Ergodicity of some dynamics of DNA sequences

Mikael Falconnet[1], Nina Gantert[2] and Ellen Saada[3]

[1]LPO Gustave Eiffel, Bordeaux, France
[2]Technische Universität München, Fakultät für Mathematik, Germany
[3]CNRS, UMR 8145 laboratoire MAP5, Université de Paris – Campus Saint-Germain-des-Prés, France

To Guy Fayolle, on the occasion of his birthday

Abstract

We define interacting particle systems on configurations of the integer lattice (with values in some finite alphabet) by the superimposition of two dynamics: a substitution process with finite range rates, and a circular permutation mechanism (called "cut-and-paste") with possibly unbounded range. The model is motivated by the dynamics of DNA sequences: we consider an ergodic model for substitutions, the RN+YpR model ([BGP08]), with three particular cases, the models JC+`cpg`, T92+`cpg`, and RNc+YpR. We investigate whether they remain ergodic with the additional cut-and-paste mechanism, which models insertions and deletions of nucleotides. Using either duality or attractiveness techniques, we provide various sets of sufficient conditions, concerning only the substitution rates, for ergodicity of the superimposed process. They imply ergodicity of the models JC+`cpg`, T92+`cpg` as well as the attractive RNc+YpR, all with an additional cut-and-paste mechanism.

1 Introduction

Motivated by biological models for the evolution of DNA sequences, this paper contributes to the classical topic of ergodicity of interacting particle systems (see e.g. [Lig05]). Let us first give the biological context, then explain the interacting particle system it induces, and how we tackle its ergodicity. This last point corresponds to the intriguing question of the ergodicity for a superimposition of an ergodic particle system (here a generalized spin system) and of a non ergodic one (here cyclic permutations which generalize exclusion processes). The spin values belong to a finite alphabet of size larger than two, which adds a difficulty.

The biological set-up. The study of evolutionary relationships among living organisms has entered the genomic age in the past decades. To model the evolution of DNA sequences remains an important and difficult part of phylogenetic analysis. Generally, every method of phylogenetic reconstruction at the molecular level is based on a probabilistic model, for codons or for nucleotides. Jukes & Cantor, in [JC69], were the first to introduce a probabilistic model, the (JC) model, to study the changes in DNA sequences. In the (JC) model, DNA sequences are encoded as elements of $\mathscr{A}^N$, where the positive integer N stands for the number of nucleotides in one strand of the DNA molecule, and $\mathscr{A}$ for the nucleotide alphabet $\{\mathtt{a}, \mathtt{t}, \mathtt{c}, \mathtt{g}\}$, where the letters represent adenine, thymine, cytosine and guanine respectively. The (JC) model deals with the product of N independent identically distributed continuous-time Markov chains modelling single site nucleotide substitutions.

The substitution-rate matrix in the (JC) model is the simplest possible because all substitutions occur at the same rate. Since then, other nucleotide substitution processes have been introduced to refine this matrix (see for instance [Kim80, Fel81, HKY85, Tam92, TN93, Yan94]) until the generalized time reversible (GTR) model introduced in [Tav86] which is such that the Markov process is reversible, and with no more restriction on the structure of the matrix. In all these processes, the independence assumption on sites was kept. As a consequence, there exists a unique stationary probability measure for the process, which is product. This means for instance that in a long DNA sequence at equilibrium the frequency of a dinucleotide $x\mathrm{p}y$ should be the product of the x and y frequencies (where, for subsets X and Z of $\mathscr{A}$, $X\mathrm{p}Z$ is the collection of dinucleotides in $X \times Z$, and we write $x\mathrm{p}y$ instead of $\{x\}\mathrm{p}\{y\}$).

But this is actually not the case in some biological contexts. Indeed, since the studies of [JKK61] and [STK62], it is well known that the dinucleotide cpg is less frequently present in many mammals DNA than it would be expected from base composition. Support for the cpg deficiency to be related to DNA methylation was provided in [Bir80]: the substitution rate of cytosine by thymine is higher in methylated cpg than in other dinucleotides. Therefore, more realistic substitution models incorporating such neighboring effects have been introduced by [DG00] with their Tamura+cpg model. To evade the dependency between neighbors, in [DG00] there is an approximation for frequencies of trinucleotides to capture some features of the true model. Bérard et al. in [BGP08] extended the latter model to the RN+YpR model and assessed rigorously the effect of neighbor-dependent substitutions. There, DNA sequences are encoded as (doubly infinite) elements of $\mathscr{A}^{\mathbb{Z}}$ and their dynamics are studied through the techniques of interacting particle systems. The properties of the RN+YpR model have been used to infer phylogenetic distances in [Fal10] or cpg hypermutability rates in [BG12].

But substitutions are not the only way to alter DNA sequences. For example, one may add several extra nucleotides to a DNA sequence by insertions, or remove them by deletions. In the model of [TKF91], single sites are inserted or deleted with rates independent of their positions, and this is superimposed to independent substitutions. There, DNA sequences have a variable (but finite) length along time and are encoded as elements of $\bigcup_N \mathscr{A}^N$. We would like to do for neighbor-dependent substitution processes on DNA sequences what was done in [TKF91] for independent substitutions. The DNA sequences are viewed, as above, as elements of $\mathscr{A}^{\mathbb{Z}}$. To avoid mathematical difficulties due to the insertion-deletion mechanisms (an insertion of a single nucleotide induces a shift of the whole sequence and leads to infinite range interactions), we require that an insertion and a deletion occur at the same time. To that purpose, we introduce a mechanism that we call "cut-and-paste", whose name is inspired from the classification proposed by [Fin89] for transposable elements. The latter, discovered by [McC53], are a type of DNA that can move around within the genome and can be distinguished in two classes : class I is commonly called "copy-and-paste", and class II, "cut-and-paste". In our settings, we consider the simplest cut-and-paste mechanism, that is, the transfer of one nucleotide into the sequence as shown on Figure 1. There, the transfer of nucleotide `t` can be seen as its deletion and reinsertion three nucleotides further.

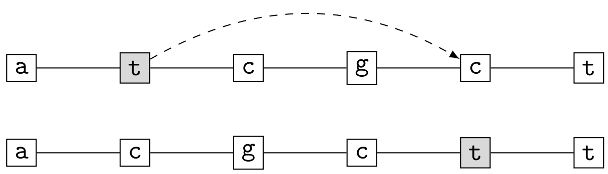

Figure 1: One example of the cut-and-paste of a nucleotide into the sequence

The modelling by a particle system, and the question of its ergodicity. We consider configurations that can take values in some finite alphabet on each site of the integer lattice, as detailed above; thus the alphabet will consist of the nucleotides `a`, `c`, `g` and `t`, but the questions we ask are not relying on this specific choice. On this configuration space, we superimpose two dynamics. The first one is given by a substitution process, and the second one by a cut-and-paste mechanism with a rate $\rho > 0$. The substitution process is quite general and includes for instance the stochastic Ising model with finite rate of dependence. The cut-and-paste process can have unbounded range and permutes the values of a certain interval of sites; it can be seen as a generalization of an exclusion process.

In this superimposition, the first dynamics is ergodic but not the second one. Note that ergodicity of our process, at least in the case of finite range permutations, would follow from the "Positive Rates Conjecture" which says that in dimension 1, a finite range interacting particle system with strictly positive transition rates is ergodic, see [Lig05], page 201. The status of this conjecture is not clear to us, see [Gra01]. It has been proved for attractive nearest-neighbor spin systems (a spin system is an interacting particle system with two states such that only one coordinate can change in each transition, see [Lig05], Chapter II) in [Gra82]. But, even if the permutations have finite range, our model does not fit into this frame, since the cut-and-paste mechanism changes several sites at the same time, and moreover we work with an alphabet of more than two values. There is a general sufficient condition for ergodicity, the so-called Dobrushin's "$M < \varepsilon$" condition (see [Dob71], or [Lig05], Chapter I). However, this condition is far from being necessary: For instance, it gives ergodicity of the one-dimensional nearest-neighbor stochastic Ising model only for high values of the temperature. Note that if a process does fit the "$M < \varepsilon$" condition, adding a small amount of "stirring" will not alter the validity of the latter, hence the superimposition with a stirring mechanism (which permutes the values on two neighboring sites) is ergodic for ρ small enough.

To derive ergodicity conditions for our superimposition of two dynamics, we use two powerful techniques: the first one is based on duality, and the second one on attractiveness and couplings. Our results involve only the parameters of the substitution process, thus they are valid for *any* value of ρ. First we use duality through a generalization of a coupling technique with branching processes due to [Fer90] to show that, in a certain range of parameters of the substitution process, the superimposition is ergodic. We then apply this general result to our generic example, the RN+YpR substitution model with an additional cut-and-paste mechanism. Second, in the spirit of [GS10, Bor11], we derive sufficient conditions on the substitution rates for attractiveness of the superimposition; assuming them, we obtain by different approaches sufficient conditions on theses rates for ergodicity. While our first results are general, the following ones concern the RN+YpR model with cut-and-paste mechanism. They imply that its particular cases, the models JC+`cpg`, T92+`cpg`, as well as the RNc+YpR model in case it is attractive, are ergodic when they are superimposed with a cut-and-paste mechanism.

The paper is organized as follows. In Section 2 we define the model and the examples, for which we state ergodicity results (Theorems 2.3 and 2.5). In Section 3, we give the set-up for generalized duality, then state general ergodicity results (Theorems 3.1 and 3.4), proved in Section 5; we then prove Theorem 2.3. In Section 4, we give the set-up for attractiveness, then state various results (some general, some for our examples) for ergodicity, proved in Section 6, and we prove Theorem 2.5.

2 Definitions and examples

In Section 2.1, we introduce our set-up: two types of interacting particle systems on $X = \mathscr{A}^{\mathbb{Z}}$, namely a substitution and a cut-and-paste processes, that we superimpose. We then define in Section 2.2 the examples that will illustrate our analysis along the paper. In Section 2.3, we state for them the ergodicity results that follow from our analysis of the superimposed process, done in the rest of the paper.

2.1 The particle system

For an analytic study of the construction and basic properties of these systems, we refer to the seminal book [Lig05], on which we rely for sufficient existence conditions of our dynamics. We will give in Section 5 a graphical construction of the latter, which will be the first step to prove ergodicity results through duality.

2.1.1 Substitution process

In such a process, only one coordinate changes in each transition. The transition mechanism is specified by a non-negative function $c(\cdot)$ defined on $\mathbb{Z} \times \mathscr{A} \times X$. For $\eta \in X$, $x \in \mathbb{Z}$ and $a \in \mathscr{A}$, $c(x, a, \eta)$ represents the rate at which the coordinate $\eta(x)$ flips to a when the system is in state η, that is the rate at which η changes to η_a^x defined by

$$\eta_a^x(z) = \begin{cases} \eta(z) & \text{if} \quad z \neq x, \\ a & \text{if} \quad z = x. \end{cases}$$

We assume that the rates $c(x, a, \cdot)$ are translation invariant, i.e. $c(x, a, \eta) = c(0, a, \tau_x \eta)$ where τ_x denotes the shift by x on X (given by $\tau_x \eta(y) = \eta(x + y)$, for all $y \in \mathbb{Z}$). From now on, we write $c(a, \eta)$ for $c(0, a, \eta)$. We further assume that, for any target a, the function $c(a, \cdot)$ depends on $\eta \in X$ only through a finite set $S(a) \subset \mathbb{Z}$ depending on a.

The pregenerator $\mathscr{L}_1$ of the substitution process is defined on a cylinder function f on X (that is, a function depending on a finite number of coordinates) by

$$\mathscr{L}_1 f(\eta) = \sum_{x \in \mathbb{Z}} \sum_{a \in \mathscr{A}} c(a, \tau_x \eta) \big[f(\eta_a^x) - f(\eta) \big]. \tag{2.1}$$

Let

$$m = \inf\{c(a, \eta) \,:\, a \in \mathscr{A}, \eta \in X\}, \tag{2.2}$$
$$K = \sup\{c(a, \eta) \,:\, a \in \mathscr{A}, \eta \in X\}, \tag{2.3}$$
$$s = \sup\{|S(a)| \,:\, a \in \mathscr{A}\}, \tag{2.4}$$

where $|S|$ denotes the cardinality of the set S. Because the alphabet $\mathscr{A}$ is finite, we have that

$$s < +\infty, \quad K < +\infty. \tag{2.5}$$

This is a sufficient condition for the existence of a Markov process with pregenerator $\mathscr{L}_1$ (see Theorem I.3.9 in [Lig05]). For Theorem 3.1 and for our examples, we will assume

$$m > 0, \tag{2.6}$$

which is a standard assumption for ergodicity results (see e.g. [Gra01], [Gra82]).

2.1.2 Cut-and-paste process

In this process, not only one coordinate changes in each transition (as it can be seen on Figure 1 where the transfer of nucleotide `t` induces a change of four coordinates because of the shift to the left for the segment `cgc`). The transition mechanisms are circular

permutations $\sigma_{x,y}$ (for $x \neq y \in \mathbb{Z}$) of finitely many sites of $\mathbb{Z}$ and are specified by a transition probability matrix p on $\mathbb{Z} \times \mathbb{Z}$.
For $\eta \in X$ and a pair of sites $x \neq y$, let $\sigma_{x,y}(\eta)$ be the configuration defined by

$$\sigma_{x,y}(\eta)(z) = \eta\big[\sigma_{x,y}^{-1}(z)\big], \quad \forall z \in \mathbb{Z},$$

where $\sigma_{x,y}$ is defined for any $x < y$ as

$$\sigma_{x,y}(z) = \begin{cases} z & \text{if} \quad z \notin \{x, x+1, \dots, y\}, \\ y & \text{if} \quad z = x, \\ z-1 & \text{if} \quad x < z \leqslant y, \end{cases}$$

and for any $x > y$ as

$$\sigma_{x,y}(z) = \begin{cases} z & \text{if} \quad z \notin \{y, y+1, \dots, x\}, \\ y & \text{if} \quad z = x, \\ z+1 & \text{if} \quad y \leqslant z < x. \end{cases}$$

The case $x < y$ corresponds to a shift to the left of the coordinates from site $x+1$ to site y as it can be seen on the top of Figure 2, whereas $x > y$ corresponds to a shift to the right of the coordinates from y to $x-1$ as it can be seen on the bottom of Figure 2.
The rate at which η changes to $\sigma_{x,y}(\eta)$ is $p(x,y)$. We assume that p is translation

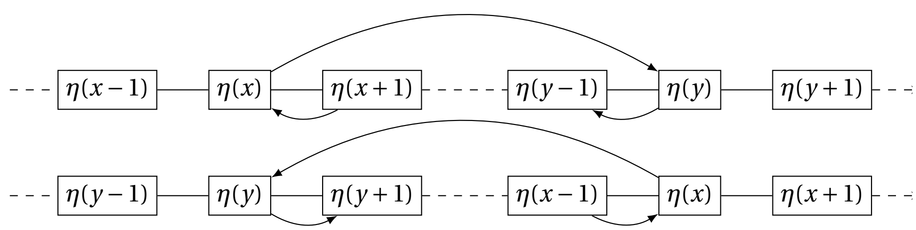

Figure 2: Transfer of the coordinate $\eta(x)$ from site x to site y when $x < y$ (top panel) or when $x > y$ (bottom panel)

invariant on $\mathbb{Z}$, that is, $p(x,y) = p(0, y-x)$ for all $x, y \in \mathbb{Z}$. We do not assume that p is symmetric, hence the "rate of transfer" $p(x,y)$ at which the coordinate $\eta(x)$ is transferred to site y might depend not only on the distance $|y-x|$ but also on the direction of the transfer.
The pregenerator $\mathscr{L}_2$ of a cut-and-paste process is defined on a cylinder function f on X by

$$\mathscr{L}_2 f(\eta) = \sum_{x,y \in \mathbb{Z}} p(x,y)\big[f(\sigma_{x,y}(\eta)) - f(\eta)\big]. \tag{2.7}$$

A sufficient condition for the existence of the cut-and-paste process, that we assume from now on, is (see [Lig05], [AFS04])

$$\sum |x| p(0,x) < \infty. \tag{2.8}$$

Remark 2.1. *(i) We mentioned that p is not necessarily symmetric because [Kov05] studies invariant measures of particle systems that interact via finite range permutations, which is a more general system than a cut-and-paste process, but with the restriction that the permutation σ has the same rate of occurrence than σ^{-1}, which is equivalent in our case to the symmetry of p.*
Ergodicity of cellular automata that are superimpositions of Glauber dynamics and permutations is studied in [Fer91].
(ii) For simplicity, we stick here to circular permutations. As it will become clear from the proofs, we could also consider more general permutations of $\eta(x), \eta(x+1), \dots, \eta(y)$ or $\eta(y), \eta(y+1), \dots \eta(x)$, respectively, resulting in different biological interpretations than the one we gave in the introduction.
(iii) The cut-and-paste process generalizes a stirring process, which is such that $p(.,.)$ is symmetric and nearest-neighbor, so that only two coordinates change in a transition.

2.1.3 The superimposition

Fix a constant $\rho > 0$ and define the pregenerator $\mathscr{L}$ on a cylinder function f on X as

$$\mathscr{L} f = \mathscr{L}_1 f + \rho \mathscr{L}_2 f. \tag{2.9}$$

We are interested in the ergodic properties of the interacting particle system with pregenerator $\mathscr{L}$, that is, the superimposition of a substitution and a cut-and-paste process. This superimposition is well defined by assumptions (2.5) and (2.8). We denote by $\mathscr{P}(X)$ the set of probability measures on X, by $\mathscr{S}$ the set of translation invariant probability measures on X. For the superimposition we denote by $\mathscr{I}$ the set of invariant probability measures.

Recall that (Definition I.1.9 in [Lig05]) ergodicity of a Markov process with values in $X = \mathscr{A}^{\mathbb{Z}}$ means that there is exactly *one* invariant probability measure, denoted e.g. by μ, and for each starting point, the law of the process converges to this invariant probability measure.

This ergodic process is moreover *exponentially ergodic* ([Fer90], page 1526) if for any bounded cylinder function f on X, there exist positive constants $a_1 = a_1(f), a_2$ such that for any initial probability measure ν, for any $t \geqslant 0$, we have

$$\left| \int f d(\nu S(t)) - \int f d\mu \right| < a_1 \mathrm{e}^{-a_2 t}. \tag{2.10}$$

The simplest ergodic substitution processes are the independent ones. As expected, the superimposition of a cut-and-paste mechanism does not affect the ergodic properties and the invariant probability measure of such processes. In this case, we do not need assumption (2.6), and moreover $s = 1$ (see (2.4)).

Lemma 2.2. *Assume that the rate function $c(\cdot)$ can be written as*

$$c(x, a, \eta) = q(\eta(x), a), \tag{2.11}$$

where $Q = (q(a,b))_{a,b\in\mathscr{A}}$ is the infinitesimal generator of an irreducible continuous-time Markov chain on $\mathscr{A}$ with unique invariant probability measure π. Then the process $(\eta_t)_{t\geqslant 0}$ with generator $\mathscr{L}$ given by (2.9) is ergodic and its unique invariant probability measure is the product measure $\pi^{\otimes\mathbb{Z}}$.

We omit the simple proof of this lemma. The examples of substitution processes we will work with from now on are not independent ones, they will satisfy $s > 1$.

2.2 Examples of substitution models

In this section we first define our generic example, the RN+YpR model, that was introduced and studied in [BGP08], to which we refer for biological motivation. According to the values of its parameters, this mathematical model contains many known biological situations: We define more and more particular cases of it, the models RNc+YpR, T92+`cpg`, and JC+`cpg`.

2.2.1 The RN+YpR model

First, RN stands for Rzhetsky-Nei [RN95] and means that the 4×4 matrix of substitution rates which characterize the independent evolution of the sites must satisfy 4 equalities, summarized as follows: for every pair of neighboring nucleotides a and $b \neq a$, the substitution rate from a to b may depend on a but only through the fact that a is a purine (`a` or `g`, symbol R) or a pyrimidine (`c` or `t`, symbol Y). For instance, the substitution rates from `c` to `a` and from `t` to `a` must coincide, as well as from `a` and from `g` to `c`, from `c` and from `t` to `g`, and finally from `a` and from `g` to `t`. The 4 remaining rates, corresponding to purine-purine and to pyrimidine-pyrimidine substitutions, are free. The matrix of substitution rates is given by

$$\begin{array}{c} \\ \mathtt{a} \\ \mathtt{t} \\ \mathtt{c} \\ \mathtt{g} \end{array}\begin{array}{c} \begin{array}{cccc} \mathtt{a} & \mathtt{t} & \mathtt{c} & \mathtt{g} \end{array} \\ \begin{pmatrix} \cdot & v_{\mathtt{t}} & v_{\mathtt{c}} & w_{\mathtt{g}} \\ v_{\mathtt{a}} & \cdot & w_{\mathtt{c}} & v_{\mathtt{g}} \\ v_{\mathtt{a}} & w_{\mathtt{t}} & \cdot & v_{\mathtt{g}} \\ w_{\mathtt{a}} & v_{\mathtt{t}} & v_{\mathtt{c}} & \cdot \end{pmatrix} \end{array}, \quad \text{with} \quad 0 \leqslant v_a \leqslant w_a \quad \text{for all} \quad a \in \mathscr{A}.$$

Second, the influence mechanism is called YpR, which stands for the fact that one allows any specific substitution rate between any two YpR dinucleotides (`cg`, `ca`, `tg` and `ta`) for a total of 8 independent parameters. Hence, every dinucleotide `cg` moves to `ca` at rate $r^{\mathtt{c}}_{\mathtt{a}}$ and to `tg` at rate $r^{\mathtt{g}}_{\mathtt{t}}$; every dinucleotide `ta` moves to `ca` at rate $r^{\mathtt{a}}_{\mathtt{c}}$ and to `tg` at rate $r^{\mathtt{t}}_{\mathtt{g}}$; every dinucleotide `ca` moves to `cg` at rate $r^{\mathtt{c}}_{\mathtt{g}}$ and to `ta` at rate $r^{\mathtt{a}}_{\mathtt{t}}$; every dinucleotide `tg` moves to `cg` at rate $r^{\mathtt{g}}_{\mathtt{c}}$ and to `ta` at rate $r^{\mathtt{t}}_{\mathtt{a}}$. We have

$$S(\mathtt{a}) = \{-1, 0\} \quad \text{and} \quad c(\mathtt{a}, \eta) = \begin{cases} v_{\mathtt{a}} & \text{if } \eta(0) \in \{\mathtt{c}, \mathtt{t}\} = Y, \\ w_{\mathtt{a}} & \text{if } \eta(0) = \mathtt{g} \text{ and } \eta(-1) \notin \{\mathtt{c}, \mathtt{t}\}, \\ w_{\mathtt{a}} + r^{a}_{\mathtt{a}} & \text{if } \eta(0) = \mathtt{g} \text{ and } \eta(-1) = a \in \{\mathtt{c}, \mathtt{t}\}, \end{cases}$$

$$S(\mathtt{t})=\{0,1\}\quad\text{and}\quad c(\mathtt{t},\eta)=\begin{cases} v_{\mathtt{t}} & \text{if } \eta(0)\in\{\mathtt{a},\mathtt{g}\}=R,\\ w_{\mathtt{t}} & \text{if } \eta(0)=\mathtt{c} \text{ and } \eta(1)\notin\{\mathtt{a},\mathtt{g}\},\\ w_{\mathtt{t}}+r^{a}_{\mathtt{t}} & \text{if } \eta(0)=\mathtt{c} \text{ and } \eta(1)=a\in\{\mathtt{a},\mathtt{g}\},\end{cases}$$

$$S(\mathtt{c})=\{0,1\}\quad\text{and}\quad c(\mathtt{c},\eta)=\begin{cases} v_{\mathtt{c}} & \text{if } \eta(0)\in\{\mathtt{a},\mathtt{g}\}=R,\\ w_{\mathtt{c}} & \text{if } \eta(0)=\mathtt{t} \text{ and } \eta(1)\notin\{\mathtt{a},\mathtt{g}\},\\ w_{\mathtt{c}}+r^{a}_{\mathtt{c}} & \text{if } \eta(0)=\mathtt{t} \text{ and } \eta(1)=a\in\{\mathtt{a},\mathtt{g}\},\end{cases}$$

$$S(\mathtt{g})=\{-1,0\}\quad\text{and}\quad c(\mathtt{g},\eta)=\begin{cases} v_{\mathtt{g}} & \text{if } \eta(0)\in\{\mathtt{c},\mathtt{t}\}=Y,\\ w_{\mathtt{g}} & \text{if } \eta(0)=\mathtt{a} \text{ and } \eta(-1)\notin\{\mathtt{c},\mathtt{t}\},\\ w_{\mathtt{g}}+r^{a}_{\mathtt{g}} & \text{if } \eta(0)=\mathtt{a} \text{ and } \eta(-1)=a\in\{\mathtt{c},\mathtt{t}\}.\end{cases}$$

$$K=\max\left\{w_a+r^b_a, w_b+r^a_b\;:\;a\in Y, b\in R\right\},\quad m=\min_{a\in\mathscr{A}} v_a,\quad\text{and}\quad s=2. \tag{2.12}$$

We assume from now on that (2.6) is satisfied, that is $\min_{a\in\mathscr{A}} v_a>0$.

2.2.2 The RNc+YpR model

In this model, see [BGP08] page 79, the substitution rates respect the “strand complementarity of nucleotides”, so that the rates of YpR substitutions from `cg` to `ca` and to `tg` coincide, from `ta` to `ca` and to `tg` coincide, from `ca` and from `tg` to `cg` coincide, from `ca` and from `tg` to `ta` coincide. Therefore

$$w_{\mathtt{a}}=w_{\mathrm{W}};\ v_{\mathtt{a}}=v_{\mathrm{W}}\quad;\quad r^{\mathtt{t}}_{\mathtt{a}}=r_{\mathrm{U}};\ r^{\mathtt{c}}_{\mathtt{a}}=r_{\mathrm{W}}, \tag{2.13}$$
$$w_{\mathtt{t}}=w_{\mathrm{W}};\ v_{\mathtt{t}}=v_{\mathrm{W}}\quad;\quad r^{\mathtt{a}}_{\mathtt{t}}=r_{\mathrm{U}};\ r^{\mathtt{g}}_{\mathtt{t}}=r_{\mathrm{W}}, \tag{2.14}$$
$$w_{\mathtt{c}}=w_{\mathrm{S}};\ v_{\mathtt{c}}=v_{\mathrm{S}}\quad;\quad r^{\mathtt{a}}_{\mathtt{c}}=r_{\mathrm{S}};\ r^{\mathtt{g}}_{\mathtt{c}}=r_{\mathrm{V}}, \tag{2.15}$$
$$w_{\mathtt{g}}=w_{\mathrm{S}};\ v_{\mathtt{g}}=v_{\mathrm{S}}\quad;\quad r^{\mathtt{t}}_{\mathtt{g}}=r_{\mathrm{S}};\ r^{\mathtt{c}}_{\mathtt{g}}=r_{\mathrm{V}}. \tag{2.16}$$

2.2.3 The T92+cpg model

This model adds neighboring effects to the classical T92 model developed by Tamura in [Tam92], which consisted in an independent evolution of the sites. The values of the above rates (2.13)–(2.16) for the T92+`cpg` model are, for some $\theta\in[0,1]$,

$$w_{\mathtt{a}}=(1-\theta)w;\ v_{\mathtt{a}}=(1-\theta)v\quad;\quad r^{\mathtt{t}}_{\mathtt{a}}=0;\ r^{\mathtt{c}}_{\mathtt{a}}=r, \tag{2.17}$$
$$w_{\mathtt{t}}=(1-\theta)w;\ v_{\mathtt{t}}=(1-\theta)v\quad;\quad r^{\mathtt{a}}_{\mathtt{t}}=0;\ r^{\mathtt{g}}_{\mathtt{t}}=r, \tag{2.18}$$
$$w_{\mathtt{c}}=\theta w;\ v_{\mathtt{c}}=\theta v\quad;\quad r^{\mathtt{a}}_{\mathtt{c}}=r^{\mathtt{g}}_{\mathtt{c}}=0, \tag{2.19}$$
$$w_{\mathtt{g}}=\theta w;\ v_{\mathtt{g}}=\theta v\quad;\quad r^{\mathtt{t}}_{\mathtt{g}}=r^{\mathtt{c}}_{\mathtt{g}}=0. \tag{2.20}$$

2.2.4 The JC+cpg model

Again this model adds neighboring effects to the JC model (see Section 1). The values of the above rates for the JC+`cpg` model correspond to the doubled values of the ones of the T92+$\texttt{cpg}$ model for $\theta = 1/2$:

$$w_{\texttt{a}} = v_{\texttt{a}} = v \quad ; \quad r^{\texttt{t}}_{\texttt{a}} = 0;\ r^{\texttt{c}}_{\texttt{a}} = r, \tag{2.21}$$
$$w_{\texttt{t}} = v_{\texttt{t}} = v \quad ; \quad r^{\texttt{a}}_{\texttt{t}} = 0;\ r^{\texttt{g}}_{\texttt{t}} = r, \tag{2.22}$$
$$w_{\texttt{c}} = v_{\texttt{c}} = v \quad ; \quad r^{\texttt{a}}_{\texttt{c}} = r^{\texttt{g}}_{\texttt{c}} = 0, \tag{2.23}$$
$$w_{\texttt{g}} = v_{\texttt{g}} = v \quad ; \quad r^{\texttt{t}}_{\texttt{g}} = r^{\texttt{c}}_{\texttt{g}} = 0. \tag{2.24}$$

2.3 Ergodicity results for these substitution models with additional cut-and-paste mechanism

Assuming that $\min_{a\in\mathscr{A}} v_a > 0$ (that is, (2.6)), [BGP08] proved that the RN+YpR model, and as a consequence the RNc+YpR, T92+$\texttt{cpg}$ and JC+$\texttt{cpg}$ models are ergodic for all substitution rates. As pointed out in [BGP08, Theorem 6], considering only the evolution of Y and R instead of the one of the four elements of $\mathscr{A}$ gives that the (only) invariant probability measure is a product measure.
Our main results for these examples are the following.

Theorem 2.3. *For any $\rho > 0$, the RN+YpR model with cut-and-paste mechanism is exponentially ergodic (recall (2.10)) if*

$$\min_{a\in\mathscr{A}} v_a > 0 \quad \textit{and} \quad \max\left(\mathscr{Y} \cup \mathscr{R}\right) < \sum_{a\in\mathscr{A}} v_a, \tag{2.25}$$

where

$$\mathscr{Y} \;=\; \{r^{\texttt{a}}_a \vee r^{\texttt{g}}_a - r^{\texttt{a}}_a \wedge r^{\texttt{g}}_a,\ r^{\texttt{a}}_a \wedge r^{\texttt{g}}_a \,:\, a \in Y\}, \tag{2.26}$$
$$\mathscr{R} \;=\; \{r^{\texttt{c}}_b \vee r^{\texttt{t}}_b - r^{\texttt{c}}_b \wedge r^{\texttt{t}}_b, r^{\texttt{c}}_b \wedge r^{\texttt{t}}_b \,:\, b \in R\}. \tag{2.27}$$

If the second inequality in (2.25) is an equality, then the RN+YpR model with cut-and-paste mechanism is ergodic.

We will prove Theorem 2.3 in Section 3.3, using the duality technique introduced and developed in Subsections 3.1 and 3.2 for general substitution processes with cut-and-paste mechanism. The next corollary is a direct application of Theorem 2.3.

Corollary 2.4. *Assume that the rates of the RNc+YpR model (defined in (2.13)–(2.16)) satisfy*

$$\max\left(r_{\mathrm{U}} \vee r_{\mathrm{W}} - r_{\mathrm{U}} \wedge r_{\mathrm{W}},\ r_{\mathrm{U}} \wedge r_{\mathrm{W}},\ r_{\mathrm{S}} \vee r_{\mathrm{V}} - r_{\mathrm{S}} \wedge r_{\mathrm{V}},\ r_{\mathrm{S}} \wedge r_{\mathrm{V}}\right) \leqslant 2(v_{\mathrm{S}} + v_{\mathrm{W}}). \tag{2.28}$$

Then, for any $\rho > 0$, the RNc+YpR model with cut-and-paste mechanism is ergodic.

The attractiveness and coupling techniques will be introduced and developed in Subsection 4.1 for general substitution processes with cut-and-paste mechanism, and specialized in Subsection 4.2 to the RN+YpR model. We will prove Theorem 2.5 in Section 4.2.

Theorem 2.5. *For any $\rho > 0$, we have the following.*

(i) The T92+`cpg` *model, and as a consequence the JC+*`cpg` *model, both with cut-and-paste mechanism, are ergodic for all substitution rates.*

(ii) Assume that the rates of the RNc+YpR model (defined in (2.13)–(2.16)) satisfy attractiveness conditions, namely

$$\text{either} \qquad r_{\mathrm{U}} \leqslant r_{\mathrm{W}};\ r_{\mathrm{S}} = r_{\mathrm{V}} = 0, \tag{2.29}$$

$$\text{or} \qquad r_{\mathrm{S}} \leqslant r_{\mathrm{V}};\ r_{\mathrm{U}} = r_{\mathrm{W}} = 0. \tag{2.30}$$

Then, the RNc+YpR model with cut-and-paste mechanism is ergodic.

More precisely, we will derive various sets of conditions for the substitution rates of the processes which imply ergodicity (Propositions 4.1, 4.9, 4.12), and we will apply them to derive Theorem 2.5.

3 Ergodicity through generalized duality

The starting point of this approach is a graphical construction of the dynamics, using a Harris representation [Har72], done in Section 5.1. To state our results in Section 3.2, we have to introduce the necessary notation in Section 3.1.

3.1 Set-up

In the pregenerator $\mathscr{L}_1$ defined by (2.1), for each $a \in \mathscr{A}$, write the substitution rate $c(a, \eta)$ depending on the finite set $S(a) \subset \mathbb{Z}$ as

$$c(a, \eta) = \sum_{j \in J(a)} \lambda_j(a) \mathbf{1}_{\{\eta \in A_j(a)\}}, \tag{3.1}$$

where $\lambda_{j+1}(a) \geqslant \lambda_j(a)$ and $A_j(a)$ are cylinder sets of X depending on $S_j(a) \subseteq S(a)$ such that the family $\{A_j(a)\}_{j \in J(a)}$ is a partition of X. Thus,

$$\sum_{j \in J(a)} \mathbf{1}_{\{\eta \in A_j(a)\}} = 1. \tag{3.2}$$

By convention, the first label in the set $J(a) \subset \mathbb{N}$ is 0. The number of elements of $J(a)$ is uniformly bounded by $s^{|\mathscr{A}|}$, where s was defined by (2.4). Indeed, since the rate $c(a, \eta)$ depends at most on s sites, there are at most $s^{|\mathscr{A}|}$ different cylinder sets $A_j(a)$. We set

$$\begin{aligned} \overline{\lambda}_j(a) &= \lambda_j(a) - \lambda_{j-1}(a) \quad \text{for } j \in J(a) \setminus \{0\}, \text{ and } \quad \overline{\lambda}_0(a) = \lambda_0(a), && (3.3)\\ \lambda(a) &= \max\left\{\lambda_j(a) \,:\, j \in J(a)\right\} - \lambda_0(a) = \lambda_{|J(a)|-1} - \lambda_0(a), && (3.4) \end{aligned}$$

and

$$\overline{\lambda}_0 = \sum_{a\in\mathscr{A}} \overline{\lambda}_0(a). \tag{3.5}$$

We first assume (2.6); then the latter quantity is positive since $\overline{\lambda}_0(a) \geqslant m$. Finally, set

$$\overline{\lambda} = \max_{a\in\mathscr{A}} \max_{j\in J(a)\setminus\{0\}} \overline{\lambda}_j(a). \tag{3.6}$$

3.2 Results

Theorem 3.1. *Assume* (2.6) *and that*

$$(s-1)\overline{\lambda} < \overline{\lambda}_0. \tag{3.7}$$

Then, for any $\rho > 0$, the process $(\eta_t)_{t\geqslant 0}$ with generator $\mathscr{L}$ given by (2.9) *is exponentially ergodic. If $(s-1)\overline{\lambda} = \overline{\lambda}_0$ then the process is ergodic.*

Remark 3.2. *(i) A stronger condition for exponential ergodicity than* (3.7) *is that m, K, s (given by* (2.2)–(2.4)*) satisfy*

$$m > 0 \quad \textit{and} \quad (s-1)(K-m) < |\mathscr{A}|m. \tag{3.8}$$

This last inequality is the natural extension of the one in [Fer90, Theorem 2.1], done for a two-letter alphabet.
(ii) Condition (3.7) *is a priori non trivial if $s > 1$ and $|J(a)| > 1$ for some $a \in \mathscr{A}$.*

Refinement of Theorem 3.1. Assume that $\mathscr{L}_1$ can be decomposed as a sum of several generators

$$\mathscr{L}_1 = \sum_{i=1}^{d} \mathscr{L}_1^{(i)}, \quad \text{with} \quad \mathscr{L}_1^{(i)} f(\eta) = \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} c^{(i)}(a, \tau_x\eta)\big[f(\eta_a^x) - f(\eta)\big]. \tag{3.9}$$

For any i in $\{1, \dots, d\}$, define $m^{(i)}$, $K^{(i)}$, $s^{(i)}$ as in (2.2)–(2.4), replacing $c(\cdot,\cdot)$ by $c^{(i)}(\cdot,\cdot)$, defined through $J^{(i)}(a)$, $\lambda_j^{(i)}(a)$, $A_j^{(i)}(a)$ as in (3.1). In particular we have

$$\sum_{j\in J^{(i)}(a)} \mathbf{1}_{\{\eta\in A_j^{(i)}(a)\}} = 1. \tag{3.10}$$

Remark 3.3. *Such a decomposition of the generator might be useful if, for some i in $\{1, \dots, d\}$, the rate $c^{(i)}(\cdot,\cdot)$ depends on less sites than the original rate $c(\cdot,\cdot)$, that is, $s^{(i)} < s$, or if $s^{(i)} = 1$. See Theorem 3.4 below.*

As in (3.3), we set, for i in $\{1,\dots,d\}$,

$$\begin{aligned}\overline{\lambda}_j^{(i)}(a) &= \lambda_j^{(i)}(a)-\lambda_{j-1}^{(i)}(a) \quad \text{for } j\in J^{(i)}(a)\setminus\{0\}, \text{ and } \quad \overline{\lambda}_0^{(i)}(a)=\lambda_0^{(i)}(a), &(3.11)\\ \lambda^{(i)}(a) &= \max\left\{\lambda_j^{(i)}(a)\,:\, j\in J^{(i)}(a)\right\}-\lambda_0^{(i)}(a)=\lambda_{|J^{(i)}(a)|-1}^{(i)}-\lambda_0^{(i)}(a). &(3.12)\end{aligned}$$

We replace assumption (2.6) by

$$\sum_{i=1}^d m^{(i)}>0, \tag{3.13}$$

which can be true whereas for some i, $m^{(i)}=0$. According to (3.5), we would now define $\overline{\lambda}_0^{(i)}=\sum_{a\in\mathscr{A}}\overline{\lambda}_0^{(i)}(a)$, but this quantity could be zero for some i in $\{1,\dots,d\}$. Therefore, the right quantities to consider are

$$\overline{\lambda}_{0,d}(a)=\sum_{i=1}^d \overline{\lambda}_0^{(i)}(a) \quad \text{and} \quad \overline{\lambda}_{0,d}=\sum_{a\in\mathscr{A}}\overline{\lambda}_{0,d}(a), \tag{3.14}$$

which are positive by (3.13) since $\overline{\lambda}_{0,d}(a)\geqslant\sum_{i=1}^d m^{(i)}$.

Finally, we set

$$\overline{\lambda}^{(i)}=\max_{a\in\mathscr{A}}\max_{j\in J^{(i)}(a)\setminus\{0\}}\overline{\lambda}_j^{(i)}(a). \tag{3.15}$$

Then we have

Theorem 3.4. *Assume* (3.13) *and that*

$$\sum_{i=1}^d (s^{(i)}-1)\overline{\lambda}^{(i)}<\overline{\lambda}_{0,d}. \tag{3.16}$$

Then, for any $\rho>0$, the process $(\eta_t)_{t\geqslant 0}$ with generator $\mathscr{L}$ given by (2.9) *is exponentially ergodic. If $\sum_{i=1}^d (s^{(i)}-1)\overline{\lambda}^{(i)}=\overline{\lambda}_{0,d}$ then the process is ergodic.*

Remark 3.5. *(i) A stronger condition for exponential ergodicity than* (3.16) *is that $(K^{(i)},s^{(i)},m^{(i)})_{i=1}^d$ satisfy*

$$\sum_{i=1}^d m^{(i)}>0 \quad \textit{and} \quad \sum_{i=1}^d (s^{(i)}-1)(K^{(i)}-m^{(i)})<|\mathscr{A}|\sum_{i=1}^d m^{(i)}. \tag{3.17}$$

As in Remark 3.2, this last inequality is a natural extension of the one in [Fer90, Theorem 2.2].
(ii) It is shown in [Fer90] in the case of a two-letter alphabet that (3.16) *improves, even without stirring, the usual "$M<\varepsilon$" condition for ergodicity.*

Table 1: Values of the quantities introduced in Section 2 for the RN+YpR model

a	$S(a)$	$j \in J(a)$	$S_j(a)$	$A_j(a)$	$\lambda_j(a)$	$\overline{\lambda}_j(a)$
$\mathtt{a}$	$\{-1,0\}$	0	$\{0\}$	$\{\eta : \eta(0) \in Y\}$	$v_{\mathtt{a}}$	$v_{\mathtt{a}}$
		1	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = b\mathtt{g}, b \in R\}$	$w_{\mathtt{a}}$	$w_{\mathtt{a}} - v_{\mathtt{a}}$
		2	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = d\mathtt{g}, d \in Y, r^d_{\mathtt{a}} = r^{\mathtt{c}}_{\mathtt{a}} \wedge r^{\mathtt{t}}_{\mathtt{a}}\}$	$w_{\mathtt{a}} + r^d_{\mathtt{a}}$	$r^d_{\mathtt{a}}$
		3	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = e\mathtt{g}, e \in Y, r^e_{\mathtt{a}} = r^{\mathtt{c}}_{\mathtt{a}} \vee r^{\mathtt{t}}_{\mathtt{a}}\}$	$w_{\mathtt{a}} + r^e_{\mathtt{a}}$	$r^e_{\mathtt{a}} - r^d_{\mathtt{a}}$
$\mathtt{t}$	$\{0,1\}$	0	$\{0\}$	$\{\eta : \eta(0) \in R\}$	$v_{\mathtt{t}}$	$v_{\mathtt{t}}$
		1	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{c}b, b \in Y\}$	$w_{\mathtt{t}}$	$w_{\mathtt{t}} - v_{\mathtt{t}}$
		2	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{c}d, d \in R, r^d_{\mathtt{t}} = r^{\mathtt{a}}_{\mathtt{t}} \wedge r^{\mathtt{g}}_{\mathtt{t}}\}$	$w_{\mathtt{t}} + r^d_{\mathtt{t}}$	$r^d_{\mathtt{t}}$
		3	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{c}e, e \in R, r^e_{\mathtt{t}} = r^{\mathtt{a}}_{\mathtt{t}} \vee r^{\mathtt{g}}_{\mathtt{t}}\}$	$w_{\mathtt{t}} + r^e_{\mathtt{t}}$	$r^e_{\mathtt{t}} - r^d_{\mathtt{t}}$
$\mathtt{c}$	$\{0,1\}$	0	$\{0\}$	$\{\eta : \eta(0) \in R\}$	$v_{\mathtt{c}}$	$v_{\mathtt{c}}$
		1	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{t}b, b \in Y\}$	$w_{\mathtt{c}}$	$w_{\mathtt{c}} - v_{\mathtt{c}}$
		2	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{t}d, d \in R, r^d_{\mathtt{c}} = r^{\mathtt{a}}_{\mathtt{c}} \wedge r^{\mathtt{g}}_{\mathtt{c}}\}$	$w_{\mathtt{c}} + r^d_{\mathtt{c}}$	$r^d_{\mathtt{c}}$
		3	$\{0,1\}$	$\{\eta : \eta(0)\eta(1) = \mathtt{t}e, e \in R, r^e_{\mathtt{c}} = r^{\mathtt{a}}_{\mathtt{c}} \vee r^{\mathtt{g}}_{\mathtt{c}}\}$	$w_{\mathtt{c}} + r^e_{\mathtt{c}}$	$r^e_{\mathtt{c}} - r^d_{\mathtt{c}}$
$\mathtt{g}$	$\{-1,0\}$	0	$\{0\}$	$\{\eta : \eta(0) \in Y\}$	$v_{\mathtt{g}}$	$v_{\mathtt{g}}$
		1	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = b\mathtt{a}, b \in R\}$	$w_{\mathtt{g}}$	$w_{\mathtt{g}} - v_{\mathtt{g}}$
		2	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = d\mathtt{a}, d \in Y, r^d_{\mathtt{g}} = r^{\mathtt{c}}_{\mathtt{g}} \wedge r^{\mathtt{t}}_{\mathtt{g}}\}$	$w_{\mathtt{g}} + r^d_{\mathtt{g}}$	$r^d_{\mathtt{g}}$
		3	$\{-1,0\}$	$\{\eta : \eta(-1)\eta(0) = e\mathtt{a}, e \in Y, r^e_{\mathtt{g}} = r^{\mathtt{c}}_{\mathtt{g}} \vee r^{\mathtt{t}}_{\mathtt{g}}\}$	$w_{\mathtt{g}} + r^e_{\mathtt{g}}$	$r^e_{\mathtt{g}} - r^d_{\mathtt{g}}$
$s = 2,\quad \overline{\lambda} = \max\left(\{w_a - v_a : a \in \mathscr{A}\} \cup \mathscr{Y} \cup \mathscr{R}\right),\quad$ and $\quad \overline{\lambda}_0 = \sum_{a\in\mathscr{A}} v_a$						

3.3 Application to the RN+YpR model with cut-and-paste mechanism

Proof. (of Theorem 2.3). This is an application of Theorem 3.4. The required notation for the RN+YpR model are contained in Table 1 (which refers to Section 2.2.1). We thus have that condition (3.7) is equivalent to (recall (2.26), (2.27))

$$\min_{a\in\mathscr{A}} v_a > 0 \quad \text{and} \quad \max\left(\{w_a - v_a : a \in \mathscr{A}\} \cup \mathscr{Y} \cup \mathscr{R}\right) < \sum_{a\in\mathscr{A}} v_a, \tag{3.18}$$

while condition (3.8) is equivalent to

$$\min_{a\in\mathscr{A}} v_a > 0 \quad \text{and} \quad \max\left\{w_a + r^b_a, w_b + r^a_b : a \in Y, b \in R\right\} - \min_{a\in\mathscr{A}} v_a < 4 \min_{a\in\mathscr{A}} v_a.$$

One can write $\mathscr{L}_1 = \mathscr{L}_1^{(1)} + \mathscr{L}_1^{(2)}$ with the rates

$$c^{(1)}(\mathtt{a}, \eta) = \begin{cases} r^a_{\mathtt{a}} & \text{if } \eta(0) = \mathtt{g} \text{ and } \eta(-1) = a \in \{\mathtt{c}, \mathtt{t}\}, \\ 0 & \text{else,} \end{cases}$$

$$c^{(1)}(\mathtt{t},\eta)=\begin{cases} r^a_{\mathtt{t}} & \text{if } \eta(0)=\mathtt{c} \text{ and } \eta(1)=a\in\{\mathtt{a},\mathtt{g}\},\\ 0 & \text{else,}\end{cases}$$

$$c^{(1)}(\mathtt{c},\eta)=\begin{cases} r^a_{\mathtt{c}} & \text{if } \eta(0)=\mathtt{t} \text{ and } \eta(1)=a\in\{\mathtt{a},\mathtt{g}\},\\ 0 & \text{else,}\end{cases}$$

$$c^{(1)}(\mathtt{g},\eta)=\begin{cases} r^a_{\mathtt{g}} & \text{if } \eta(0)=\mathtt{a} \text{ and } \eta(-1)=a\in\{\mathtt{c},\mathtt{t}\}.\\ 0 & \text{else,}\end{cases}$$

and

$$c^{(2)}(a,\eta)=\begin{cases} w_a & \text{if } \{a,\eta(0)\}=\{\mathtt{a},\mathtt{g}\} \text{ or } \{a,\eta(0)\}=\{\mathtt{t},\mathtt{c}\},\\ v_a & \text{else.}\end{cases}$$

(see the next two tables). As a consequence, we have

$$K^{(1)}=\max\{r^b_a,r^a_b\,:\,a\in R,b\in Y\},\quad m^{(1)}=0,\quad s^{(1)}=2, \tag{3.19}$$

$$K^{(2)}=\max_{a\in\mathscr{A}} w_a,\quad m^{(2)}=\min_{a\in\mathscr{A}} v_a,\quad s^{(2)}=1. \tag{3.20}$$

As claimed in Remark 3.3, we are in an interesting case because $s^{(2)}=1<s=2$.

a	$S^{(1)}(a)$	$j\in J^{(1)}(a)$	$A^{(1)}_j(a)$	$\lambda^{(1)}_j(a)$	$\overline{\lambda}^{(1)}_j(a)$
$\mathtt{a}$	$\{-1,0\}$	0	$\{\eta\,:\,(\eta(-1),\eta(0))\notin Y\times\{\mathtt{g}\}\}$	0	0
		1	$\{\eta\,:\,\eta(-1)\eta(0)=d\mathtt{g},\,d\in Y, r^d_{\mathtt{a}}=r^{\mathtt{c}}_{\mathtt{a}}\wedge r^{\mathtt{t}}_{\mathtt{a}}\}$	$r^d_{\mathtt{a}}$	$r^d_{\mathtt{a}}$
		2	$\{\eta\,:\,\eta(-1)\eta(0)=e\mathtt{g},\,e\in Y, r^e_{\mathtt{a}}=r^{\mathtt{c}}_{\mathtt{a}}\vee r^{\mathtt{t}}_{\mathtt{a}}\}$	$r^e_{\mathtt{a}}$	$r^e_{\mathtt{a}}-r^d_{\mathtt{a}}$
$\mathtt{t}$	$\{0,1\}$	0	$\{\eta\,:\,(\eta(0),\eta(1))\notin\{\mathtt{c}\}\times R\}$	0	0
		1	$\{\eta\,:\,\eta(0)\eta(1)=\mathtt{c}d,\,d\in R, r^d_{\mathtt{t}}=r^{\mathtt{a}}_{\mathtt{t}}\wedge r^{\mathtt{g}}_{\mathtt{t}}\}$	$r^d_{\mathtt{t}}$	$r^d_{\mathtt{t}}$
		2	$\{\eta\,:\,\eta(0)\eta(1)=\mathtt{c}e,\,e\in R, r^e_{\mathtt{t}}=r^{\mathtt{a}}_{\mathtt{t}}\vee r^{\mathtt{g}}_{\mathtt{t}}\}$	$r^e_{\mathtt{t}}$	$r^e_{\mathtt{t}}-r^d_{\mathtt{t}}$
$\mathtt{c}$	$\{0,1\}$	0	$\{\eta\,:\,(\eta(0),\eta(1))\notin\{\mathtt{t}\}\times R\}$	0	0
		1	$\{\eta\,:\,\eta(0)\eta(1)=\mathtt{t}d,\,d\in R, r^d_{\mathtt{c}}=r^{\mathtt{a}}_{\mathtt{c}}\wedge r^{\mathtt{g}}_{\mathtt{c}}\}$	$r^d_{\mathtt{c}}$	$r^d_{\mathtt{c}}$
		2	$\{\eta\,:\,\eta(0)\eta(1)=\mathtt{t}e,\,e\in R, r^e_{\mathtt{c}}=r^{\mathtt{a}}_{\mathtt{c}}\vee r^{\mathtt{g}}_{\mathtt{c}}\}$	$r^e_{\mathtt{c}}$	$r^e_{\mathtt{c}}-r^d_{\mathtt{c}}$
$\mathtt{g}$	$\{-1,0\}$	0	$\{\eta\,:\,(\eta(-1),\eta(0))\notin Y\times\{\mathtt{a}\}\}$	0	0
		1	$\{\eta\,:\,\eta(-1)\eta(0)=d\mathtt{a},\,d\in Y, r^d_{\mathtt{g}}=r^{\mathtt{c}}_{\mathtt{g}}\wedge r^{\mathtt{t}}_{\mathtt{g}}\}$	$r^d_{\mathtt{g}}$	$r^d_{\mathtt{g}}$
		2	$\{\eta\,:\,\eta(-1)\eta(0)=e\mathtt{a},\,e\in Y, r^e_{\mathtt{g}}=r^{\mathtt{c}}_{\mathtt{g}}\vee r^{\mathtt{t}}_{\mathtt{g}}\}$	$r^e_{\mathtt{g}}$	$r^e_{\mathtt{g}}-r^d_{\mathtt{g}}$
$s^{(1)}=2,\quad \overline{\lambda}^{(1)}=\max(\mathscr{Y}\cup\mathscr{R}),\quad$ and $\quad\overline{\lambda}^{(1)}_0=0$					

a	$S^{(2)}(a)$	$J^{(2)}(a)$	$A^{(2)}_j(a)$	$\lambda^{(2)}_j(a)$	$\overline{\lambda}^{(2)}_j$
$\mathtt{a}$	$\{0\}$	0	$\{\eta\,:\,\eta(0)\neq\mathtt{g}\}$	$v_{\mathtt{a}}$	$v_{\mathtt{a}}$
		1	$\{\eta\,:\,\eta(0)=\mathtt{g}\}$	$w_{\mathtt{a}}$	$w_{\mathtt{a}}-v_{\mathtt{a}}$
$\mathtt{t}$	$\{0\}$	0	$\{\eta\,:\,\eta(0)\neq\mathtt{c}\}$	$v_{\mathtt{t}}$	$v_{\mathtt{t}}$
		1	$\{\eta\,:\,\eta(0)=\mathtt{c}\}$	$w_{\mathtt{t}}$	$w_{\mathtt{t}}-v_{\mathtt{t}}$
$\mathtt{c}$	$\{0\}$	0	$\{\eta\,:\,\eta(0)\neq\mathtt{t}\}$	$v_{\mathtt{c}}$	$v_{\mathtt{c}}$
		1	$\{\eta\,:\,\eta(0)=\mathtt{t}\}$	$w_{\mathtt{c}}$	$w_{\mathtt{c}}-v_{\mathtt{c}}$
$\mathtt{g}$	$\{0\}$	0	$\{\eta\,:\,\eta(0)\neq\mathtt{a}\}$	$v_{\mathtt{g}}$	$v_{\mathtt{g}}$
		1	$\{\eta\,:\,\eta(0)=\mathtt{a}\}$	$w_{\mathtt{g}}$	$w_{\mathtt{g}}-v_{\mathtt{g}}$
$s^{(2)}=1,\quad\overline{\lambda}^{(2)}=\sum_{a\in\mathscr{A}}(w_a-v_a),\quad$ and $\quad\overline{\lambda}^{(2)}_0=\sum_{a\in\mathscr{A}}v_a$					

Condition (3.16) becomes (2.25). Note that since $s^{(2)} = 1$, w_a is not present in condition (2.25), which is thus weaker than condition (3.18), while condition (3.17) becomes (recall (3.19)–(3.20))

$$\min_{a\in\mathscr{A}} v_a > 0 \quad \text{and} \quad \max\{r_a^b, r_b^a : a \in R, b \in Y\} < 4 \min_{a\in\mathscr{A}} v_a. \tag{3.21}$$

□

4 Ergodicity through attractiveness

In this section we present an alternative approach to ergodicity, based on the attractiveness of the process when it is present. For the sake of simplicity, we restrict ourselves to the finite alphabet $\mathscr{A} = \{\mathtt{a}, \mathtt{c}, \mathtt{g}, \mathtt{t}\}$. While our first result (in Section 4.1) is general, the next ones (in Sections 4.2 and 4.3) depend on the specific dynamics of our generic example, the RN+YpR model with cut-and-paste mechanism. The results stated in this section are proved in Section 6.

4.1 Set-up and first result

We first recall the general set-up for attractiveness, relying on [Lig05]. It requires a total order $\leqslant$ on $\mathscr{A}$ which induces a partial order on X. Let us assume that such an order has been defined for the 4 elements of $\mathscr{A}$, that we write $1 < 2 < 3 < 4$ for the moment.

Let η and ξ be two configurations of X. We say that $\eta \leqslant \xi$ if for any $x \in \mathbb{Z}$, we have $\eta(x) \leqslant \xi(x)$. We define $\mathscr{M}$ as the class of bounded monotone functions f on X, that is, for all configurations η and ξ such that $\eta \leqslant \xi$, we have $f(\eta) \leqslant f(\xi)$. The partial order on X induces a stochastic order on $\mathscr{P}(X)$. For two elements μ and ν of $\mathscr{P}(X)$, we say that $\mu \preceq \nu$ if, for any $f \in \mathscr{M}$, we have $\mu(f) \leqslant \nu(f)$.

According to Theorem II.2.2 in [Lig05], for any Feller process on X with semigroup $\{S(t) : t \geqslant 0\}$, the following two statements are equivalent. For any $\mu, \nu \in \mathscr{P}(X)$,

(i) If $f \in \mathscr{M}$, then $S(t)f \in \mathscr{M}$, for any $t \geqslant 0$.

(ii) If $\mu \preceq \nu$, then $\mu S(t) \leqslant \nu S(t)$, for any $t \geqslant 0$.

A Feller process on X with semigroup $\{S(t) : t \geqslant 0\}$ is said to be *attractive* if the equivalent conditions above are satisfied.

An attractive process possesses two special extremal invariant probability measures, the lower and the upper one, that is for us $\mu_l = \lim_{t\to\infty} \delta_{\underline{1}} S(t)$ and $\mu_u = \lim_{t\to\infty} \delta_{\underline{4}} S(t)$, where $\delta_{\underline{1}}$ (resp. $\delta_{\underline{4}}$) denotes the Dirac measure on the configuration η such that $\eta(x) = 1$ (resp. $\eta(x) = 4$) for all $x \in \mathbb{Z}$. Note that they are translation invariant, since the dynamics is translation invariant. They satisfy $\mu_l \preceq \mu_u$ and any invariant probability measure μ is such that $\mu_l \preceq \mu \preceq \mu_u$. The process is ergodic if and only if $\mu_l = \mu_u$.

To deal with our generic example, there are many possibilities to define an order on $\mathscr{A}$. We have to choose one order that will induce attractiveness of the dynamics. We thus define the strict order $<$ on $\mathscr{A}$ as

$$\mathtt{c} < \mathtt{t} < \mathtt{a} < \mathtt{g} \tag{4.1}$$

that we use from now on. Other possible choices will be detailed in Section 4.3.

Proposition 4.1. *Assume that a Feller process on X is attractive with respect to the order in (4.1) and has translation invariant rates. If any invariant and translation invariant probability measure μ satisfies*

$$\mu(\eta(0) = \mathtt{a}) \quad = \quad \mu(\eta(0) = \mathtt{t}), \tag{4.2}$$
$$\mu(\eta(0) = \mathtt{c}) \quad = \quad \mu(\eta(0) = \mathtt{g}), \tag{4.3}$$

then the process is ergodic, that is, $\mu_l = \mu_u$.

Remark 4.2. *(i) Conditions (4.2)–(4.3) imply that $\mu(Y) = \mu(R) = 1/2$.*
(ii) Proposition 4.1 can be extended to an alphabet of size $2n$ or $2n+1$ for an integer $n \geqslant 1$ as follows. Denoting $\mathscr{A} = \{a_1, a_2, \cdots, a_k\}$ with the order $a_1 < a_2 < \cdots < a_k$ and $k \in \{2n, 2n+1\}$, if we assume that any invariant and translation invariant probability measure μ satisfies

$$\mu(\eta(0) = a_i) \quad = \quad \mu(\eta(0) = a_{k-i}) \tag{4.4}$$

for all $1 \leqslant i \leqslant n$, then the process is ergodic.

4.2 The attractive RN+YpR model with cut-and-paste mechanism

The next proposition will enable to concentrate on substitution models to check attractiveness.

Proposition 4.3. *Assume that the RN+YpR model is attractive. Then its superimposition with a cut-and-paste process is attractive as well.*

Proposition 4.4. *Assume that $\mathscr{A}$ is endowed with the order (4.1). Under the conditions*

$$0 \quad = \quad r_{\mathtt{c}}^{\mathtt{g}} = r_{\mathtt{c}}^{\mathtt{a}}, \tag{4.5}$$
$$r_{\mathtt{t}}^{\mathtt{a}} \quad \leqslant \quad r_{\mathtt{t}}^{\mathtt{g}}, \tag{4.6}$$
$$r_{\mathtt{a}}^{\mathtt{t}} \quad \leqslant \quad r_{\mathtt{a}}^{\mathtt{c}}, \tag{4.7}$$
$$r_{\mathtt{g}}^{\mathtt{c}} = r_{\mathtt{g}}^{\mathtt{t}} \quad = \quad 0, \tag{4.8}$$

the RN+YpR model is attractive.

Remark 4.5. *(i) The rates r_a^b have to satisfy inequalities (see (4.6)–(4.7)), except the ones corresponding to a transition to the biggest or to the smallest element of $\mathscr{A}$ (with respect to the order (4.1)), that have to be 0 (see (4.5) and (4.8)).*
(ii) When the transition probability $p(.,.)$ is nearest-neighbor, an application of Theorem 2.4 in [Bor11] gives that conditions (4.5)–(4.8) are also necessary for attractiveness of the RN+YpR model with cut-and-paste mechanism.

The next lemma is an immediate application of Proposition 4.4.

Lemma 4.6. *The T92+$\mathtt{cpg}$ model, hence also the JC+$\mathtt{cpg}$ model, are attractive.*

In view of our next results, we compute the first moments of any translation invariant and invariant probability measure for the RN+YpR model with cut-and-paste mechanism.

Proposition 4.7. *Let $\mu \in \mathscr{I} \cap \mathscr{S}$. It satisfies*

$$\mu(R) = \mu(\eta(0) \in \{\mathtt{a},\mathtt{g}\}) = \frac{v_{\mathrm{R}}}{v_{\mathrm{Y}} + v_{\mathrm{R}}}, \tag{4.9}$$

$$\mu(Y) = \mu(\eta(0) \in \{\mathtt{c},\mathtt{t}\}) = \frac{v_{\mathrm{Y}}}{v_{\mathrm{Y}} + v_{\mathrm{R}}}, \tag{4.10}$$

and

$$\mu(\eta(0) = \mathtt{a}) = \frac{v_{\mathtt{a}}\,\mu(Y) + w_{\mathtt{a}}\,\mu(R) - r_{\mathrm{R}}}{w_{\mathrm{R}} + v_{\mathrm{Y}}}, \tag{4.11}$$

$$\mu(\eta(0) = \mathtt{c}) = \frac{v_{\mathtt{c}}\,\mu(R) + w_{\mathtt{c}}\,\mu(Y) - r_{\mathrm{Y}}}{w_{\mathrm{Y}} + v_{\mathrm{R}}}, \tag{4.12}$$

$$\mu(\eta(0) = \mathtt{g}) = \frac{v_{\mathtt{g}}\,\mu(Y) + w_{\mathtt{g}}\,\mu(R) + r_{\mathrm{R}}}{w_{\mathrm{R}} + v_{\mathrm{Y}}}, \tag{4.13}$$

$$\mu(\eta(0) = \mathtt{t}) = \frac{v_{\mathtt{t}}\,\mu(R) + w_{\mathtt{t}}\,\mu(Y) + r_{\mathrm{Y}}}{w_{\mathrm{Y}} + v_{\mathrm{R}}}, \tag{4.14}$$

where we abbreviated

$$v_{\mathrm{Y}} = v_{\mathtt{t}} + v_{\mathtt{c}} \quad ; \quad v_{\mathrm{R}} = v_{\mathtt{a}} + v_{\mathtt{g}}, \tag{4.15}$$

$$w_{\mathrm{Y}} = w_{\mathtt{t}} + w_{\mathtt{c}} \quad ; \quad w_{\mathrm{R}} = w_{\mathtt{a}} + w_{\mathtt{g}}, \tag{4.16}$$

and

$$\begin{aligned} -r_{\mathrm{Y}} = -r_{\mathrm{Y}}(\mu) &= \sum_{a\in R} r_{\mathtt{c}}^a \mu\left((\eta(0),\eta(1)) = (\mathtt{t},a)\right) \\ &\qquad - \sum_{a\in R} r_{\mathtt{t}}^a \mu\left((\eta(0),\eta(1)) = (\mathtt{c},a)\right), \end{aligned} \tag{4.17}$$

$$\begin{aligned} -r_{\mathrm{R}} = -r_{\mathrm{R}}(\mu) &= \sum_{a\in Y} r_{\mathtt{a}}^a \mu\left((\eta(-1),\eta(0)) = (a,\mathtt{g})\right) \\ &\qquad - \sum_{a\in Y} r_{\mathtt{g}}^a \mu\left((\eta(-1),\eta(0)) = (a,\mathtt{a})\right). \end{aligned} \tag{4.18}$$

Remark 4.8. *The values (4.9)–(4.14) were already computed in [BGP08, Proposition 14]: indeed since we consider first moments of a translation invariant probability measure for a translation invariant dynamics, the cut-and-paste mechanism disappears from our computations. However we cannot obtain two-points correlations as in [BGP08, Proposition 15], since there the cut-and-paste mechanism is present in computations.*

We can now prove part (i) of Theorem 2.5; the first part of (ii) will be proved later on, thanks to Proposition 4.12.

Proof. (of Theorem 2.5, (i)). This is an application of Proposition 4.1: We have to check that equalities (4.2)–(4.3) are satisfied for both examples. Since the JC+cpg model is a particular case of the T92+cpg model, it is enough to consider the latter. By Lemma 4.6 and Proposition 4.3, the T92+cpg model with cut-and-paste mechanism is attractive. Then we compute, using that μ is translation invariant,

$$\begin{aligned}
v_{\mathrm{Y}} = v_{\mathrm{R}} = v \quad &; \quad w_{\mathrm{Y}} = w_{\mathrm{R}} = w, \\
\mu(Y) = \mu(R) \quad &= \quad 1/2, \\
r_{\mathrm{Y}} = r\mu((\eta(0), \eta(1)) = (\mathtt{c}, \mathtt{g})) \quad &= \quad -r_{\mathrm{R}} = r\mu((\eta(-1), \eta(0)) = (\mathtt{c}, \mathtt{g})), \\
\mu(\eta(0) = \mathtt{a}) \quad &= \quad \frac{1-\theta}{2} + \frac{r_{\mathrm{Y}}}{v+w} = \mu(\eta(0) = \mathtt{t}), \\
\mu(\eta(0) = \mathtt{c}) \quad &= \quad \frac{\theta}{2} + \frac{r_{\mathrm{R}}}{v+w} = \mu(\eta(0) = \mathtt{g}).
\end{aligned}$$

Thus (4.2)–(4.3) are satisfied for the T92+cpg model with cut-and-paste mechanism. □

Thanks to Proposition 4.7, it is possible to relax the assumptions of Proposition 4.1 to derive ergodicity:

Proposition 4.9. *Assume that the RN+YpR model with cut-and-paste mechanism is attractive with respect to the order in (4.1). If μ_l and μ_u satisfy either (4.2) or (4.3), then the process is ergodic.*

Remark 4.10. *Under the assumption of Proposition 4.9, we may have $\mu(Y) \neq \mu(R)$, where $\mathscr{I} = \{\mu\}$.*

Another way to look for ergodicity is to investigate the monotone coupling measure ν of two ordered probability measures μ_1 and μ_2. The probability measure ν on $X \times X$ is a *monotone coupling measure* of μ_1 and μ_2, if its marginals are μ_1 and μ_2, and it satisfies

$$\nu((\eta, \xi) : \eta \leqslant \xi) = 1. \tag{4.19}$$

Such a coupling exists by Strassen's Theorem since $\mu_1 \preceq \mu_2$ (see [Lig05, Theorem II.2.4]).

Proposition 4.11. *Assume that $\mathscr{A}$ is endowed with the order (4.1). Let ν be a monotone coupling measure of two translation invariant probability measures μ_1 and μ_2 such that $\mu_1 \preceq \mu_2$, and which both satisfy (4.9)–(4.10). Then we have*

$$\nu((\eta(0), \xi(0)) \in \{(\mathtt{c}, \mathtt{g}), (\mathtt{t}, \mathtt{g}), (\mathtt{c}, \mathtt{a}), (\mathtt{t}, \mathtt{a})\}) = 0. \tag{4.20}$$

This proposition applies to the coupling measure ν of the lower and upper invariant probability measures μ_l and μ_u, when the dynamics is attractive (with respect to the order in (4.1)). Thus, proving that $\nu((\eta(0), \xi(0)) \in \{(\mathtt{a}, \mathtt{g}), (\mathtt{c}, \mathtt{t})\}) = 0$ would imply ergodicity.

Proposition 4.12. *Let ν be a monotone coupling measure of μ_l and μ_u when the process is attractive with respect to the order in (4.1). Assume that the rates satisfy one of the 3 following conditions,*

- *(a)* $r^{\mathtt{c}}_{\mathtt{a}} = r^{\mathtt{t}}_{\mathtt{a}}$*;*
- *(b)* $r^{\mathtt{g}}_{\mathtt{t}} = r^{\mathtt{a}}_{\mathtt{t}}$*;*
- *(c)* $\Big((\alpha)$ *and* $(\gamma)\Big)$ *or* (β) *or* (δ)*;*

where

$$\begin{array}{ll} (\alpha) & (r^{\mathtt{g}}_{\mathtt{t}} - r^{\mathtt{a}}_{\mathtt{t}}) - r^{\mathtt{c}}_{\mathtt{a}} \leqslant 0, \\ (\beta) & 0 < (r^{\mathtt{g}}_{\mathtt{t}} - r^{\mathtt{a}}_{\mathtt{t}}) - r^{\mathtt{c}}_{\mathtt{a}} \leqslant v_{\mathtt{t}} + v_{\mathtt{c}} + w_{\mathtt{g}} + w_{\mathtt{a}}, \\ (\gamma) & (r^{\mathtt{c}}_{\mathtt{a}} - r^{\mathtt{t}}_{\mathtt{a}}) - r^{\mathtt{g}}_{\mathtt{t}} \leqslant 0, \\ (\delta) & 0 < (r^{\mathtt{c}}_{\mathtt{a}} - r^{\mathtt{t}}_{\mathtt{a}}) - r^{\mathtt{g}}_{\mathtt{t}} \leqslant w_{\mathtt{t}} + w_{\mathtt{c}} + v_{\mathtt{g}} + v_{\mathtt{a}}. \end{array}$$

Then we have

$$\nu((\eta(0), \xi(0)) \in \{(\mathtt{a}, \mathtt{g}), (\mathtt{c}, \mathtt{t})\}) = 0, \tag{4.21}$$

hence the RN+YpR model with cut-and-paste mechanism is ergodic.

Remark 4.13. *This proposition gives another proof of Theorem 2.5, (i), since the T92+*`cpg` *model, hence the JC+*`cpg` *model, both with cut-and-paste mechanism, satisfy also (α) and (γ) in the set (c) of conditions of Proposition 4.12.*

4.3 Other order relations for the RN+YpR model with cut-and-paste mechanism

We chose the order (4.1) on $\mathscr{A}$, that gave relations on the rates r_a^b for attractiveness, and eventually ergodicity. In Theorem 2.5 we saw that those relations gave attractiveness and ergodicity for the JC+`cpg`, T92+`cpg`, and some RNc+YpR models with cut-and-paste mechanism. Are there other possible orders on $\mathscr{A}$ that would give attractiveness of the RN+YpR model? There are a priori 24 possibilities.

In what follows, we refer to the proof of Proposition 4.4, done in Section 6. There, we detail the coupling transitions starting from two ordered configurations (with respect to the order (4.1)), and forbid transitions that would break this order between the configurations. Going to the coupling tables in this proof, we see that we cannot take an order relation that would 'separate' the values in Y and R : Let us try for instance $\mathtt{c} < \mathtt{a} < \mathtt{t} < \mathtt{g}$; then we cannot forbid the transition from $(\mathtt{c}, \mathtt{a})$ to $(\mathtt{t}, \mathtt{a})$. This fact forbids 16 possibilities of order.

Then, once we do not separate the values in Y and R, we are left with the following 8 possibilities,

$$\begin{array}{ll}
(O1) & \mathtt{c} < \mathtt{t} < \mathtt{a} < \mathtt{g}, \\
(O2) & \mathtt{g} < \mathtt{a} < \mathtt{t} < \mathtt{c}, \\
(O3) & \mathtt{t} < \mathtt{c} < \mathtt{a} < \mathtt{g}, \\
(O4) & \mathtt{t} < \mathtt{c} < \mathtt{g} < \mathtt{a}, \\
(O5) & \mathtt{c} < \mathtt{t} < \mathtt{g} < \mathtt{a}, \\
(O6) & \mathtt{a} < \mathtt{g} < \mathtt{c} < \mathtt{t}, \\
(O7) & \mathtt{a} < \mathtt{g} < \mathtt{t} < \mathtt{c}, \\
(O8) & \mathtt{g} < \mathtt{a} < \mathtt{c} < \mathtt{t},
\end{array}$$

with the attractiveness conditions they induce, by proceeding as in Proposition 4.4 and doing the ad-hoc permutations. Indeed, there we worked with the order (4.1), that we now denote as order *(O1)*; if we write it as $1 < 2 < 3 < 4$, the attractiveness conditions (4.5)–(4.8) are written

$$0 = r_1^3 = r_1^4, \tag{4.22}$$
$$r_2^3 \leqslant r_2^4, \tag{4.23}$$
$$r_3^1 \geqslant r_3^2, \tag{4.24}$$
$$r_4^1 = r_4^2 = 0, \tag{4.25}$$

and the conditions in Proposition 4.12 become

$$(\widetilde{a})\ r_3^1 = r_3^2\,; \quad (\widetilde{b})\ r_2^4 = r_2^3\,; \quad (\widetilde{c})\ \Big((\widetilde{\alpha})\ \textit{and}\ (\widetilde{\gamma})\Big)\ \textit{or}\ (\widetilde{\beta})\ \textit{or}\ (\widetilde{\delta}); \tag{4.26}$$

where

$$(\widetilde{\alpha}) \qquad (r_2^4 - r_2^3) - r_3^1 \leqslant 0, \tag{4.27}$$
$$(\widetilde{\beta}) \qquad 0 < (r_2^4 - r_2^3) - r_3^1 \leqslant v_2 + v_1 + w_4 + w_3, \tag{4.28}$$
$$(\widetilde{\gamma}) \qquad (r_3^1 - r_3^2) - r_3^4 \leqslant 0, \tag{4.29}$$
$$(\widetilde{\delta}) \qquad 0 < (r_3^1 - r_3^2) - r_3^4 \leqslant w_2 + w_1 + v_4 + v_3. \tag{4.30}$$

If we take also into account the constraint that we want to keep the result for the JC+`cpg` and T92+`cpg` models with cut-and-paste mechanism, then, among the 7 possibilities after *(O1)*, only one is possible, which is *(O2)*, that is, $\mathtt{g} < \mathtt{a} < \mathtt{t} < \mathtt{c}$.
But the other orders enable to deal with other dynamics, for instance the RNc+YpR model with cut-and-paste mechanism, for which we can now derive attractiveness conditions then prove Theorem 2.5, (ii).

Lemma 4.14. *The RNc+YpR model is attractive if its rates satisfy either* (2.29) *or* (2.30).

Proof. We have to check conditions (4.22)–(4.25) respectively for the orders $(O1)$ to $(O8)$: The orders $(O1)$ and $(O2)$ yield (2.29), the orders $(O4)$ and $(O6)$ yield (2.30), while the other orders yield a trivial case, where all the rates are equal to 0. □

Proof. (of Theorem 2.5, (ii)). The two possible sets of rates of the attractive RNc+YpR model with cut-and-paste mechanism (that is (2.29) or (2.30), see Lemma 4.14) satisfy $(\widetilde{\alpha})$ and $(\widetilde{\gamma})$ in the set $(\widetilde{c})$ of the above conditions (4.26), for the respective orders $(O1)$ and $(O4)$. □

5 Proofs through generalized duality

To prove Theorems 3.1 and 3.4 we proceed as follows: In Section 5.1, we provide a graphical construction of the process, which yields a generalized dual of the process. Then this dual is dominated by a branching process, for which we derive a condition for extinction. This one implies exponential ergodicity of the process; see Section 5.2.
Although our proofs are quite similar in spirit to those of [Fer90], we chose to give details for the sake of completeness, and to highlight the places were they are different.

5.1 Graphical constructions and dual process

We adapt to our context the graphical constructions of [Fer90]. We start in section 5.1.1 with the substitution process in our two different cases: either the minimal substitution rate is positive, or the pregenerator $\mathscr{L}_1$ is decomposed in a sum of pregenerators. Then, in Section 5.1.2 we provide the graphical construction of the cut-and-paste process. Finally, in Section 5.1.3 we construct a (generalized) dual process (that is, a non-Markovian dynamics) for the process with pregenerator $\mathscr{L}$.

5.1.1 Graphical construction of the process with pregenerator $\mathscr{L}_1$

• Under Assumption (2.6): the minimal substitution rate is positive

Recall the decomposition (3.1) of the rate function $c(\cdot,\cdot)$ introduced in Section 3.1 and the notations there. For $x \in \mathbb{Z}, a \in \mathscr{A}, j \in J(a)$, let

$$A_j(x,a) = \tau_x^{-1} A_j(a) \tag{5.1}$$

i.e. $\eta \in A_j(x,a)$ if and only if $\tau_x \eta \in A_j(a)$. First injecting (3.1) in (2.1), then using (3.2) yields a rewriting of the pregenerator $\mathscr{L}_1$ as

$$\begin{aligned}
\mathscr{L}_1 f(\eta) &= \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \sum_{j\in J(a)\setminus\{0\}} \overline{\lambda}_j(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell(x,a)\}} \left[f(\eta_a^x) - f(\eta)\right] \\
&\quad + \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \overline{\lambda}_0(a) \sum_{\ell\geqslant 0} \mathbf{1}_{\{\eta\in A_\ell(x,a)\}} \left[f(\eta_a^x) - f(\eta)\right] \\
&= \mathscr{L}_1^{b,1} f(\eta) + \mathscr{L}_1^{n,1} f(\eta)
\end{aligned} \tag{5.2}$$

where

$$\mathscr{L}_1^{b,1} f(\eta) = \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \sum_{j\in J(a)\setminus\{0\}} \overline{\lambda}_j(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell(x,a)\}} \big[f(\eta_a^x) - f(\eta)\big] \tag{5.3}$$

$$\mathscr{L}_1^{n,1} f(\eta) = \sum_{x\in\mathbb{Z}} \overline{\lambda}_0 \sum_{a\in\mathscr{A}} \frac{\overline{\lambda}_0(a)}{\overline{\lambda}_0} \left[f(\eta_a^x) - f(\eta)\right]. \tag{5.4}$$

For the branching process, $\mathscr{L}_1^{b,1}$ will induce the births, while $\mathscr{L}_1^{n,1}$ corresponds to a "noise" part that will induce the deaths. Note that, contrary to [Fer90], our rates for this noise part are not uniform (this is why we will get a better r.h.s. in (3.7) than in (3.8), as it will be explained in the proof of Remark 3.2).

We now define, on the graphical representation, marks (j,a) and (δ,a) (for $a \in \mathscr{A}, j \in J(a) \setminus \{0\}$) that will induce respectively births and deaths in the branching process. The corresponding families of random variables are all mutually independent.
Let $\mathfrak{M} = \{(M_u(x,a), u \geqslant 0) : x \in \mathbb{Z}, a \in \mathscr{A}\}$ be a family of independent Poisson point processes (PPPs) such that the rate of $(M_u(x,a), u \geqslant 0)$ is $\lambda(a)$ defined in (3.4).
Let $\mathscr{U} = \{(U_n(x,a), n \geqslant 0) : x \in \mathbb{Z}, a \in \mathscr{A}\}$ be a family of independent random variables, all uniformly distributed on $[0,1]$. The n-th occurrence of $(M_u(x,a), u \geqslant 0)$ is marked (j,a) with $j \in J(a) \setminus \{0\}$ if

$$\frac{\lambda_{j-1}(a) - \lambda_0(a)}{\lambda(a)} < U_n(x,a) < \frac{\lambda_j(a) - \lambda_0(a)}{\lambda(a)}.$$

Thus the (j,a) marks are distributed according to a PPP with rate $\overline{\lambda}_j(a)$.

Let $\mathfrak{M}^0 = \{(M_u^0(x), u \geqslant 0) : x \in \mathbb{Z}\}$ be a family of independent PPPs such that the rate of $(M_u^0(x), u \geqslant 0)$ is $\overline{\lambda}_0$ defined by (3.5). Let $\mathscr{U}^0 = \{(U_n^0(x), n \geqslant 0) : x \in \mathbb{Z}\}$ be a family of independent discrete random variables with values in $\mathscr{A}$ such that

$$\mathbb{P}(U_n^0(x) = a) = \frac{\lambda_0(a)}{\overline{\lambda}_0}. \tag{5.5}$$

The n-th occurrence of $(M_u^0(x), u \geqslant 0)$ is marked (δ, a) if $U_n^0(x) = a$.

The evolution of the process $(\eta_t)_{t\geqslant 0}$ is now determined by this graphical representation as follows. Let $\omega \in \Omega$ be a configuration of the marked PPPs. Fix a site $x \in \mathbb{Z}$ and a time $t > 0$, and let $0 < T_1 \leqslant \cdots \leqslant T_{d-1} < t$ be the times of the successive marks present at site x in the time interval $[0, t]$. Set $T_0 = 0$ and $T_d = t$. Then we define $\eta_s(x) = \eta_{T_i}(x)$, for $s \in [T_i, T_{i+1}), i < d$, where $\eta_{T_i}(x)$ is constructed with the following recipe.

Suppose that the configuration at time T_i^- is $\eta_{T_i^-}$. By definition of T_i, a mark of ω is present at site x at time T_i. There are two possibilities:

1(a). A (j, a)-mark with $j \in J(a) \setminus \{0\}$. If the configuration $\eta_{T_i^-}$ belongs to at least one of the sets $A_\ell(x, a)$ with $\ell \geqslant j$, then substitute the letter at x by a, so that $\eta_{T_i} = (\eta_{T_i^-})_a^x$. Otherwise, nothing happens.

1(b). A (δ, a)-mark. Substitute the letter at x by a, so that $\eta_{T_i} = (\eta_{T_i^-})_a^x$. Note that this substitution is independent of $\eta_{T_i^-}$, whence the term "noise" for $\mathscr{L}_1^{n,1}$.

The fact that this recipe produces the desired substitution rates $c(\cdot, \cdot)$ comes from the rewriting (5.2) of (2.1), and the thinning property of Poisson processes. Moreover a percolation argument (see e.g. [Dur93], Section 2) implies that only a finite (random) number of sites influence the evolution of a fixed site, hence the previous description yields a well-defined dynamics.

• Under Assumptions (3.9) and (3.13): decomposition of the pregenerator as a sum

Recall the decomposition of the rate functions $c^{(i)}(\cdot, \cdot)$ introduced in Section 3.2 and the notations there. Proceeding as for (5.2) gives a rewriting of the pregenerators $\mathscr{L}_1^{(i)}$ and $\mathscr{L}_1$ as

$$\begin{aligned}
\mathscr{L}_1^{(i)} f(\eta) &= \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \sum_{j\in J^{(i)}(a)\setminus\{0\}} \overline{\lambda}_j^{(i)}(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell^{(i)}(x,a)\}} \big[f(\eta_a^x) - f(\eta)\big] \\
&\quad + \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \overline{\lambda}_0^{(i)}(a) \sum_{\ell\geqslant 0} \mathbf{1}_{\{\eta\in A_\ell^{(i)}(x,a)\}} \big[f(\eta_a^x) - f(\eta)\big] \\
\mathscr{L}_1 f(\eta) &= \mathscr{L}_1^{b,2} f(\eta) + \mathscr{L}_1^{n,2} f(\eta)
\end{aligned} \tag{5.6}$$

where

$$\mathscr{L}_1^{b,2} f(\eta) = \sum_{i=1}^{d} \sum_{x\in\mathbb{Z}} \sum_{a\in\mathscr{A}} \sum_{j\in J^{(i)}(a)\setminus\{0\}} \overline{\lambda}_j^{(i)}(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell^{(i)}(x,a)\}} \big[f(\eta_a^x) - f(\eta)\big] \tag{5.7}$$

$$\mathscr{L}_1^{n,2} f(\eta) = \sum_{x\in\mathbb{Z}} \overline{\lambda}_{0,d} \sum_{a\in\mathscr{A}} \frac{\overline{\lambda}_{0,d}(a)}{\overline{\lambda}_{0,d}} \left[f(\eta_a^x) - f(\eta)\right]. \tag{5.8}$$

and, as in (5.1), for $x \in \mathbb{Z}, a \in \mathscr{A}, j \in J^{(i)}(a)$, $\eta \in A_j^{(i)}(x, a)$ if and only if $\tau_x \eta \in A_j^{(i)}(a)$.

Let $\mathfrak{M}^{(i)} = \{(M_u^{(i)}(x,a), u \geqslant 0) : x \in \mathbb{Z}, a \in \mathscr{A}\}$ be a family of independent PPPs such that the rate of $(M_u^{(i)}(x,a), u \geqslant 0)$ is $\lambda^{(i)}(a)$ (see (3.12)). Let $\mathscr{U}^{(i)} = \{(U_n^{(i)}(x,a), n \geqslant 0) : x \in \mathbb{Z}, a \in \mathscr{A}\}$ be a family of independent random variables, all with uniform law on $[0,1]$. The n-th occurrence of $(M_u^{(i)}(x,a), u \geqslant 0)$ is marked (j,a,i) with $j \in J^{(i)}(a) \setminus \{0\}$ if

$$\frac{\lambda_{j-1}^{(i)}(a) - \lambda_0^{(i)}(a)}{\lambda^{(i)}(a)} < U_n^{(i)}(a) < \frac{\lambda_j^{(i)}(a) - \lambda_0^{(i)}(a)}{\lambda^{(i)}(a)}.$$

Thus the (j,a,i) marks are distributed according to a PPP with rate $\overline{\lambda}_j^{(i)}(a)$.

Let $\mathfrak{M}^{0,d} = \{(M_u^{0,d}(x), u \geqslant 0) : x \in \mathbb{Z}\}$ be a family of independent PPPs such that the rate of $(M_u^{0,d}(x), u \geqslant 0)$ is $\overline{\lambda}_{0,d}$ defined by (3.14). Let $\mathscr{U}^{0,d} = \{(U_n^{0,d}(x), n \geqslant 0) : x \in \mathbb{Z}\}$ be a family of independent discrete random variables with values in $\mathscr{A}$ such that

$$\mathbb{P}(U_n^{0,d}(x) = a) = \frac{\overline{\lambda}_{0,d}(a)}{\overline{\lambda}_{0,d}}. \tag{5.9}$$

The n-th occurrence of $(M_u^{0,d}(x), u \geqslant 0)$ is marked (δ, a) if $U_n^{0,d}(x) = a$.

Let $\omega \in \Omega$ be a configuration of the marked PPPs. Fix a site $x \in \mathbb{Z}$ and a time $t > 0$, and let $0 < T_1 \leqslant \cdots \leqslant T_{d-1} < t$ be the times of the successive marks present at site x in the time interval $[0,t]$. Set $T_0 = 0$ and $T_d = t$. Then we define $\eta_s(x) = \eta_{T_i}(x)$, for $s \in [T_i, T_{i+1})$ where $\eta_{T_i}(x)$ is constructed with the following recipe.

Suppose that the configuration at time T_i^- is $\eta_{T_i^-}$. By definition of T_i, a mark of ω is present at site x at time T_i. There are two possibilities:

1(a). A (j,a,i)-mark with $j \in J^{(i)}(a) \setminus \{0\}$. If the configuration $\eta_{T_i^-}$ belongs to at least one of the sets $A_\ell^{(i)}(x,a)$ with $\ell \geqslant j$, then substitute the letter at x by a, so that $\eta_{T_i} = (\eta_{T_i^-})_a^x$. Otherwise, nothing happens.

1(b). A (δ,a)-mark. Substitute the letter at x by a, so that $\eta_{T_i} = (\eta_{T_i^-})_a^x$.

5.1.2 Graphical construction of the process with pregenerator $\mathscr{L}_2$

Once again, we use a Harris graphical construction based on a family of independent PPPs indexed by $\mathbb{Z} \times \mathbb{Z}$; it is adapted from [AFS04] (which deals with an exclusion process). A Borel-Cantelli argument shows that only a finite number of Poisson processes are involved in the computation of the evolution of a site x until a fixed time t.

Let $\mathscr{N} = \{(N_u(x,y), u \geqslant 0) : (x,y) \in \mathbb{Z}\}$ be a family of independent PPPs such that the rate of the process indexed by (x,y) is $p(x,y)$. At each of its arrival times and each site z such that $x \leqslant z \leqslant y$ (or $y \leqslant z \leqslant x$ if $x > y$), we put a mark $(\circlearrowleft, x, y)$.

Let $\omega \in \Omega$ be a configuration of the marked PPP. Fix a site $z \in \mathbb{Z}$ and a time $t > 0$, and let $0 < T_1 \leqslant \cdots \leqslant T_{d-1} < t$ be the times of the successive marks present at site z

in the time interval $[0,t]$. Set $T_0 = 0$ and $T_d = t$. Then we define $\eta_u(x) = \eta_{T_i}(x)$, for $u \in [T_i, T_{i+1})$ where $\eta_{T_i}(x)$ is constructed with the following recipe.

Suppose that the configuration at time T_i^- is $\eta_{T_i^-}$. By definition of T_i, a mark $(\circlearrowright, x, y)$ of ω is present at site z at time T_i. There are two possibilities:

2(a). $x < y$. In this case, the contents of sites x, $x+1$, $\ldots$, y are right circularly permuted so that $\eta_{T_i} = \sigma_{x,y}(\eta_{T_i^-})$.

2(b). $x > y$. In this case, the contents of sites y, $y+1$, $\ldots$, x are left circularly permuted so that $\eta_{T_i} = \sigma_{x,y}(\eta_{T_i^-})$.

This recipe produces the desired cut-and-paste rates given by $p(.,.)$.

On Figure 3, given a configuration ω of marked PPPs, one can see the evolution of sites 1 to 5.

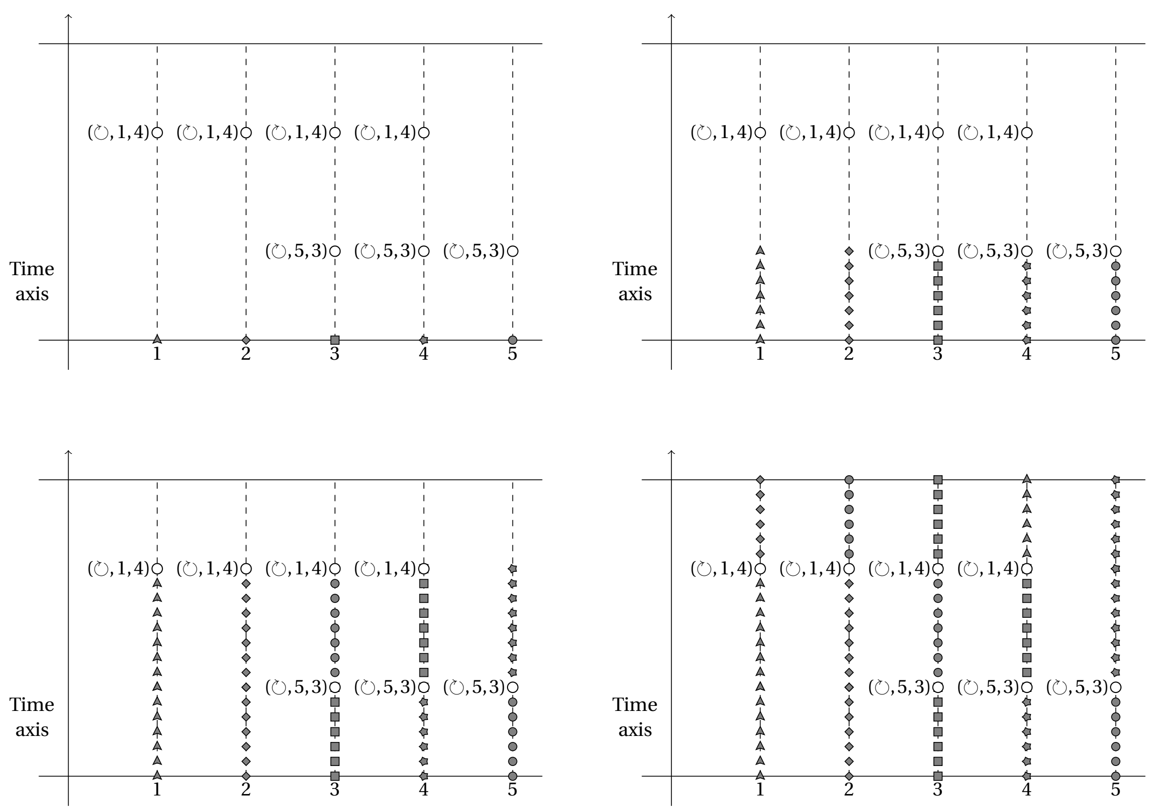

Figure 3: Evolution of the sites 1 to 5 under pregenerator $\mathscr{L}_2$ given a realization of the marked PPPs. It should be read by line from the left to the right. Here, ■ stands for a, ▲ for t, ✿ for c and ◆ for g.

5.1.3 Construction of the dual process

Thanks to the constructions provided in Sections 5.1.1 and 5.1.2, to have a graphical representation of the process with pregenerator $\mathscr{L}$, it suffices to multiply the rates of $\mathscr{N}$ by ρ, and to assume that $\mathfrak{M}$, $\mathfrak{M}^0$, $\mathscr{U}$, $\mathscr{U}^0$ and $\mathscr{N}$, are mutually independent.

Now, we turn to the construction of the generalized dual process of $\mathscr{L}$. It is a marked branching structure constructed on the space $\mathbb{Z} \times [0, +\infty)$.

Fix a finite set of sites $D \subset \mathbb{Z}$ and a time t. Suppose that for the time interval $[0, t]$, we have a realization ω of the marked PPPs described above. We reverse the time direction calling $\widehat{u} = t - u$ and we construct a space-time branching structure contained in $\mathbb{Z} \times [\widehat{0}, \widehat{t}]$ with base $(D, \widehat{0})$ and top $(D_{\widehat{t}}, \widehat{t})$. We proceed by induction with the following recipe.

Suppose that the spatial projection of the structure at time $\widehat{u}$ is $D_{\widehat{u}}$. Let $\widehat{T}$ be the first Poisson mark after $\widehat{u}$ involving some site of $D_{\widehat{u}}$. There are the following possibilities.

1(a). A (j, a)-mark involving site $x \in D_{\widehat{T}^-}$ with $j \in J(a) \setminus \{0\}$. In this case, the point $(x, \widehat{T})$ is marked (j, a) and the set $D_{\widehat{T}}$ will be $D_{\widehat{T}^-} \cup \left(\bigcup_{\ell \geqslant j} S_\ell(x, a) \right)$, where $S_\ell(x, a) = \tau_x^{-1} S_\ell(a)$.

1(b). A (δ, a)-mark involving site $x \in D_{\widehat{T}^-}$. In this case, the point $(x, \widehat{T})$ is marked (δ, a) and the set $D_{\widehat{T}}$ will be $D_{\widehat{T}^-} \setminus \{x\}$.

2(a). A $(\circlearrowright, y, z)$-mark involving $x \in D_{\widehat{T}^-}$ with $y < z$. In this case, all the points of $D_{\widehat{T}} \cap [y, z] \times \{T\}$ are marked with $(\circlearrowright, y, z)$ and the set $D_{\widehat{T}}$ will be $\sigma_{y,z}^{-1}(D_{\widehat{T}^-}) = \sigma_{z,y}(D_{\widehat{T}^-})$.

2(b). A $(\circlearrowleft, y, z)$-mark involving $x \in D_{\widehat{T}^-}$ with $y > z$. In this case, all the points of $D_{\widehat{T}} \cap [z, y] \times \{T\}$ are marked with $(\circlearrowleft, y, z)$ and the set $D_{\widehat{T}}$ will be $\sigma_{z,y}^{-1}(D_{\widehat{T}^-}) = \sigma_{y,z}(D_{\widehat{T}^-})$.

According to this construction, for each finite set D and time t, we are defining a map from the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ into the space of all possible marked branching structures on $\mathbb{Z} \times [\widehat{0}, \widehat{t}]$,

$$\widehat{D}^D_{[\widehat{0},\widehat{t}]} : (\omega, t) \mapsto \left(N, (\widehat{t}_k, x_k, y_k, z_k, j_k, D_k), k = 1, \ldots, N\right), \tag{5.10}$$

where

- N is the number of marks in the interval $[\widehat{0}, \widehat{t}]$,
- $\widehat{t}_k$ is the time of the occurrence of the kth mark,
- x_k is one site involved with the kth mark,

- y_k and z_k are the boundaries of the sites involved with the kth mark, if this mark is a circle arrow mark (if not $y_k = z_k = x_k$),
- j_k is the type of the kth mark (δ, (j,a), or $\circlearrowright$),
- D_k is the set of sites in the spatial projection of the structure between times $\widehat{t}_k$ and $\widehat{t}_{k+1}$.

Note that there are no ambiguities if at least two sites $x, x' \in D_{\widehat{T}^-}$ are involved in the same circle arrow mark $(\circlearrowright, y, z)$ with $y < z$. Indeed in both cases, the same set of points are marked with $(\circlearrowright, y, z)$ and the set $D_{\widehat{T}}$ will be $\sigma_{y,z}^{-1}(D_{\widehat{T}^-})$. It is similar with a circle arrow mark $(\circlearrowright, y, z)$ with $y > z$.

One can check that if D and D' are finite sets of sites such that $D \subset D'$, then the marked branching structure

$$\widehat{D}^D_{[\widehat{0},\widehat{t}]}(\omega, t) = \big(N, (\widehat{t}_k, x_k, y_k, z_k, j_k, D_k), k = 1, \dots, N\big)$$

is a subset of the marked branching structure

$$\widehat{D}^{D'}_{[\widehat{0},\widehat{t}]}(\omega, t) = \big(N', (\widehat{t}'_k, x'_k, y'_k, z'_k, j'_k, D'_k), k = 1, \dots, N'\big)\,,$$

in the sense that $N \leqslant N'$, and for any k in $\{1, \dots, N\}$, there exists $r(k)$ in $\{1, \dots, N'\}$ such that

$$(\widehat{t}_k, x_k, y_k, z_k, j_k) = (\widehat{t}'_{r(k)}, x'_{r(k)}, y'_{r(k)}, z'_{r(k)}, j'_{r(k)}) \quad \text{and} \quad D_k \subset D'_{r(k)}.$$

The central idea of the construction is exactly the same as in [Fer90]: when we go back in time and the generalized dual process meets a (δ, a)-mark at site x, it is not necessary to go further to know the value of $\eta(x)$, because it is determined at that point by an independent random variable.

5.2 Proofs of Theorems 3.1 and 3.4, and of Remarks 3.2 *(i)* and 3.5 *(i)*

Proof. (of Theorem 3.1). The proof is adapted from [Fer90]. If the spatial projection of the dual structure started at time $t = \widehat{0}$ is empty at time $0 = \widehat{t}$, that is, $D^D_{\widehat{t}} = \emptyset$, then $\eta_t(D)$ does not depend on $\eta_0 = \eta$. This implies that a sufficient condition for the exponential ergodicity of the process is that, for all finite set D, there exists positive constants α_1, α_2 such that

$$\mathbb{P}(D^D_{\widehat{t}} = \emptyset) \leqslant \alpha_1 |D| \mathrm{e}^{-t\alpha_2}. \tag{5.11}$$

Hence, we are interested in the evolution of the cardinal of $D^D_{\widehat{u}}$. Since the marks coming from the cut-and-paste process do not change the cardinal of $D^D_{\widehat{u}}$ along time, they are not involved in the following.

The process $|D^D_{\widehat{u}}|$ can be dominated by a branching process $Z^{|D|}_u \in \mathbb{N}$, that is, $|D^D_{\widehat{u}}| \leqslant Z^{|D|}_u$ for all u with probability one. This branching process is defined as follows. At rate $\overline{\lambda} + \overline{\lambda}_0$ (defined in (3.6) and (3.5)) each branch dies and is replaced by either s new branches with probability $\overline{\lambda}/(\overline{\lambda} + \overline{\lambda}_0)$ or 0 new branches with probability $\overline{\lambda}_0/(\overline{\lambda} + \overline{\lambda}_0)$.

Indeed, in the generalized dual process presented in Section 5.1.3, a site x is removed from $D^D_{\widehat{u}}$ at time $\widehat{T}$ when a (δ, a)-mark appears at $(x, \widehat{T})$. The (δ, a)-marks are distributed according to a PPP with rate $\lambda_0(a)$. Therefore, the total rate at which site x is removed from $D^D_{\widehat{u}}$ is $\sum_{a\in\mathscr{A}} \lambda_0(a) = \overline{\lambda}_0$. Hence in the dominating branching process, a branch dies at rate $\overline{\lambda}_0$.

Now, we focus on the apparition of new branches. In the generalized dual process, when a (j, a)-mark appears at $(x, \widehat{T})$, a maximum number of $|\bigcup_{\ell\geqslant j} S_j(x, a)| = s_j(a)$ sites might be added to $D^D_{\widehat{u}}$. This happens at rate $\overline{\lambda}_j(a)$. Therefore for the dominating branching process, we use the bounds $\overline{\lambda}_j(a) \leqslant \overline{\lambda}$ and $s_j(a) \leqslant s$, so that a branch is replaced by s branches at rate $\overline{\lambda}$.

The initial state of the branching process is $Z^{|D|}_0 = |D|$. A sufficient condition for (5.11) is that the average number of branches created at each branching is less than 1. This happens when

$$\frac{s\overline{\lambda}}{\overline{\lambda} + \overline{\lambda}_0} < 1. \tag{5.12}$$

□

Proof. (of Theorem 3.4). The construction in this case of the generalized dual process follows the line of construction in Section 5.1.3, except that one should replace 1(a) by

1(a). A (j, a, i)-mark involving site $x \in D_{\widehat{T}^-}$ with $j \in J^{(i)}(x, a) \setminus \{0\}$. In this case, the point $(x, \widehat{T})$ is marked (j, a, i) and the set $D_{\widehat{T}}$ will be $D_{\widehat{T}^-} \cup S^{(i)}(x, a)$, where $S^{(i)}(x, a) = \tau_x^{-1} S^{(i)}(a)$.

To prove (3.7) in Theorem 3.1, we said that a branch in the dominating branching process dies and is either replaced by s or 0 branches. To prove here (3.16) we say that a branch in the dominating branching process dies at rate $\Theta = \overline{\lambda}_{0,d} + \sum_{i=1}^d \overline{\lambda}^{(i)}$ (recall (3.14) and (3.15)), and is either replaced by $s^{(1)}, \ldots, s^{(d)}$ or 0 branches at respective rates $\overline{\lambda}^{(1)}/\Theta, \ldots, \overline{\lambda}^{(d)}/\Theta$ and $\overline{\lambda}_{0,d}/\Theta$.

A sufficient condition for (5.11) is now

$$\sum_{i=1}^{d} \frac{s^{(i)}\overline{\lambda}^{(i)}}{\Theta} < 1, \tag{5.13}$$

which is equivalent to (3.16). □

Proof. (of Remarks 3.2(i) and 3.5(i)). To recover analogous results to Theorems 2.1 and 2.2 in [Fer90] requires two main changes in the steps to prove Theorems 3.1 and 3.4: We have to take a 'uniform noise' when rewriting the generator $\mathscr{L}_1$, then to take a different bound to define the birth rates of the branching process. We define the quantities we have to modify with an upper index F (we keep the other ones unchanged), and consider in parallel the two results.

We have to replace the definitions of $\overline{\lambda}_0(a)$ in (3.3), and of $\overline{\lambda}_0^{(i)}(a)$ in (3.11), by

$$\overline{\lambda}_0^{\mathrm{F}}(a) = \lambda_0(a) - m, \qquad \overline{\lambda}_0^{(i),\mathrm{F}}(a) = \lambda_0^{(i)}(a) - m^{(i)}, \tag{5.14}$$

the one of $\lambda(a)$ in (3.4), and of $\lambda^{(i)}(a)$ in (3.12), by

$$\lambda^{\mathrm{F}}(a) \;=\; \max\big\{\lambda_j(a) \,:\, j \in J(a)\big\} - m = \lambda_{|J(a)|-1} - m, \tag{5.15}$$

$$\lambda^{(i),\mathrm{F}}(a) \;=\; \max\big\{\lambda_j^{(i)}(a) \,:\, j \in J^{(i)}(a)\big\} - m^{(i)} = \lambda^{(i)}_{|J^{(i)}(a)|-1} - m^{(i)}; \tag{5.16}$$

also take, instead of (3.5) and of (3.14),

$$\overline{\lambda}_0^{\mathrm{F}} = |\mathscr{A}|m, \qquad \overline{\lambda}_{0,d}^{\mathrm{F}} = |\mathscr{A}|\sum_{i=1}^{d} m^{(i)}, \tag{5.17}$$

and, instead of (3.6) and of (3.15),

$$\overline{\lambda}^{\mathrm{F}} \;=\; \max_{a\in\mathscr{A}}\Big(\overline{\lambda}_0^{\mathrm{F}}(a) + \sum_{j\in J(a)\setminus\{0\}} \overline{\lambda}_j(a)\Big) = \max_{a\in\mathscr{A}} \lambda^{\mathrm{F}}(a) = K - m, \tag{5.18}$$

$$\overline{\lambda}^{(i),\mathrm{F}} \;=\; \max_{a\in\mathscr{A}}\Big(\overline{\lambda}_0^{(i),\mathrm{F}}(a) + \sum_{j\in J^{(i)}(a)\setminus\{0\}} \overline{\lambda}_j^{(i)}(a)\Big) = \max_{a\in\mathscr{A}} \lambda^{(i),\mathrm{F}}(a) = K^{(i)} - m^{(i)}. \tag{5.19}$$

Then, instead of (5.2) and of (5.6), rewrite the pregenerator $\mathscr{L}_1$ as

$$\begin{aligned}
\mathscr{L}_1 f(\eta) \;=\; & \sum_{x\in\mathbb{Z}}\sum_{a\in\mathscr{A}}\sum_{j\in J(a)\setminus\{0\}} \overline{\lambda}_j(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell(x,a)\}}\big[f(\eta_a^x) - f(\eta)\big] \\
& + \sum_{x\in\mathbb{Z}}\sum_{a\in\mathscr{A}} \overline{\lambda}_0^{\mathrm{F}}(a) \sum_{\ell\geqslant 0} \mathbf{1}_{\{\eta\in A_\ell(x,a)\}}\big[f(\eta_a^x) - f(\eta)\big] \\
& + \sum_{x\in\mathbb{Z}} \overline{\lambda}_0^{\mathrm{F}} \sum_{a\in\mathscr{A}} \frac{1}{|\mathscr{A}|}\left[f(\eta_a^x) - f(\eta)\right].
\end{aligned} \tag{5.20}$$

and

$$\begin{aligned}
\mathscr{L}_1 f(\eta) \;=\; & \sum_{i=1}^{d}\sum_{x\in\mathbb{Z}}\sum_{a\in\mathscr{A}}\sum_{j\in J^{(i)}(a)\setminus\{0\}} \overline{\lambda}_j^{(i)}(a) \sum_{\ell\geqslant j} \mathbf{1}_{\{\eta\in A_\ell^{(i)}(x,a)\}}\big[f(\eta_a^x) - f(\eta)\big] \\
& + \sum_{i=1}^{d}\sum_{x\in\mathbb{Z}}\sum_{a\in\mathscr{A}} \overline{\lambda}_0^{(i),\mathrm{F}}(a) \sum_{\ell\geqslant 0} \mathbf{1}_{\{\eta\in A_\ell^{(i)}(x,a)\}}\big[f(\eta_a^x) - f(\eta)\big] \\
& + \sum_{x\in\mathbb{Z}} \overline{\lambda}_{0,d}^{\mathrm{F}} \sum_{a\in\mathscr{A}} \frac{1}{|\mathscr{A}|}\left[f(\eta_a^x) - f(\eta)\right].
\end{aligned} \tag{5.21}$$

In the graphical construction, replace the definitions (5.5) and (5.9) by

$$\mathbb{P}(U_n^{0,\mathrm{F}}(x)=a)=\frac{1}{|\mathscr{A}|};\qquad \mathbb{P}(U_n^{0,d,\mathrm{F}}(x)=a)=\frac{1}{|\mathscr{A}|}, \tag{5.22}$$

so that these discrete random variables become uniform. Finally, the conditions for extinction (5.12) and (5.13) become

$$\frac{s\overline{\lambda}^{\mathrm{F}}}{\overline{\lambda}^{\mathrm{F}}+\overline{\lambda}_0^{\mathrm{F}}}<1,\qquad \frac{\sum_{i=1}^d s^{(i)}\overline{\lambda}^{(i),\mathrm{F}}}{\overline{\lambda}_{0,d}^{\mathrm{F}}+\sum_{i=1}^d\overline{\lambda}^{(i),\mathrm{F}}}<1. \tag{5.23}$$

□

6 Proofs through attractiveness

Proof. (of Proposition 4.1). The lower and upper invariant probability measures are $\mu_l=\lim_{t\to\infty}\delta_{\underline{\mathtt{c}}}S(t)$ and $\mu_u=\lim_{t\to\infty}\delta_{\underline{\mathtt{g}}}S(t)$, where $\delta_{\underline{\mathtt{c}}}$ (resp. $\delta_{\underline{\mathtt{g}}}$) denotes the Dirac measure on the configuration η such that $\eta(x)=\mathtt{c}$ (resp. $\eta(x)=\mathtt{g}$) for all $x\in\mathbb{Z}$. They are translation invariant. To prove ergodicity, we have to show that $\mu_l=\mu_u$.

• *Step 1:* We derive consequences of attractiveness.
Note that the functions $\phi(\eta)=\mathbf{1}_{\{\eta(0)\geqslant\mathtt{g}\}}$ and $\varphi(\eta)=\mathbf{1}_{\{\eta(0)>\mathtt{c}\}}$ belong to $\mathscr{M}$, and since $\mathtt{c}$ and $\mathtt{g}$ are respectively the smallest and largest elements of $\mathscr{A}$ with respect to the order (4.1), we have $\phi(\eta)=\mathbf{1}_{\{\eta(0)=\mathtt{g}\}}$ and $\varphi(\eta)=1-\mathbf{1}_{\{\eta(0=\mathtt{c}\}}$, hence

$$\mu_l(\eta(0)=\mathtt{g})\leqslant\mu_u(\eta(0)=\mathtt{g})\qquad\text{and}\qquad \mu_u(\eta(0)=\mathtt{c})\leqslant\mu_l(\eta(0)=\mathtt{c}). \tag{6.1}$$

Similarly, the function $\psi(\eta)=\mathbf{1}_{\{\eta(0)>\mathtt{t}\}}$ belongs to $\mathscr{M}$, and, by the order (4.1), it satisfies $\psi(\eta)=\mathbf{1}_{\{\eta(0)=\mathtt{a}\}}+\mathbf{1}_{\{\eta(0)=\mathtt{g}\}}=1-\mathbf{1}_{\{\eta(0)=\mathtt{c}\}}-\mathbf{1}_{\{\eta(0)=\mathtt{t}\}}$. We thus have that

$$\mu_l(\eta(0)=\mathtt{a})+\mu_l(\eta(0)=\mathtt{g})\ \leqslant\ \mu_u(\eta(0)=\mathtt{a})+\mu_u(\eta(0)=\mathtt{g}), \tag{6.2}$$
$$\mu_l(\eta(0)=\mathtt{c})+\mu_l(\eta(0)=\mathtt{t})\ \geqslant\ \mu_u(\eta(0)=\mathtt{c})+\mu_u(\eta(0)=\mathtt{t}). \tag{6.3}$$

• *Step 2:* We take into account the assumptions on the rates.
Combining (6.1) with (4.3) implies

$$\mu_l(\eta(0)=\mathtt{c})=\mu_u(\eta(0)=\mathtt{c})=\mu_l(\eta(0)=\mathtt{g})=\mu_u(\eta(0)=\mathtt{g}). \tag{6.4}$$

Then combining (6.4) first with (6.2), and then with (6.3) yields

$$\mu_l(\eta(0)=\mathtt{a})\ \leqslant\ \mu_u(\eta(0)=\mathtt{a}), \tag{6.5}$$
$$\mu_l(\eta(0)=\mathtt{t})\ \geqslant\ \mu_u(\eta(0)=\mathtt{t}). \tag{6.6}$$

Finally, combining (6.5)–(6.6) with (4.2) implies

$$\mu_l(\eta(0)=\mathtt{t})=\mu_u(\eta(0)=\mathtt{t})=\mu_l(\eta(0)=\mathtt{a})=\mu_u(\eta(0)=\mathtt{a}). \tag{6.7}$$

• *Step 3:* We now proceed in the same spirit as in Corollary II.2.8 in [Lig05]. Let ν be a monotone coupling measure of μ_l and μ_u. It thus has to satisfy

$$\nu((\eta,\xi):\eta(0)\neq\xi(0))=\nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{t}),(\mathtt{c},\mathtt{a}),(\mathtt{c},\mathtt{g}),(\mathtt{t},\mathtt{a}),(\mathtt{t},\mathtt{g}),(\mathtt{a},\mathtt{g})\}), \tag{6.8}$$

so that by (6.4),

$$\begin{aligned}\nu((\eta,\xi):\eta(0)=\xi(0)=\mathtt{g}) &= \mu_l(\eta(0)=\mathtt{g})=\mu_u(\xi(0)=\mathtt{g}), &(6.9)\\ \nu((\eta,\xi):\eta(0)=\xi(0)=\mathtt{c}) &= \mu_u(\xi(0)=\mathtt{c})=\mu_l(\eta(0)=\mathtt{c}). &(6.10)\end{aligned}$$

By (6.8), (6.9) and (6.10), and because $\mathtt{c}$ and $\mathtt{g}$ are respectively the smallest and largest elements of $\mathscr{A}$ with respect to the order (4.1),

$$\begin{aligned}\nu((\eta,\xi):\eta(0)\neq\mathtt{g},\xi(0)=\mathtt{g}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{g}),(\mathtt{t},\mathtt{g}),(\mathtt{a},\mathtt{g})\})\\ &= \mu_u(\xi(0)=\mathtt{g})-\nu(\eta(0)=\xi(0)=\mathtt{g})=0, &(6.11)\\ \nu((\eta,\xi):\eta(0)=\mathtt{c},\xi(0)\neq\mathtt{c}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{t}),(\mathtt{c},\mathtt{a}),(\mathtt{c},\mathtt{g})\})\\ &= \mu_l(\eta(0)=\mathtt{c})-\nu(\eta(0)=\xi(0)=\mathtt{c})=0, &(6.12)\end{aligned}$$

so that, using also again (6.8) with respectively (6.12) and (6.11), we obtain,

$$\begin{aligned}\nu((\eta,\xi):\eta(0)\neq\mathtt{a},\xi(0)=\mathtt{a}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{a}),(\mathtt{t},\mathtt{a})\})\\ &= 0+\nu(\eta(0)=\mathtt{t},\xi(0)=\mathtt{a}), &(6.13)\\ \nu((\eta,\xi):\eta(0)=\mathtt{a},\xi(0)\neq\mathtt{a}) &= \nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=0. &(6.14)\end{aligned}$$

Since we also have

$$\begin{aligned}\nu((\eta,\xi):\eta(0)\neq\mathtt{a},\xi(0)=\mathtt{a}) &= \mu_u(\xi(0)=\mathtt{a})-\nu(\eta(0)=\xi(0)=\mathtt{a}), &(6.15)\\ \nu((\eta,\xi):\eta(0)=\mathtt{a},\xi(0)\neq\mathtt{a}) &= \mu_l(\eta(0)=\mathtt{a})-\nu(\eta(0)=\xi(0)=\mathtt{a}), &(6.16)\end{aligned}$$

combining (6.13), (6.14) with (6.7) implies that

$$\nu(\eta(0)=\mathtt{t},\xi(0)=\mathtt{a})=0. \tag{6.17}$$

Note that all the possible cases on the r.h.s. of (6.8) have probability 0: $(\mathtt{c},\mathtt{g}),(\mathtt{t},\mathtt{g}),(\mathtt{a},\mathtt{g})$ by (6.11), $(\mathtt{c},\mathtt{t}),(\mathtt{c},\mathtt{a})$ by (6.12), $(\mathtt{t},\mathtt{a})$ by (6.17). We conclude that

$$\nu((\eta,\xi):\eta(0)\neq\xi(0))=0,$$

hence $\mu_l=\mu_u$. □

• **Monotonicity of the RN+YpR model, and of the RN+YpR model with cut-and-paste mechanism**

Proof. (of Proposition 4.3). We denote by $\overline{\mathscr{L}_1}$ the pregenerator of the monotone coupled dynamics for the RN+YpR substitution process (which exists by our assumption). Let similarly $\overline{\mathscr{L}}_2$ denote the pregenerator of the coupled cut-and-paste dynamics through basic coupling, that is, the same transition takes place for both copies of the process: it is defined by, for a cylinder function g on $X \times X$,

$$\overline{\mathscr{L}}_2 g(\eta, \xi) = \sum_{x,y \in \mathbb{Z}} p(x,y) \big[g(\sigma_{x,y}(\eta), \sigma_{x,y}(\xi)) - g(\eta, \xi) \big]. \tag{6.18}$$

This coupled dynamics is monotone, since a transition does not change the way in which the values of the two processes on each site are coupled. For the complete dynamics (that is the RN+YpR model with cut-and-paste mechanism) we consider the combination of both couplings, that is the pregenerator

$$\overline{\mathscr{L}} g = \overline{\mathscr{L}}_1 g + \rho \overline{\mathscr{L}}_2 g. \tag{6.19}$$

Being the sum of two pregenerators of attractive dynamics, it yields also an attractive dynamics. We denote by $(\overline{S}(t), t \geqslant 0)$ its semi-group. □

Proof. (of Proposition 4.4). To derive attractiveness, we construct a coupled dynamics $(\eta_t, \xi_t)_{t \geqslant 0}$ starting from ordered configurations $\eta_0 \leqslant \xi_0$, through basic coupling. Then we find conditions on the rates prohibiting the coupled transitions breaking the increasing order between coupled configurations. We will denote by $\overline{\mathscr{L}}_1$ the induced coupled generator (its existence was assumed in Proposition 4.3). Similarly with (2.1), this generator is defined on a cylinder function f on $X \times X$ by

$$\overline{\mathscr{L}}_1 f(\eta, \xi) = \sum_{x \in \mathbb{Z}} \sum_{(a,b) \in \mathscr{A} \times \mathscr{A}} \overline{c}((a,b), \tau_x(\eta, \xi)) \big[f(\eta_a^x, \xi_b^x) - f(\eta, \xi) \big]. \tag{6.20}$$

We now define the coupled rates $\overline{c}((a,b), \tau_x(\eta, \xi))$ for the transitions $(\eta, \xi) \to (\eta_a^x, \xi_b^x)$.

By translation invariance of the dynamics, it is enough to look at site 0. Thus we write the coupled transitions and their rates (according to basic coupling) in the following 3 tables. There, we indicate with the symbol $(*)$ the coupled transitions to be forbidden for attractiveness; we derive after each table the corresponding sufficient conditions that these forbidden interactions induce.

We rely on the rates given in Table 1 for the RN+YpR model.

Transition				Rate
$(\mathtt{c},\mathtt{c})$	$\to$	$(\mathtt{a},\mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{g},\mathtt{g})$		$v_{\mathtt{g}}$
	$\to$	$(\mathtt{t},\mathtt{t})$		$\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$
	$\to$	$(\mathtt{c},\mathtt{t})$		$c(\mathtt{t},\xi)-\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$
	$\to$	$(\mathtt{t},\mathtt{c})$	$(*)$	$c(\mathtt{t},\eta)-\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$
$(\mathtt{c},\mathtt{t})$	$\to$	$(\mathtt{a},\mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{g},\mathtt{g})$		$v_{\mathtt{g}}$
	$\to$	$(\mathtt{t},\mathtt{t})$		$c(\mathtt{t},\eta)$
	$\to$	$(\mathtt{c},\mathtt{c})$		$c(\mathtt{c},\xi)$
$(\mathtt{c},\mathtt{a})$	$\to$	$(\mathtt{a},\mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{t},\mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{t},\mathtt{a})$		$c(\mathtt{t},\eta)-v_{\mathtt{t}}$
	$\to$	$(\mathtt{c},\mathtt{c})$		$v_{\mathtt{c}}$
	$\to$	$(\mathtt{g},\mathtt{g})$		$v_{\mathtt{g}}$
	$\to$	$(\mathtt{c},\mathtt{g})$		$c(\mathtt{g},\xi)-v_{\mathtt{g}}$
$(\mathtt{c},\mathtt{g})$	$\to$	$(\mathtt{a},\mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{c},\mathtt{a})$		$c(\mathtt{a},\xi)-v_{\mathtt{a}}$
	$\to$	$(\mathtt{t},\mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{t},\mathtt{g})$		$c(\mathtt{t},\eta)-v_{\mathtt{t}}$
	$\to$	$(\mathtt{c},\mathtt{c})$		$v_{\mathtt{c}}$
	$\to$	$(\mathtt{g},\mathtt{g})$		$v_{\mathtt{g}}$

Under basic coupling, both configurations undergo the same transition according to the maximal possible rate, and then uncoupled transitions are added to fit the correct transitions for each marginal. We detail this construction in the first 5 lines of this first table, the others are similar. There, we start from $(\eta(0),\xi(0))=(\mathtt{c},\mathtt{c})$. The rate for a transition from $\mathtt{c}$ to $\mathtt{a}$ or to $\mathtt{g}$ in Table 1 does not depend on the value of the configuration on neighboring sites, therefore here we have transitions respectively to $(\mathtt{a},\mathtt{a})$ or $(\mathtt{g},\mathtt{g})$ with the rates $v_{\mathtt{a}}$ or $v_{\mathtt{g}}$, and these rates yield the correct rate for each marginal transition. But the rate for a transition from $\mathtt{c}$ to $\mathtt{t}$ in Table 1 depends on the value of the configuration on site 1. Therefore the maximal rate for a coupled transition from $(\mathtt{c},\mathtt{c})$ to $(\mathtt{t},\mathtt{t})$ is $\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$, and, to obtain the correct rate for each marginal transition, it has to be supplemented by respective uncoupled transitions to $(\mathtt{c},\mathtt{t})$ and $(\mathtt{t},\mathtt{c})$, with rates $c(\mathtt{t},\xi)-\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$ and $c(\mathtt{t},\eta)-\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}$.

But a transition to $(\mathtt{t},\mathtt{c})$ would break the increasing order between the coupled configurations. To forbid it, that is, for its rate to be 0, we need to have

$$c(\mathtt{t},\eta)=\min\{c(\mathtt{t},\eta),c(\mathtt{t},\xi)\}. \tag{6.21}$$

There are two possibilities, according to the value of the coupled configuration on 1, which is such that $\eta(1)\leqslant\xi(1)$. Either $\eta(1)\in Y=\{\mathtt{c},\mathtt{t}\}$, hence $c(\mathtt{t},\eta)=w_{\mathtt{t}}\leqslant c(\mathtt{t},\xi)$, and

(6.21) is satisfied; or $\eta(1) \in R = \{\mathtt{a}, \mathtt{g}\}$, hence $c(\mathtt{t}, \eta) = w_{\mathtt{t}} + r_{\mathtt{t}}^{\eta(1)}$ and $c(\mathtt{t}, \xi) = w_{\mathtt{t}} + r_{\mathtt{t}}^{\xi(1)}$: a necessary and sufficient condition for (6.21) to be satisfied is

$$r_{\mathtt{t}}^{\mathtt{a}} \leqslant r_{\mathtt{t}}^{\mathtt{g}}.$$

Transition				Rate
$(\mathtt{t}, \mathtt{t})$	$\to$	$(\mathtt{a}, \mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{c}, \mathtt{c})$		$\min\{c(\mathtt{c}, \eta), c(\mathtt{c}, \xi)\}$
	$\to$	$(\mathtt{c}, \mathtt{t})$		$c(\mathtt{c}, \eta) - \min\{c(\mathtt{c}, \eta), c(\mathtt{c}, \xi)\}$
	$\to$	$(\mathtt{t}, \mathtt{c})$	$(*)$	$c(\mathtt{c}, \xi) - \min\{c(\mathtt{c}, \eta), c(\mathtt{c}, \xi)\}$
	$\to$	$(\mathtt{g}, \mathtt{g})$		$v_{\mathtt{g}}$
$(\mathtt{t}, \mathtt{a})$	$\to$	$(\mathtt{a}, \mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{t}, \mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{c}, \mathtt{c})$		$v_{\mathtt{c}}$
	$\to$	$(\mathtt{c}, \mathtt{a})$		$c(\mathtt{c}, \eta) - v_{\mathtt{c}}$
	$\to$	$(\mathtt{g}, \mathtt{g})$		$v_{\mathtt{g}}$
	$\to$	$(\mathtt{t}, \mathtt{g})$		$c(\mathtt{g}, \xi) - v_{\mathtt{g}}$
$(\mathtt{t}, \mathtt{g})$	$\to$	$(\mathtt{a}, \mathtt{a})$		$v_{\mathtt{a}}$
	$\to$	$(\mathtt{t}, \mathtt{a})$		$c(\mathtt{a}, \xi) - v_{\mathtt{a}}$
	$\to$	$(\mathtt{c}, \mathtt{c})$		$v_{\mathtt{c}}$
	$\to$	$(\mathtt{c}, \mathtt{g})$		$c(\mathtt{c}, \eta) - v_{\mathtt{c}}$
	$\to$	$(\mathtt{t}, \mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{g}, \mathtt{g})$		$v_{\mathtt{g}}$

In this second table, the transition from $(\mathtt{t}, \mathtt{t})$ to $(\mathtt{t}, \mathtt{c})$ has to be forbidden for attractiveness, which requires

$$c(\mathtt{c}, \xi) = \min\{c(\mathtt{c}, \eta), c(\mathtt{c}, \xi)\}. \tag{6.22}$$

We have that

$$c(\mathtt{c}, \eta) = w_{\mathtt{c}} + r_{\mathtt{c}}^{\eta(1)} \mathbf{1}_{\{\eta(1) \in R\}}, \quad c(\mathtt{c}, \xi) = w_{\mathtt{c}} + r_{\mathtt{c}}^{\xi(1)} \mathbf{1}_{\{\xi(1) \in R\}}, \quad \text{with} \quad \eta(1) \leqslant \xi(1).$$

Thus (6.22) is satisfied if and only if

$$r_{\mathtt{c}}^{\mathtt{a}} = r_{\mathtt{c}}^{\mathtt{g}} = 0.$$

Transition				Rate
$(\mathtt{a},\mathtt{a})$	$\to$	$(\mathtt{t},\mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{c},\mathtt{c})$		$v_{\mathtt{c}}$
	$\to$	$(\mathtt{g},\mathtt{g})$		$\min\{c(\mathtt{g},\eta),c(\mathtt{g},\xi)\}$
	$\to$	$(\mathtt{a},\mathtt{g})$		$c(\mathtt{g},\xi)-\min\{c(\mathtt{g},\eta),c(\mathtt{g},\xi)\}$
	$\to$	$(\mathtt{g},\mathtt{a})$	$(*)$	$c(\mathtt{g},\eta)-\min\{c(\mathtt{g},\eta),c(\mathtt{g},\xi)\}$
$(\mathtt{a},\mathtt{g})$	$\to$	$(\mathtt{a},\mathtt{a})$		$c(\mathtt{a},\xi)$
	$\to$	$(\mathtt{g},\mathtt{g})$		$c(\mathtt{g},\eta)$
	$\to$	$(\mathtt{t},\mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{c},\mathtt{c})$		$v_{\mathtt{c}}$
$(\mathtt{g},\mathtt{g})$	$\to$	$(\mathtt{a},\mathtt{a})$		$\min\{c(\mathtt{a},\eta),c(\mathtt{a},\xi)\}$
	$\to$	$(\mathtt{a},\mathtt{g})$		$c(\mathtt{a},\eta)-\min\{c(\mathtt{a},\eta),c(\mathtt{a},\xi)\}$
	$\to$	$(\mathtt{g},\mathtt{a})$	$(*)$	$c(\mathtt{a},\xi)-\min\{c(\mathtt{a},\eta),c(\mathtt{a},\xi)\}$
	$\to$	$(\mathtt{t},\mathtt{t})$		$v_{\mathtt{t}}$
	$\to$	$(\mathtt{c},\mathtt{c})$		$v_{\mathtt{c}}$

In this third table, the transitions from $(\mathtt{a},\mathtt{a})$ and from $(\mathtt{g},\mathtt{g})$ to $(\mathtt{g},\mathtt{a})$ have to be forbidden for attractiveness, which requires

$$\begin{aligned} c(\mathtt{g},\eta) &= \min\{c(\mathtt{g},\eta),c(\mathtt{g},\xi)\}, && (6.23)\\ c(\mathtt{a},\xi) &= \min\{c(\mathtt{a},\eta),c(\mathtt{a},\xi)\}. && (6.24)\end{aligned}$$

We have that

$$\begin{aligned} c(\mathtt{g},\eta) &= w_{\mathtt{g}}+r_{\mathtt{g}}^{\eta(-1)}\mathbf{1}_{\{\eta(-1)\in Y\}},\ c(\mathtt{g},\xi)=w_{\mathtt{g}}+r_{\mathtt{g}}^{\xi(-1)}\mathbf{1}_{\{\xi(-1)\in Y\}},\ \text{with } \eta(-1)\leqslant\xi(-1),\\ c(\mathtt{a},\eta) &= w_{\mathtt{a}}+r_{\mathtt{a}}^{\eta(-1)}\mathbf{1}_{\{\eta(-1)\in Y\}},\ c(\mathtt{a},\xi)=w_{\mathtt{a}}+r_{\mathtt{a}}^{\xi(-1)}\mathbf{1}_{\{\xi(-1)\in Y\}},\ \text{with } \eta(-1)\leqslant\xi(-1).\end{aligned}$$

Thus (6.23) and (6.24) are respectively satisfied if and only if

$$\begin{aligned} r_{\mathtt{g}}^{\mathtt{c}}=r_{\mathtt{g}}^{\mathtt{t}} &= 0,\\ r_{\mathtt{a}}^{\mathtt{t}} &\leqslant r_{\mathtt{a}}^{\mathtt{c}}.\end{aligned}$$

The proposition is proved. □

Proof. (of Proposition 4.7). We first compute $\mathscr{L}f(\eta)$ for $f(\eta)=\mathbf{1}_{\{\eta(0)=a\}}$, with $a\in\mathscr{A}$. Recall that we write $c(b,\eta)$ for $c(0,b,\eta)$.

$$\begin{aligned} \mathscr{L}_1 f(\eta) &= \sum_{b\in\mathscr{A}} c(b,\eta)[f(\eta_b^0)-f(\eta)]\\ &= -\sum_{b\in\mathscr{A},\, b\neq a} c(b,\eta)f(\eta)+c(a,\eta)\mathbf{1}_{\{\eta(0)\neq a\}}, && (6.25)\\ \mathscr{L}_2 f(\eta) &= \sum_{x\in\mathbb{Z}} p(x,0)[\mathbf{1}_{\{\eta(x)=a\}}-\mathbf{1}_{\{\eta(0)=a\}}]+\sum_{x,y\in\mathbb{Z},\, x<0<y} p(x,y)[\mathbf{1}_{\{\eta(-1)=a\}}-\mathbf{1}_{\{\eta(0)=a\}}]\\ &\quad+\sum_{x,y\in\mathbb{Z},\, y<0<x} p(x,y)[\mathbf{1}_{\{\eta(1)=a\}}-\mathbf{1}_{\{\eta(0)=a\}}]. && (6.26)\end{aligned}$$

Because μ is translation invariant, using (6.26), we have that $\int \mathscr{L}_2 f(\eta) d\mu(\eta) = 0$. Therefore, by the invariance of μ, we have $0 = \int \mathscr{L} f(\eta) d\mu(\eta) = \int \mathscr{L}_1 f(\eta) d\mu(\eta)$. Thus we only need to compute $\mathscr{L}_1 f(\eta)$, relying on the values for the rates given in Table 1, and using (6.25).

$$\begin{aligned}
\mathscr{L}_1(\mathbf{1}_{\{\eta(0)=\mathtt{a}\}}) &= -\mathbf{1}_{\{\eta(0)=\mathtt{a}\}}(v_\mathtt{t}+v_\mathtt{c}+w_\mathtt{g}) - \sum_{a\in Y}\mathbf{1}_{\{(\eta(-1),\eta(0))=(a,\mathtt{a})\}} r^a_\mathtt{g} \\
&\quad +\mathbf{1}_{\{\eta(0)\in Y\}} v_\mathtt{a} + \mathbf{1}_{\{\eta(0)=\mathtt{g}\}} w_\mathtt{a} + \sum_{a\in Y}\mathbf{1}_{\{(\eta(-1),\eta(0))=(a,\mathtt{g})\}} r^a_\mathtt{a}, \qquad (6.27)\\
\mathscr{L}_1(\mathbf{1}_{\{\eta(0)=\mathtt{g}\}}) &= -\mathbf{1}_{\{\eta(0)=\mathtt{g}\}}(w_\mathtt{a}+v_\mathtt{t}+v_\mathtt{c}) - \sum_{a\in Y}\mathbf{1}_{\{(\eta(-1),\eta(0))=(a,\mathtt{g})\}} r^a_\mathtt{a} \\
&\quad +\mathbf{1}_{\{\eta(0)\in Y\}} v_\mathtt{g} + \mathbf{1}_{\{\eta(0)=\mathtt{a}\}} w_\mathtt{g} + \sum_{a\in Y}\mathbf{1}_{\{(\eta(-1),\eta(0))=(a,\mathtt{a})\}} r^a_\mathtt{g}, \qquad (6.28)\\
\mathscr{L}_1(\mathbf{1}_{\{\eta(0)=\mathtt{c}\}}) &= -\mathbf{1}_{\{\eta(0)=\mathtt{c}\}}(v_\mathtt{a}+v_\mathtt{g}+w_\mathtt{t}) - \sum_{a\in R}\mathbf{1}_{\{(\eta(0),\eta(1))=(\mathtt{c},a)\}} r^a_\mathtt{t} \\
&\quad +\mathbf{1}_{\{\eta(0)\in R\}} v_\mathtt{c} + \mathbf{1}_{\{\eta(0)=\mathtt{t}\}} w_\mathtt{c} + \sum_{a\in R}\mathbf{1}_{\{(\eta(0),\eta(1))=(\mathtt{t},a)\}} r^a_\mathtt{c}, \qquad (6.29)\\
\mathscr{L}_1(\mathbf{1}_{\{\eta(0)=\mathtt{t}\}}) &= -\mathbf{1}_{\{\eta(0)=\mathtt{t}\}}(v_\mathtt{a}+w_\mathtt{c}+v_\mathtt{g}) - \sum_{a\in R}\mathbf{1}_{\{(\eta(0),\eta(1))=(\mathtt{t},a)\}} r^a_\mathtt{c} \\
&\quad +\mathbf{1}_{\{\eta(0)\in R\}} v_\mathtt{t} + \mathbf{1}_{\{\eta(0)=\mathtt{c}\}} w_\mathtt{t} + \sum_{a\in R}\mathbf{1}_{\{(\eta(0),\eta(1))=(\mathtt{c},a)\}} r^a_\mathtt{t}. \qquad (6.30)
\end{aligned}$$

We write $0 = \int \mathscr{L}_1 f(\eta) d\mu(\eta)$ starting from (6.27)–(6.30). This gives the following linear system, whose last line states that μ is a probability measure.

$$\begin{cases}
-(v_\mathrm{Y}+w_\mathtt{g})\mu(\eta(0)=\mathtt{a}) + w_\mathtt{a}\mu(\eta(0)=\mathtt{g}) + v_\mathtt{a}\mu(Y) & = r_\mathrm{R},\\
w_\mathtt{g}\mu(\eta(0)=\mathtt{a}) - (v_\mathrm{Y}+w_\mathtt{a})\mu(\eta(0)=\mathtt{g}) + v_\mathtt{g}\mu(Y) & = -r_\mathrm{R},\\
-(v_\mathrm{R}+w_\mathtt{t})\mu(\eta(0)=\mathtt{c}) + w_\mathtt{c}\mu(\eta(0)=\mathtt{t}) + v_\mathtt{c}\mu(R) & = r_\mathrm{Y},\\
w_\mathtt{t}\mu(\eta(0)=\mathtt{c}) - (v_\mathrm{R}+w_\mathtt{c})\mu(\eta(0)=\mathtt{t}) + v_\mathtt{t}\mu(R) & = -r_\mathrm{Y},\\
\mu(\eta(0)=\mathtt{a}) + \mu(\eta(0)=\mathtt{g}) & = \mu(R),\\
\mu(\eta(0)=\mathtt{c}) + \mu(\eta(0)=\mathtt{t}) & = \mu(Y),\\
\mu(\eta(0)=\mathtt{a}) + \mu(\eta(0)=\mathtt{g}) + \mu(\eta(0)=\mathtt{c}) + \mu(\eta(0)=\mathtt{t}) & = 1.
\end{cases} \qquad (6.31)$$

Combining the addition of lines 1 and 2 (taking into account lines 5 and 6) with the last line of (6.31) gives a system whose solution is (4.9)–(4.10). We then insert those values into (6.31). Solving the system composed by its lines 1 and 5 (resp. its lines 3 and 6) yields (4.11), (4.13) (resp. (4.12), (4.14)). □

Proof. (of Proposition 4.9). We assume that (4.3) is satisfied; the case where (4.2) is satisfied is similar, and its proof is left to the reader.

We go through the 3 steps of the proof of Proposition 4.1: Step 1 is still valid, with (6.2)–(6.3) which become

$$\begin{aligned}\mu_l(\eta(0)=\mathtt{a})+\mu_l(\eta(0)=\mathtt{g}) &= \mu_u(\eta(0)=\mathtt{a})+\mu_u(\eta(0)=\mathtt{g}), \qquad (6.32)\\ \mu_l(\eta(0)=\mathtt{c})+\mu_l(\eta(0)=\mathtt{t}) &= \mu_u(\eta(0)=\mathtt{c})+\mu_u(\eta(0)=\mathtt{t}). \qquad (6.33)\end{aligned}$$

because of (4.9)–(4.10). Thus, in Step 2, (6.4) is still valid, as well as (6.5)–(6.6). But combining them with (6.32)–(6.33) implies that

$$\mu_l(\eta(0)=\mathtt{t})=\mu_u(\eta(0)=\mathtt{t}) \quad \text{and} \quad \mu_l(\eta(0)=\mathtt{a})=\mu_u(\eta(0)=\mathtt{a}). \tag{6.34}$$

Finally, Step 3 is valid, which ends the proof. $\square$

Proof. (of Proposition 4.11). Because ν is a monotone coupling measure of μ_1 and μ_2, it satisfies (4.19) and (6.8). Thus we have that

$$\begin{cases}\mu_1(\eta(0)=\mathtt{c}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{c}),(\mathtt{c},\mathtt{t}),(\mathtt{c},\mathtt{a}),(\mathtt{c},\mathtt{g})\}),\\ \mu_2(\xi(0)=\mathtt{c}) &= \nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{c})),\end{cases} \tag{6.35}$$

$$\begin{cases}\mu_1(\eta(0)=\mathtt{t}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{t},\mathtt{t}),(\mathtt{t},\mathtt{a}),(\mathtt{t},\mathtt{g})\}),\\ \mu_2(\xi(0)=\mathtt{t}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{t}),(\mathtt{t},\mathtt{t})\}),\end{cases} \tag{6.36}$$

$$\begin{cases}\mu_1(\eta(0)=\mathtt{a}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{a},\mathtt{a}),(\mathtt{a},\mathtt{g})\}),\\ \mu_2(\xi(0)=\mathtt{a}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{a}),(\mathtt{t},\mathtt{a}),(\mathtt{a},\mathtt{a})\}),\end{cases} \tag{6.37}$$

$$\begin{cases}\mu_1(\eta(0)=\mathtt{g}) &= \nu((\eta(0),\xi(0))=(\mathtt{g},\mathtt{g})),\\ \mu_2(\xi(0)=\mathtt{g}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{g}),(\mathtt{t},\mathtt{g}),(\mathtt{a},\mathtt{g}),(\mathtt{g},\mathtt{g})\}).\end{cases} \tag{6.38}$$

Hence

$$\begin{aligned}\mu_1(\eta(0)\in\{\mathtt{a},\mathtt{g}\}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{a},\mathtt{a}),(\mathtt{a},\mathtt{g}),(\mathtt{g},\mathtt{g})\}),\\ \mu_2(\xi(0)\in\{\mathtt{a},\mathtt{g}\}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{a}),(\mathtt{t},\mathtt{a}),(\mathtt{a},\mathtt{a}),(\mathtt{c},\mathtt{g}),(\mathtt{t},\mathtt{g}),(\mathtt{a},\mathtt{g}),(\mathtt{g},\mathtt{g})\}.)\end{aligned}$$

We then conclude thanks to (4.9)–(4.10) that yield

$$\mu_1(\eta(0)\in\{\mathtt{a},\mathtt{g}\})=\mu_2(\xi(0)\in\{\mathtt{a},\mathtt{g}\})=\frac{v_{\mathrm{R}}}{v_{\mathrm{Y}}+v_{\mathrm{R}}}\,.$$

Finally, (6.35)–(6.38) can be simplified in

$$\begin{cases}\mu_1(\eta(0)=\mathtt{c}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{c}),(\mathtt{c},\mathtt{t})\}),\\ \mu_2(\xi(0)=\mathtt{c}) &= \nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{c})),\\ \mu_1(\eta(0)=\mathtt{t}) &= \nu((\eta(0),\xi(0))=(\mathtt{t},\mathtt{t})),\\ \mu_2(\xi(0)=\mathtt{t}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{c},\mathtt{t}),(\mathtt{t},\mathtt{t})\}),\\ \mu_1(\eta(0)=\mathtt{g}) &= \nu((\eta(0),\xi(0))=(\mathtt{g},\mathtt{g})),\\ \mu_2(\xi(0)=\mathtt{g}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{a},\mathtt{g}),(\mathtt{g},\mathtt{g})\}),\\ \mu_1(\eta(0)=\mathtt{a}) &= \nu((\eta(0),\xi(0))\in\{(\mathtt{a},\mathtt{a}),(\mathtt{a},\mathtt{g})\}),\\ \mu_2(\xi(0)=\mathtt{a}) &= \nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{a})).\end{cases} \tag{6.39}$$

□

Proof. *(of Proposition 4.12).*

• We compute $\overline{\mathscr{L}}_1 f(\eta,\xi)$ for the function $f(\eta,\xi)=\mathbf{1}_{\{\eta(0)=a,\xi(0)=b\}}$, for $a,b\in\mathscr{A}$. We will then apply this computation to $(a,b)=(\mathtt{a},\mathtt{g})$ and to $(a,b)=(\mathtt{c},\mathtt{t})$. We have (recall the notation (6.20))

$$\begin{aligned}
\overline{\mathscr{L}}_1 f(\eta,\xi) &= \sum_{(a',b')\in\mathscr{A}\times\mathscr{A}} \overline{c}((a',b'),(\eta,\xi))\left[f(\eta^0_{a'},\xi^0_{b'})-f(\eta,\xi)\right]\\
&= -\sum_{(a',b')\neq(a,b)} \overline{c}((a',b'),(\eta,\xi))f(\eta,\xi)+\overline{c}((a,b),(\eta,\xi))\left[1-f(\eta,\xi)\right]\\
&= -\sum_{(a',b')\in\mathscr{A}\times\mathscr{A}} \overline{c}((a',b'),(\eta,\xi))f(\eta,\xi)+\overline{c}((a,b),(\eta,\xi)) \qquad (6.40)
\end{aligned}$$

where we used

$$f(\eta^0_{a'},\xi^0_{b'})=\mathbf{1}_{\{\eta^0_{a'}(0)=a,\xi^0_{b'}(0)=b\}}=\mathbf{1}_{\{a'=a,b'=b\}}.$$

• We now apply (6.40) to $(a,b)=(\mathtt{a},\mathtt{g})$, according to the coupling rates given in the tables in the proof of Proposition 4.4, then using the attractiveness conditions from Proposition 4.4.

$$\begin{aligned}
\overline{\mathscr{L}}_1 f(\eta,\xi) &= -\mathbf{1}_{\{\eta(0)=\mathtt{a},\xi(0)=\mathtt{g}\}}\left[c(\mathtt{a},\xi)+c(\mathtt{g},\eta)+v_{\mathtt{t}}+v_{\mathtt{c}}\right]\\
&\quad+\mathbf{1}_{\{\eta(0)=\mathtt{a},\xi(0)=\mathtt{a}\}}\left[c(\mathtt{g},\xi)-\min\left\{c(\mathtt{g},\eta),c(\mathtt{g},\xi)\right\}\right]\\
&\quad+\mathbf{1}_{\{\eta(0)=\mathtt{g},\xi(0)=\mathtt{g}\}}\left[c(\mathtt{a},\eta)-\min\left\{c(\mathtt{a},\eta),c(\mathtt{a},\xi)\right\}\right]\\
&= -\mathbf{1}_{\{\eta(0)=\mathtt{a},\xi(0)=\mathtt{g}\}}\left[w_{\mathtt{a}}+\mathbf{1}_{\{\xi(-1)=a\in Y\}}r^a_{\mathtt{a}}+w_{\mathtt{g}}+0+v_{\mathtt{t}}+v_{\mathtt{c}}\right]+0\\
&\quad+\mathbf{1}_{\{\eta(0)=\mathtt{g},\xi(0)=\mathtt{g}\}}\left[\mathbf{1}_{\{\eta(-1)=a\in Y\}}r^a_{\mathtt{a}}-\min\left\{\mathbf{1}_{\{\eta(-1)=a\in Y\}}r^a_{\mathtt{a}},\mathbf{1}_{\{\xi(-1)=b\in Y\}}r^b_{\mathtt{a}}\right\}\right]\\
&= -\mathbf{1}_{\{\eta(0)=\mathtt{a},\xi(0)=\mathtt{g}\}}\left[v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}+\mathbf{1}_{\{\xi(-1)=\mathtt{c}\}}r^{\mathtt{c}}_{\mathtt{a}}+\mathbf{1}_{\{\xi(-1)=\mathtt{t}\}}r^{\mathtt{t}}_{\mathtt{a}}\right]\\
&\quad+\mathbf{1}_{\{\eta(-1)=\mathtt{c},\xi(-1)=\mathtt{t},\eta(0)=\mathtt{g},\xi(0)=\mathtt{g}\}}\left[r^{\mathtt{c}}_{\mathtt{a}}-r^{\mathtt{t}}_{\mathtt{a}}\right]. \qquad (6.41)
\end{aligned}$$

For the last term in the last equality, we have used that since $\eta(-1)\leqslant\xi(-1)$, either they are both equal, or $\eta(-1)=\mathtt{c},\xi(-1)=\mathtt{t}$.
We now integrate (6.41) with respect to ν, using also translation invariance of ν. To lighten the formulas, we use the notation, for $x,y\in\mathbb{Z}$, $a_1,a_2,b_1,b_2\in\mathscr{A}$,

$$\nu\begin{pmatrix} x & y\\ a_1 & a_2\\ b_1 & b_2\end{pmatrix}=\nu(\eta(x)=a_1,\eta(y)=a_2,\xi(x)=b_1,\xi(y)=b_2)$$

as well as

$$\nu\begin{pmatrix} x\\ a_1\\ b_1\end{pmatrix}=\nu(\eta(x)=a_1,\xi(x)=b_1),\quad \nu\begin{pmatrix} x & y\\ a_1 & *\\ b_1 & b_2\end{pmatrix}=\nu(\eta(x)=a_1,\xi(x)=b_1,\xi(y)=b_2).$$

We also use that by Proposition 4.11 we have

$$\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}=\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & *\end{pmatrix},\quad \nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix}=\nu\begin{pmatrix}0 & 1\\ * & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix},\quad \nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}=\nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ * & \mathtt{g}\end{pmatrix}.$$

We thus get

$$\begin{aligned}0=\int\overline{\mathscr{L}}_1 f(\eta,\xi)d\nu &= -[v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{a}\\ \mathtt{g}\end{pmatrix}-r_{\mathtt{a}}^{\mathtt{t}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}-r_{\mathtt{a}}^{\mathtt{c}}\,\nu\begin{pmatrix}0 & 1\\ * & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix}\\ &\quad -r_{\mathtt{a}}^{\mathtt{t}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ * & \mathtt{g}\end{pmatrix}+[r_{\mathtt{a}}^{\mathtt{c}}-r_{\mathtt{a}}^{\mathtt{t}}]\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & *\end{pmatrix}.\end{aligned}\tag{6.42}$$

• We similarly apply (6.40) to $(a,b)=(\mathtt{c},\mathtt{t})$, then also integrate with respect to ν, using translation invariance of ν. This gives

$$\begin{aligned}0=\int\overline{\mathscr{L}}_1 f(\eta,\xi)d\nu &= -[w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{c}\\ \mathtt{t}\end{pmatrix}-r_{\mathtt{t}}^{\mathtt{a}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}-r_{\mathtt{t}}^{\mathtt{a}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & *\\ \mathtt{t} & \mathtt{a}\end{pmatrix}\\ &\quad -r_{\mathtt{t}}^{\mathtt{g}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & *\end{pmatrix}+[r_{\mathtt{t}}^{\mathtt{g}}-r_{\mathtt{t}}^{\mathtt{a}}]\,\nu\begin{pmatrix}0 & 1\\ * & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix}.\end{aligned}\tag{6.43}$$

• Adding (6.42) and (6.43) gives:

$$\begin{aligned}0 &= -[v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{a}\\ \mathtt{g}\end{pmatrix}+[(r_{\mathtt{t}}^{\mathtt{g}}-r_{\mathtt{t}}^{\mathtt{a}})-r_{\mathtt{a}}^{\mathtt{c}}]\,\nu\begin{pmatrix}0 & 1\\ * & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix}\\ &\quad -[w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{c}\\ \mathtt{t}\end{pmatrix}+[(r_{\mathtt{a}}^{\mathtt{c}}-r_{\mathtt{a}}^{\mathtt{t}})-r_{\mathtt{t}}^{\mathtt{g}}]\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & *\end{pmatrix}\\ &\quad -[r_{\mathtt{a}}^{\mathtt{t}}+r_{\mathtt{t}}^{\mathtt{a}}]\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}-r_{\mathtt{a}}^{\mathtt{t}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ * & \mathtt{g}\end{pmatrix}-r_{\mathtt{t}}^{\mathtt{a}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & *\\ \mathtt{t} & \mathtt{a}\end{pmatrix}.\end{aligned}\tag{6.44}$$

• We claim that

$$\nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=0\quad\text{if and only if}\quad\nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{t}))=0.\tag{6.45}$$

Indeed, assuming that $\nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=0$, we have also

$$\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix}=\nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ * & \mathtt{g}\end{pmatrix}=\nu\begin{pmatrix}0 & 1\\ * & \mathtt{a}\\ \mathtt{c} & \mathtt{g}\end{pmatrix}=0.$$

Hence (6.43) becomes

$$0 \;=\; -[w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{c}\\ \mathtt{t}\end{pmatrix} - r_{\mathtt{t}}^{\mathtt{a}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & *\\ \mathtt{t} & \mathtt{a}\end{pmatrix} - r_{\mathtt{t}}^{\mathtt{g}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{g}\\ \mathtt{t} & *\end{pmatrix}.$$

The above r.h.s. contains only non-positive terms: it means that each of them is equal to 0, in particular the first one, for which we know that $w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}>0$ (note that for the second and third terms, $r_{\mathtt{t}}^{\mathtt{a}}$ and/or $r_{\mathtt{t}}^{\mathtt{g}}$ could be equal to 0). This implies that $\nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{t}))=0$.
Similarly, assuming that $\nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{t}))=0$, (6.42) becomes

$$0=-[v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}]\,\nu\begin{pmatrix}0\\ \mathtt{a}\\ \mathtt{g}\end{pmatrix} - r_{\mathtt{a}}^{\mathtt{t}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{c} & \mathtt{a}\\ \mathtt{t} & \mathtt{g}\end{pmatrix} - r_{\mathtt{a}}^{\mathtt{t}}\,\nu\begin{pmatrix}0 & 1\\ \mathtt{t} & \mathtt{a}\\ * & \mathtt{g}\end{pmatrix}.$$

Since $v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}>0$, this implies that $\nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=0$.

• We then consider each of the 3 assumptions on the rates.

(a) Assuming that $r_{\mathtt{a}}^{\mathtt{c}}=r_{\mathtt{a}}^{\mathtt{t}}$, we have that the r.h.s. in (6.42) contains only non-positive terms: it means that each of them is equal to 0, in particular the first one, for which $v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}>0$. This implies that $\nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=0$, and we conclude thanks to (6.45).

(b) Assuming that $r_{\mathtt{t}}^{\mathtt{g}}=r_{\mathtt{t}}^{\mathtt{a}}$ induces a similar reasoning, to get first $\nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{t}))=0$ by considering (6.43) and using that $w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}>0$, then conclude thanks to (6.45).

(c), *(i)* As a preliminary, we examine the conditions (α), (β), (γ), (δ): It is impossible to have neither (α) nor (γ) satisfied, since it would imply

$$r_{\mathtt{t}}^{\mathtt{g}}-r_{\mathtt{t}}^{\mathtt{a}}>r_{\mathtt{a}}^{\mathtt{c}}>r_{\mathtt{t}}^{\mathtt{g}}+r_{\mathtt{a}}^{\mathtt{t}} \quad \text{hence} \quad -r_{\mathtt{t}}^{\mathtt{a}}>r_{\mathtt{a}}^{\mathtt{t}}.$$

Thus if one of them is not satisfied, the other automatically is.
If (β) is satisfied then (α) is not satisfied, hence (γ) is satisfied.
Similarly, if (δ) is satisfied then (γ) is not satisfied, hence (α) is satisfied.

(ii) If (α) and (γ) are satisfied, then the r.h.s. of (6.44) contains only non-positive terms, hence each of them is equal to 0. Since $v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}>0$ and $w_{\mathtt{t}}+w_{\mathtt{c}}+v_{\mathtt{g}}+v_{\mathtt{a}}>0$, this implies that $\nu((\eta(0),\xi(0))=(\mathtt{a},\mathtt{g}))=\nu((\eta(0),\xi(0))=(\mathtt{c},\mathtt{t}))=0$.

(iii) If (β) is satisfied, on the one hand we bound the sum of the first two terms in the r.h.s. of (6.44) by

$$\left(-[v_{\mathtt{t}}+v_{\mathtt{c}}+w_{\mathtt{g}}+w_{\mathtt{a}}]+[(r_{\mathtt{t}}^{\mathtt{g}}-r_{\mathtt{t}}^{\mathtt{a}})-r_{\mathtt{a}}^{\mathtt{c}}]\right)\,\nu\begin{pmatrix}0\\ \mathtt{a}\\ \mathtt{g}\end{pmatrix},$$

and on the other hand, by *(i)*, (γ) is satisfied. Hence the r.h.s. of (6.44), which is equal to 0, is bounded by a sum of only non-positive terms, thus each of them is equal to 0. Since $w_{\mathtt{t}} + w_{\mathtt{c}} + v_{\mathtt{g}} + v_{\mathtt{a}} > 0$, this implies that $\nu((\eta(0), \xi(0)) = (\mathtt{c}, \mathtt{t})) = 0$. We conclude thanks to (6.45).

(iv) Similarly, if (δ) is satisfied, on the one hand we bound the sum of the third and fourth terms in the r.h.s. of (6.44) by

$$(-[w_{\mathtt{t}} + w_{\mathtt{c}} + v_{\mathtt{g}} + v_{\mathtt{a}}] + [(r_{\mathtt{a}}^{\mathtt{c}} - r_{\mathtt{a}}^{\mathtt{t}}) - r_{\mathtt{t}}^{\mathtt{g}}]) \ \nu \begin{pmatrix} 0 \\ \mathtt{c} \\ \mathtt{t} \end{pmatrix},$$

and on the other hand, by *(i)*, (α) is satisfied. Hence the r.h.s. of (6.44), which is equal to 0, is bounded by a sum of only non-positive terms, thus each of them is equal to 0. Since $v_{\mathtt{t}} + v_{\mathtt{c}} + w_{\mathtt{g}} + w_{\mathtt{a}} > 0$, this implies that $\nu((\eta(0), \xi(0)) = (\mathtt{a}, \mathtt{g})) = 0$. We conclude thanks to (6.45).

The proposition is proved. $\square$

Acknowledgements. We thank the referee for carefully reading the first version of this text, and for giving us several helpful comments. We thank both MAP5 lab. and TU München for financial support and hospitality.

References

[AFS04] Enrique Andjel, Pablo A. Ferrari, and Adriano Siqueira. Law of large numbers for the simple exclusion process. *Stochastic processes and their applications*, 113:217–233, 2004.

[BG12] Jean Bérard and Laurent Guéguen. Accurate phylogenetic estimation of substitution rates with context-dependent models. *Systematic Biology*, 6(3):510–521, 2012.

[BGP08] Jean Bérard, Jean-Baptiste Gouéré, and Didier Piau. Solvable models of neighbor-dependent nucleotide substitution processes. *Mathematical Biosciences*, 211:56–88, 2008.

[Bir80] Adrian P. Bird. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research*, 8:1499–1504, 1980.

[Bor11] Davide Borrello. Stochastic order and attractiveness for particle systems with multiple births, deaths and jumps. *Electron. J. Probab.*, 16(4):106–151, 2011.

[DG00] Laurent Duret and Nicolas Galtier. The covariation between TpA deficiency, CpG deficiency, and G+C content of human isochores is due to a mathematical artifact. *Molecular Biology and Evolution*, 17:1620–1625, 2000.

[Dob71] Roland L. Dobrushin. Markov processes with a large number of locally interacting components: Existence of a limit process and its ergodicity. *Problems Inform. Transmission*, 7:149–164, 1971.

[Dur93] Rick Durrett. *Lectures on probability theory, École d'été de probabilités de Saint-Flour XXIII-1993*, chapter Ten Lectures on Particle Systems, pages 97–201. Springer, 1993.

[Fal10] Mikael Falconnet. Phylogenetic distances for neighbour dependent substitution processes. *Mathematical Biosciences*, 224(2):101–108, 2010.

[Fel81] Joseph Felsenstein. Evolutionary trees from DNA sequences : A maximum likelihood approach. *J. Mol. Evol.*, 17:368–376, 1981.

[Fer90] Pablo A. Ferrari. Ergodicity for spin systems with stirrings. *The Annals of Probability*, 18(4):1523–1538, 1990.

[Fer91] Pablo A. Ferrari. Ergodicity for a class of probabilistic cellular automata. *Rev. Mat. Apl.*, 12:93–102, 1991.

[Fin89] David J. Finnegan. Eukaryotic tranposable elements and genome evolution. *Trends Genet.*, 5:103–107, 1989.

[Gra82] Lawrence F. Gray. The positive rates problem for attractive nearest neighbor spin systems on **Z**. *Z. Wahrsch. Verw. Gebiete*, 61(3):389–404, 1982.

[Gra01] Lawrence F. Gray. A reader's guide to P. Gács's "positive rates" paper: "Reliable cellular automata with self-organization" [J. Statist. Phys. **103** (2001), no. 1-2, 45–267; MR1828729 (2002c:82058a)]. *J. Statist. Phys.*, 103(1-2):1–44, 2001.

[GS10] Thierry Gobron and Ellen Saada. Couplings, attractiveness and hydrodynamics for conservative particle systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 46(4):1132–1177, 2010.

[Har72] Ted E. Harris. Nearest-neighbor Markov interaction processes on multidimensional lattices. *Adv. Math*, 9:66–89, 1972.

[HKY85] Masahito Hasegawa, Hirohisa Kishino, and Taka-aki Yano. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, 22:160–174, 1985.

[JC69] Thomas Hughes Jukes and Charles R. Cantor. *Mammalian protein metabolism*, chapter Evolution of Protein Molecules, pages 21–132. Academic Press, New York, 1969.

[JKK61] John Josse, Armin Dale Kaiser, and Arthur Kornberg. Enzymatic synthesis of deoxyribonucleic acid. VIII. frequencies of nearest neighbor base sequences in deoxyribonucleic acid. *J. Biol. Chem.*, 236:864–875, 1961.

[Kim80] Motoo Kimura. A Simple Method for Estimating Evolutionary Rates of Base Subsitutions Through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol.*, 10:111–120, 1980.

[Kov05] Yevgeniy Kovchegov. Exclusion processes with multiple interactions. *Stochastic processes and their applications*, 115:1233–1256, 2005.

[Lig05] Thomas M. Liggett. *Interacting Particle Systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[McC53] Barbara McClintock. Induction of instability at selected loci in maize. *Genetics*, 38(6):579–599, 1953.

[RN95] Andrey Rzhetsky and Masatoshi Nei. Tests of applicability of several substitution models for DNA sequence data. *Mol. Biol. Evol.*, 12:131–51, 1995.

[STK62] Morton N. Swartz, Thomas A. Trautner, and Arthur Kornberg. Enzymatic synthesis of deoxyribonucleic acid. XI. further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J. Biol. Chem.*, 236:1961–1967, 1962.

[Tam92] Koichiro Tamura. Estimation of the number of nucleotide substitutions when there are strong transition/transversion and g+c content biases. *Molecular Biology and Evolution*, 9:678–687, 1992.

[Tav86] Simon Tavaré. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, 17:57–86, 1986.

[TKF91] Jeffrey L. Thorne, Hirohisa Kishino, and Joseph Felsenstein. An evolutionary model for maximum likelihood alignment of DNA sequences. *Journal of Molecular Evolution*, 33(2):114–124, 1991.

[TN93] Koichiro Tamura and Masatoshi Nei. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10:512–526, 1993.

[Yan94] Ziheng Yang. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.*, 39:105–111, 1994.