1$ may depend on its configuration at time n, nor do we assume that the law of the Markov chain be translation-invariant (which is equivalent to the property that $J_{i,j} = J_{0,j-i}$ for all i and j).

A particular Ising model is the one with $J_{ij} = 1_{|i-j|=1}$; we shall call this model the Ising model with closest neighbors interaction.

Let i and j be two sites of the lattice. We now compute the coefficients c_{ij}:

$$c_{ij} = \sup_{\varpi^i\in\overline{\Omega}^i} \frac{1}{1 + e^{2\frac{E(\varpi^i,1)}{RT}} e^{-\frac{2|J_{ij}|}{RT}}} - \frac{1}{1 + e^{2\frac{E(\varpi^i,1)}{RT}} e^{\frac{2|J_{ij}|}{RT}}}.$$

Very easily, one shows that the influence of site i onto site j is maximal when the cumulated influence of all other sites is zero, or in other words that:

$$c_{ij} \le \frac{1}{1 + e^{-\frac{2|J_{ij}|}{RT}}} - \frac{1}{1 + e^{\frac{2|J_{ij}|}{RT}}} = \frac{\sinh\left(\frac{2|J_{ij}|}{RT}\right)}{1 + \cosh\left(\frac{2|J_{ij}|}{RT}\right)} \le \tanh\left(\frac{2|J_{ij}|}{RT}\right). \tag{5.1}$$

As a corollary of Theorem 1.12, we have:

Corollary 5.1 (Ising model).

If $\sup_{j\in\mathbb{Z}^d} \sum_{i\in\mathbb{Z}^d} \frac{\sinh\left(\frac{2|J_{ij}|}{RT}\right)}{1 + \cosh\left(\frac{2J_{ij}}{RT}\right)} < 1$*, then the process has a unique stationary measure.*

In particular, if T is high enough there exists a unique stationary measure.

For the Ising model with closest neighbors interaction, there is a unique stationary measure ν as soon as $T > \frac{2}{k\,\mathrm{argtanh}(\frac{1}{d})}$. Moreover, for any function f from E_i to $\mathbb{R}$ and any function g from E_j to $\mathbb{R}$, we have:

$$|\mathbb{E}_\nu(fg) - \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)| \le (\delta f)(\delta g) \left(d \tanh\left(\frac{2}{RT}\right)\right)^{\frac{|i-j|-1}{2}}, \tag{5.2}$$

where $\delta f = |f(1) - f(-1)|$ (and the same for g).

Actually, we get much stronger results if the conditions of Theorem 1.12 are satisfied. For instance, the probability measures, when transported by the dynamic, get close to the unique stationary measure at exponential speed. Moreover, the precise estimate we have for the Ising model with closest neighbors interaction on $\mathbb{Z}^d$ can be generalized to any Ising model with closest neighbors interaction on a graph with finite maximum degree, such that the degree of any vertex is even.

Proof.

The first statement is a direct application of Corollary 1.13.

The second one is straightforward if we have a look at the following inequality:

$$c_{ij} \leq \sup_{j\in\mathbb{Z}^d} \sum_{i\in\mathbb{Z}^d} \tanh\left(\frac{2|J_{ij}|}{RT}\right) \leq \frac{2}{RT} \sup_{j\in\mathbb{Z}^d} \sum_{i\in\mathbb{Z}^d} |J_{ij}|.$$

The third one is basically an application of our Dobrushin-like criterion, but with a small change. Since we gain a factor 2, we need a slightly better bound for the coefficients c_{ij} than the one in equation (5.1). We can actually compute these coefficients explicitly, and we find then:

$$c_{ij} = \frac{1}{2} \tanh\left(\frac{2 1_{|i-j|=1}}{RT}\right),$$

which is what we want. The gain of a factor 2 is due to the fact that, for any sites i and j such that J_{ij} is nonzero, the value of the energy $E(\overline{\omega}^i)(1)$ can never be zero (because there is an odd number of equal influences), which leads to this improved bound.

The exponential spatial decay of correlations is a direct application of Corollary 4.5. □

Actually, if we apply directly Theorem 4.4 to the Ising model with closest neighbors interaction, we can prove that $\mathbb{E}_\nu(fg) = \mathbb{E}_\nu(f)\mathbb{E}_\nu(g)$ if $|i-j|$ is odd.

5.2 Ising model and iterated function systems

We propose at last another version of the Ising model, with an iterated function systems flavor. In order to simplify some of the arguments, we shall work on $\mathbb{Z}^d$ with translation-invariant systems.

For all i in V, we take $E_i := [-1, +1]$, and $d_i(x, y) = |x - y|$. We also define two transformations on each E_i: $f_{-1}(x) = -1 + (x+1)/3$, and $f_1(x) = 1 + (x-1)/3$. When we choose f_1, the configuration gets closer to 1, while it gets closer to -1 if we choose f_{-1}. Let X_n^i be the configuration at site i and at time n. We still want the process $((X_n^i)_{i\in\mathbb{Z}^d})_{n\in\mathbb{N}}$ to be a Markov chain on $\Omega := \bigotimes_{i\in\mathbb{Z}^d} E_i$. We put:

$$\mathbb{P}(X_{n+1}^i = f_\sigma(X_n^i) | (X_n^j)_{j\in\mathbb{Z}^d}) = \frac{e^{-\frac{E(X_n,\sigma)}{RT}}}{2\cosh\left(\frac{E(X_n,\sigma)}{RT}\right)},$$

where:

$$E(X_n, \sigma) := -\sum_{\substack{j\in\mathbb{Z}^d \\ j\neq i}} J_{ji}\sigma X_n^j - J_{ii} T \sigma X_n^i,$$

with $\sup_{j\in\mathbb{Z}^d} \sum_{i\in\mathbb{Z}^d} |J_{ij}|$ being finite. We also assume that the system is translation-invariant, which can be translated by $J_{i,j} = J_{0,j-i}$.

Now, we shall match this system with the diverse assumptions of Theorem 3.6. Let ω_0 be the configuration $(0, 0, \cdots)$. Let i be any site. The Markov chain induced on E_i is defined by the following formula:

$$\mathbb{P}(X_{n+1}^i = f_\sigma(X_n^i) | X_n^i) = \frac{e^{\frac{J_{ii}\sigma X_n^i}{R}}}{2\cosh\left(\frac{J_{ii}\sigma X_n^i}{R}\right)},$$

This iterated function system satisfies the assumptions of Theorem 1 in [1], and hence has a unique, attractive stationary probability measure. Moreover, the operator $\mathcal{L}_{\mathrm{ref,i},\overline{\omega_0}^{\mathrm{i}}}$ acting on Lipschitz functions on E_i is quasicompact. It comes for instance from a theorem of H. Hennion [7], and the remark that this operator satisfies a Lasota-Yorke inequality (for a proof of this inequality, see Subsection 3.2):

$$\left\|\mathcal{L}_{\mathrm{ref,i},\overline{\omega_0}^{\mathrm{i}}} f\right\|_{\widetilde{\mathrm{Lip}}_1(E_i)} \leq \frac{1}{3} \|f\|_{\widetilde{\mathrm{Lip}}_1(E_i)} + \frac{|J_{ii}|}{R} \|f\|_\infty .$$

Thus, $\mathcal{L}_{\mathrm{ref,i},\overline{\omega_0}^{\mathrm{i}}}$ has a unique eigenvalue on the unit circle, 1, when it acts on $\widetilde{\mathrm{Lip}}_1(E_i)$, and any other eigenvalue is of modulus strictly smaller than 1. Once we make this operator act on $\mathrm{Lip}_1(E_i)$, the eigenvalue 1 disappears, and its spectral radius is strictly smaller than 1. Hence, there exists some $K \geq 1$ and $\lambda < 1$ such that:

$$\left\|\mathcal{L}^n_{\mathrm{ref,i},\overline{\omega_0}^{\mathrm{i}}}\right\|_{\mathrm{Lip}_1(E_i)\mapsto \mathrm{Lip}_1(E_i)} \leq K\lambda^n.$$

Since the system is translation-invariant, we can take the same constants K and λ for all operators $\mathcal{L}_{\mathrm{ref,i},\overline{\omega_0}^{\mathrm{i}}}$. They are also independent of the temperature. Some computations lead to the estimates:

$$c_{ij} \leq \frac{|J_{ij}|}{RT}, \quad j \neq i;$$
$$c_{ii} \leq \frac{|J_{ii}|}{R}.$$

Now, we can state a result for such Ising models:

Corollary 5.2 (Iterated Function System Ising model)**.**
Assume that $\sum_{i\in\mathbb{Z}^d} |J_{0i}|^\alpha$ *is finite for some* $\alpha < 1$*. Then, if the temperature* T *is high enough, the process has a unique stationary measure.*
Assume that $J_{0i} = 0$ *if* $|i|$ *is large enough. Then, if the temperature is high enough, the systems exhibits an exponential spatial decay of correlations.*

It is nothing more than an application of Theorem 3.6. Juste notice that, since the system is translation-invariant, we have $\varepsilon^* = (\varepsilon^t)^*$, and that the assumption of this theorem imply that both ε^* and $\varepsilon^*(\alpha)$ go to zero as T grows to infinity.

Remark 5.3.
Let C *be the Cantor set. Let us write* $C_i := 2C - 1$*, with the usual conventions about set multiplication and addition. Then, any stationary measure for the Ising model we defined here is supported by* $\bigotimes_{i\in\mathbb{Z}^d} C_i$*. This shows, in particular, that it would be impossible to get a convergence for the total variation norm starting from every probability measure. If* μ_i*, the marginal of the probability measure* μ *on* E_i*, has a density with respect to the Lebesgue measure, then so does* $(\mathcal{L}^n\mu)_i$ *for any* n*, so that the distance in total variation between* $\mathcal{L}^n\mu$ *and the stationary measure is always* 2*. Hence, it is paramount to work with softer norms.*

Acknowledgments
This article would never have been written if Gerhard Keller had not proposed this nice subject to work on, and provided many comments and advices during my year-long stay at Erlangen. My thanks also go to Viviane Baladi, who made the afore-mentioned stay possible.

References

[1] (MR0971099) M. F. Barnsley, S. G. Demko, J. H. Elton and J. S. Geronimo, *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities*, Annales de l'Institut Henri Poincaré, Probabilités et statistiques, **3** (1988), 367–394.

[2] (MR2283252) J. Dedecker, C. Prieur and P. Raynaud de Fitte, *Parametrized Kantorovich-Rubinstein theorem and application to the coupling of random variables*, Dependence in probability and statistics, Lecture Notes in Statistics, **187** (2006), 105–121.

[3] (MR0231434) R.L. Dobrushin, *Description of a random field by means of conditional probabilities and conditions for its regularity*, Theory of Probability and its Applications, **13** (1968) 201–229.

[4] (MR0298716) R.L. Dobrushin, *Definition of a system of random variables by means of conditional distributions*, Theory of Probability and its Applications, **15** (1970) 458–486.

[5] (MR2498672) M. Dyer, L.A. Goldberg and M. Jerrum, *Matrix norms and rapid mixing for spin systems*, The Annals of Applied Probability, **19** (2009) 71–107.

[6] (MR0539733) L. Gross, *Decay of correlations in classical lattice models at high temperature*, Communications in Mathematical Physics, **68** (1979) 9–27.

[7] (MR1129880) H. Hennion, *Sur un théorème spectral et son application aux noyaux lipchitziens*, (French) Proceedings of the American Mathematical Society, **118** (1993) 627–639.

[8] (MR1679080) G. Keller and C. Liverani, *Stability of the spectrum for transfer operators*, Annali della Scuola Normale Superiore di Pisa, **28** (1999) 141–152.

[9] (MR0676186) D. Klein, *Dobrushin uniqueness techniques and the decay of correlations in continuum statistical mechanics*, Communications in Mathematical Physics, **86** (1986) 227–246.

[10] (MR2822443) R.S. MacKay, *Robustness of Markov processes on large networks*, Journal of Difference Equations and Applications, **17** (2011) 1155–1167.

[11] (MR2459454) C. Villani, *Optimal transport. Old and new*, Springer-Verlag Berlin, 2009.