Automates cellulaires probabilistes et mesures spécifiques sur des espaces symboliques

Irène Marcovici

Université Paris Diderot (Paris 7)

Sorbonne Paris Cité

École Doctorale Sciences Mathématiques de Paris Centre

Laboratoire d'Informatique Algorithmique : Fondements et Applications

Irène Marcovici

Automates cellulaires probabilistes et mesures spécifiques sur des espaces symboliques

Thèse dirigée par Jean Mairesse

Soutenue le 22 novembre 2013,

en vue d'obtenir le grade de

Docteur de l'Université Paris Diderot,

Spécialité Mathématiques Appliquées

Jury

Directeur de thèse :	Jean Mairesse
Rapporteurs :	Péter Gács (non présent à la soutenance) Jarkko Kari Ellen Saada
Examinateurs :	Marie-Pierre Béal Francis Comets Élise Janvresse

Remerciements

Mes remerciements vont tout d'abord à Jean Mairesse. Il m'a proposé un sujet de recherche ouvrant la voie à de nombreux développements, et m'a encadrée de manière attentive tout au long de ma thèse. J'ai apprécié les moments que nous avons passés à réfléchir ensemble. Ces séances de travail s'accompagnaient d'une grande liberté de recherche et m'ont permis d'acquérir progressivement une réelle autonomie, tout en goûtant à la stimulation de la recherche en groupe et aux joies des découvertes partagées. En plus de son suivi scientifique, qu'il a ainsi effectué avec soin tout en me laissant beaucoup d'indépendance, Jean a veillé à mon intégration aux communautés mathématiques et informatiques, et a su encourager mes vocations de "globe-matheuse". Pour toutes ces raisons, je tiens à lui exprimer ma plus profonde gratitude.

Péter Gács, Jarkko Kari, et Ellen Saada ont accepté d'être rapporteurs de ma thèse. J'en suis très honorée et je les remercie vivement d'avoir pris le temps de s'intéresser à mes travaux. Leurs commentaires m'ont permis d'améliorer mon manuscrit et me seront précieux pour de futures directions de recherche. Je remercie également Marie-Pierre Béal, Francis Comets et Élise Janvresse d'avoir accepté d'être examinateurs de ma thèse. Ce jury, constitué de mathématicienn-ne-s et d'informaticien-ne-s, reflète la pluridisciplinarité de ma thèse, effectuée au sein d'un laboratoire d'informatique alors que j'avais initialement une formation à dominante mathématique, sur un thème résolument transversal.

Ana Bušić a beaucoup contribué à ce que ma thèse débute sous les meilleurs auspices. Je me rappelerai des échanges avec elle et Jean par webcam depuis ma chambre étudiante d'Heidelberg, pour poursuivre les travaux de mon stage de Master. Nous nous sommes ensuite retrouvées en différentes occasions, pour d'intenses séances de travail ou lors de conférences. Les températures extérieures étaient variées (-10 °C à Turku en décembre 2010, 30 °C à Stockholm en juin 2011) mais nos échanges toujours chaleureux.

Avec Nazim Fatès, les séances de travail communes se sont poursuivies dans la bonne humeur, et m'ont insufflé de l'énergie pour approfondir de nouvelles pistes de recherche. En plus d'être très riches scientifiquement, nos rencontres étaient l'occasion de discussions philosophiques et de partages culturels, que ce soit à Paris ou ailleurs, comme à Cologne avant la conférence *STACS*.

J'ai effectué un séjour de recherche à Santiago du Chili, de mars à mai 2012, pour travailler avec Alejandro Maass. Je lui suis extrêmement reconnaissante de m'avoir donné cette opportunité, et je le remercie vivement ainsi que tous les membres du CMM pour leur chaleureux accueil. Au cours de ce séjour, j'ai également eu l'occasion de travailler avec Alexis Ballier, grâce à qui j'ai approfondi mes connaissances de dynamique symbolique, et aussi des spécialités chiliennes !

De retour en France, j'ai été contente de prolonger avec Mathieu Sablik et Benjamin Hellouin de Menibus les travaux commencés avec Alejandro. Mathieu m'a initiée aux systèmes dynamiques discrets lors d'un groupe de lecture dont il avait la charge à l'ENS Lyon, puis pendant mon premier stage de recherche, effectué sous sa direction à Marseille, au cours de l'été 2008. Le fait que j'aie continué dans cette voie témoigne de l'importance qu'ont eue pour moi ces premières expériences, et du plaisir que j'ai eu à découvrir ce domaine grâce à lui.

Je remercie Pierre-Yves Louis pour nos riches échanges à Poitiers puis à Amsterdam, à la veille du *Workshop on Probabilistic Cellular Automata* de juin 2013. Cet événement, dont il a porté l'initiative, a contribué de manière fondamentale au rapprochement des différentes communautés travaillant sur les automates cellulaires probabilistes.

C'est également un plaisir de travailler avec Vincent Delecroix, grâce à sa grande curiosité et à son enthousiasme. Il m'a aussi permis de me familiariser davantage avec *Sage*.

En juillet 2013, j'ai eu l'opportunité d'initier une collaboration avec James Martin à l'Université d'Oxford. Nos premiers échanges m'ont beaucoup apporté et je me réjouis d'avoir prochainement l'occasion de poursuivre les travaux que nous avons commencés ensemble.

Je suis membre des groupes de travail SDA2 et ALEA du GDR Informatique Mathématique. L'appartenance à ces communautés a beaucoup contribué à mon épanouissement en tant que doctorante. Les différentes rencontres annuelles organisées sont des moments importants du calendrier, toujours très enrichissantes scientifiquement et extrêmement sympathiques. C'est le cas aussi des réunions de l'ANR Magnum.

En septembre 2013, j'ai également eu la chance de participer au premier *Heidelberg Laureate Forum*. Je remercie à nouveau Alejandro, Francis et Jean, grâce à qui j'ai eu l'opportunité de participer à cette rencontre, qui restera comme un moment fort de la fin de ma thèse.

Au LIAFA, j'ai bénéficié d'un environnement de recherche extrêmement favorable, très riche à la fois sur le plan scientifique et sur le plan humain.

Je remercie chaleureusement Valérie Berthé, dont le soutien a beaucoup compté pour moi. J'ai été heureuse d'organiser le Forum des Jeunes Mathématicien-ne-s à ses côtés en novembre 2012, et c'est grâce à elle que j'ai ensuite rejoint l'association *femmes&mathématiques*.

De la rue du Chevaleret au bâtiment Sophie Germain, j'ai eu beaucoup de plaisir à partager un bureau avec Adeline, Antoine, Bruno, Cezara, Jehanne, Jérémie, Luc, Robin et Sandrine. J'ai apprécié le grand calme qui y régnait, ainsi que les débats animés qui l'interrompaient parfois, des sujets les plus loufoques aux questions existentielles. Bruno faisait aussi partie de l'expédition au Canigou (ou presque) après l'EJC de Perpignan, et je dois une spéciale dédicace à Sandrine de m'avoir fait découvrir la Côte de granit rose.

Je n'oublie pas non plus les autres membres du LIAFA, doctorant-e-s et permanent-e-s, que j'ai cotoyés tout au long de ma thèse. Déjeuners ou discussions au coin café, sorties piscines... autant de moments agréables passés ensemble. Les responsables gâteaux d'un jour ou de toujours contribuent de façon essentielle à la convivialité du laboratoire, et les séminaires thésards sont aussi des occasions sympathiques de réunir doctorants du LIAFA et de PPS.

Avec Antoine Taveneaux et Timo Jolivet, nous avons effectué simultanément la dernière ligne droite vers la soutenance, en partageant nos expériences et en nous soutenant mutuellement. On m'avait prédit un été enfermée à rédiger, mais au final, la rédaction n'a pas été un calvaire du tout, et elle ne m'a pas interdit une petite trêve estivale.

Nous devons aussi beaucoup à Noëlle Delgado et à Nathalie Rousseau. Leur présence est indispensable à la bonne marche du laboratoire, et grâce à leur efficacité, les démarches administratives n'ont jamais été un souci pour moi.

A l'Université Paris Diderot, j'ai enseigné le TD de mathématiques du cours de Jérôme Dubois, puis les TD et TP Java du cours d'algorithmique et programmation de Stéphane

Boucheron. J'ai beaucoup apprécié de travailler avec Jérôme et Stéphane, aux côtés des autres chargés de TD. Leurs enseignements, à la fois exigeants et pragmatiques, m'ont permis d'intervenir dans les meilleures conditions devant les étudiants, dont j'ai également apprécié l'attitude globalement positive.

Au cours de ma scolarité, du collège jusqu'à l'ENS Lyon, j'ai eu la chance de rencontrer des enseignants exceptionnels, qui ont su stimuler ma curiosité intellectuelle et encourager mon goût pour les mathématiques. J'ai également bénéficié des activités de l'association *Animath*. Je remercie vivement les personnes qui ont ainsi développé mon attrait pour la recherche, et parmi elles Xavier Caruso, ainsi que celles grâce à qui j'ai maintenant le plaisir d'intervenir à mon tour devant de jeunes élèves passionnés par les mathématiques.

Je pense aussi à mes amis du lycée et de l'ENS Lyon, que j'ai toujours plaisir à retrouver régulièrement.

En M2, nous étions peu nombreux à suivre le parcours probabilités, et nous avons formé un petit groupe soudé avec Alexis, Elodie, Marielle et Nicolas, à suivre les précieux enseignements qui nous étaient dispensés.

J'ai participé aux *Doctoriales* en septembre 2012 et j'apprécie également de revoir les doctorants de toutes disciplines que j'y ai rencontrés.

Au cours de mes trois années de thèse, j'ai été membre de plusieurs associations, militantes, culturelles, sportives. De nombreux kilomètres parcourus, en marchant ou en courant, sur terrain plat ou escarpé. Des personnes, des lieux. Bastille, Nation, Cité U, Hôtel de Ville, Porte Dorée, Laumière ou Jaurès. La Plagne, Mijoux, Marseille, Arras. Valais, Emmental. Beauregard, Montgardier. Des chemins qui se croisent, s'éloignent, se rejoignent. L'émotion reste vive.

Ces quelques mots sont insuffisants pour remercier comme je le souhaite mes collègues et ami-e-s. Ils le sont encore davantage pour remercier mes parents et mes sœurs. Mais ils n'ont pas besoin de remerciements pour savoir combien leur soutien est important, et je les en remercie !

Contents

Introduction

Cette thèse porte sur les automates cellulaires probabilistes et sur des mesures spécifiques sur des espaces symboliques.

Les espaces symboliques sont des ensembles de la forme $\mathcal{A}^E$, où $\mathcal{A}$ est un ensemble fini de symboles, et E un ensemble dénombrable, appelé l'ensemble des cellules. Ils apparaissent dans des contextes variés, et en particulier lors de la modélisation de phénomènes physiques et biologiques. Par exemple, dans le modèle d'Ising, qui est utilisé en mécanique statistique comme modèle mathématique du ferromagnétisme, un matériau est représenté par différents moments magnétiques ayant chacun deux états possibles, +1 ou -1, et disposés selon un graphe (qui a généralement une structure de réseau). En biologie, les espaces symboliques sont utilisés par exemple pour modéliser un ensemble de cellules qui peuvent être dans différents états (saine/infectée). Au-delà de la modélisation, dans les composants informatiques et électroniques, l'information est encodée par des configurations sur des espaces symboliques : une image numérique est ainsi constituée d'un ensemble de pixels disposés sur une grille, à qui sont attribués des couleurs, parmi un ensemble fini de couleurs possibles.

Nous nous intéressons à des mesures spécifiques sur des espaces symboliques. Par spécifique, nous entendons des mesures qui présentent des caractéristiques originales, par opposition à génériques. Les spécificités des mesures que nous considérons sont doubles.

D'une part, ces mesures ont des propriétés intrinsèques qui les rendent spéciales : elles ont une structure combinatoire particulière, mettant en jeu la topologie de l'ensemble des cellules sur lequel elles sont définies. Les mesures markoviennes auront en particulier un rôle fondamental.

D'autre part, ces mesures correspondent à des équilibres particuliers de processus stochastiques, marches aléatoires ou automates cellulaires probabilistes (ACP).

Un ACP est une chaîne de Markov sur un espace symbolique. Le temps est discret, et toutes les cellules évoluent de manière synchrone : le nouvel état de chaque cellule est choisi de manière aléatoire, indépendamment des autres cellules, selon une distribution déterminée par les états d'un nombre fini de cellules situées dans le voisinage. Les ACP sont de bons candidats pour modeliser les systèmes complexes intervenant dans des processus physiques ou biologiques, en raison du contraste entre la simplicité de leur définition et la complexité des comportements qu'ils engendrent. Ils sont utilisés pour explorer les modèles de calcul robustes aux erreurs. Enfin, ils interviennent dans différents contextes en probabilité et en combinatoire.

Considérons le cas particulier de l'ensemble des cellules $E = \mathbb{Z}$, l'alphabet $\mathcal{A} = \{0, 1\}$, et le voisinage constitué de la cellule elle-même et de sa voisine de droite (ou de manière équivalente, de la voisine de gauche et de la cellule elle-même). Alors, un ACP est entièrement déterminé par les quatre paramètres $(\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11})$, où $\theta_{ij} \in [0, 1]$ est la probabilité qu'une cellule soit mise à jour par un 1 si son voisinage est dans l'état ij. Considérons par exemple l'ACP défini par les quatre paramètres $(p, 0, 0, 0)$ pour un certain $p \in [0, 1]$, voir Figure 1.

Cet ACP peut être décrit de la manière suivante : si le voisinage d'une cellule est dans l'état 00, alors, avec probabilité p, la cellule est actualisée par un 1, et avec probabilité $1-p$, la cellule est actualisée par un 0. Dans les autres cas, la cellule est actualisée par un 0 (avec probabilité 1). Pour $p=0$ et $p=1$, il n'y a plus d'aléa : on obtient un automate cellulaire déterministe.

Les trajectoires d'un ACP sont représentées par des diagrammes espace-temps, qui vivent eux-mêmes sur des espaces symboliques, avec une dimension supplémentaire correspondant au temps. Sur la Figure 1, les cellules contenant un 0, resp. un 1, sont représentées en blanc, resp. en bleu. La ligne la plus basse est la condition initiale, choisie ici aléatoirement, et les lignes suivantes, de bas en haut, correspondent aux mises à jour successives des cellules.

Le comportement à l'équilibre d'un ACP est étudié par l'intermédiaire des mesures invariantes de la chaîne de Markov sur l'espace symbolique sur lequel il est défini. De nombreuses questions se posent. Un ACP est ergodique s'il a une unique mesure invariante, qui est attractive. Le problème de l'ergodicité des ACP est indécidable : il n'existe pas d'algorithme capable de dire, lorsqu'on lui fournit en argument les paramètres de l'ACP, s'il est ergodique ou non. Et on ne connaît pas d'outil général pour décrire les mesures invariantes d'un ACP.

Dans ce contexte, le problème des taux positifs est un véritable défi. On dit qu'un ACP a des taux strictement positifs si pour n'importe quelle valeur de son voisinage, une cellule peut être mise à jour par n'importe quel symbole avec une probabilité positive. En dimension deux, il existe des exemples simples d'ACP à taux strictement positifs qui ne sont pas ergodiques, mais pour les ACP unidimensionnels, le seul exemple connu a été proposé en 2001 par Gács (après une première publicaion en 1986), et il est très complexe. En dimension un, si on se limite aux ACP ayant un voisinage de taille 2, et définis sur un ensemble de symboles de taille 2, on ne sait pas si tous les ACP à taux strictement positifs (c'est-à-dire, les ACP tels que $\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11} \in]0,1[$) sont ergodiques.

Des outils issus de la mécanique statistique ont été développés pour étudier les ACP à taux strictement positifs. Dans ce cas, la recherche de mesures invariantes est équivalente à un problème de mécanique statistique à l'équilibre. Dans nos travaux, nous nous intéressons aussi aux ACP ayant des composantes déterministes. Certains outils de mécanique statistique peuvent être adaptés, mais l'analyse de ces ACP nécessite un soin tout particulier.

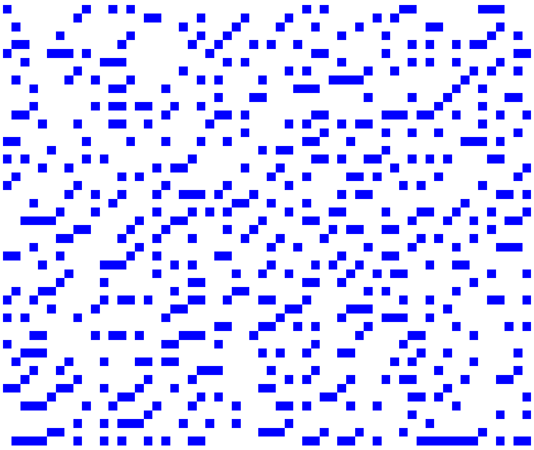
Lorsqu'il n'est pas possible d'obtenir des résultats exacts, il est naturel de se tourner vers la simulation. Simuler des ACP est un véritable défi, coûteux à la fois en temps et en espace. De plus, quand le nombre de cellules est infini, il n'est pas possible de manipuler des configurations complètes, celles-ci peuvent seulement être décrites au moyen de certaines quantités observables. Le point crucial est alors de savoir si l'on peut donner des garanties sur les résultats obtenus par simulation. En adaptant la méthode de couplage arrière de Propp et Wilson, nous proposons un algorithme permettant d'échantillonner parfaitement la mesure invariante d'un ACP ergodique, sous certaines conditions. L'algorithme est aléatoire et retourne une configuration (ou une portion de configuration) distribuée selon la mesure invariante, de telle sorte qu'en répétant la procédure, il est possible d'estimer la mesure invariante avec une précision arbitraire.

Dans des cas très particuliers, il est cependant possible de prévoir de manière théorique le comportement asymptotique d'un ACP. Par exemple, on sait caractériser les ACP ayant une mesure invariante de forme produit de Bernoulli. Nous montrons que dans ce cas, les diagrammes espace-temps définissent des mesures ayant de très faibles dépendances, qui partagent des propriétés particulières.

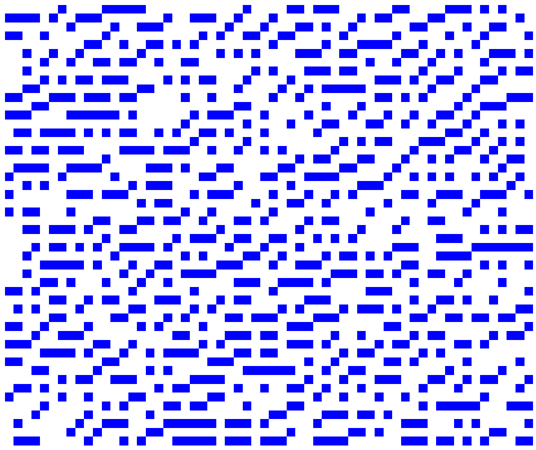
Le cas des AC déterministes ayant des mesures de Bernoulli invariantes est également intéressant. Puisque pour les AC déterministes, l'ergodicité est équivalente à la nilpotence, il est alors plus pertinent d'assouplir à la fois l'unicité de la mesure invariante et la propriété

1 avec probabilité p 0 avec probabilité $1-p$	0 (avec probabilité 1)
0 0	0 1 1 0 1 1

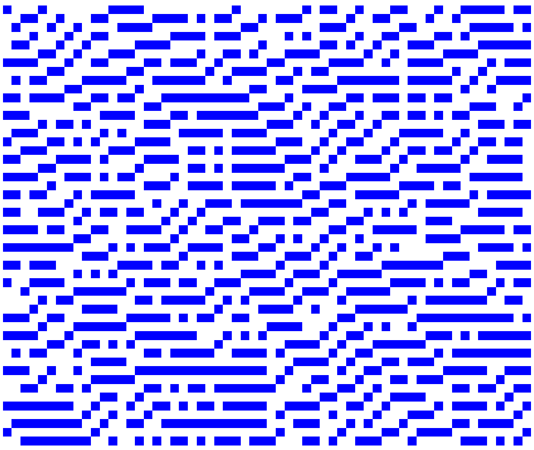
Règle locale de l'ACP.



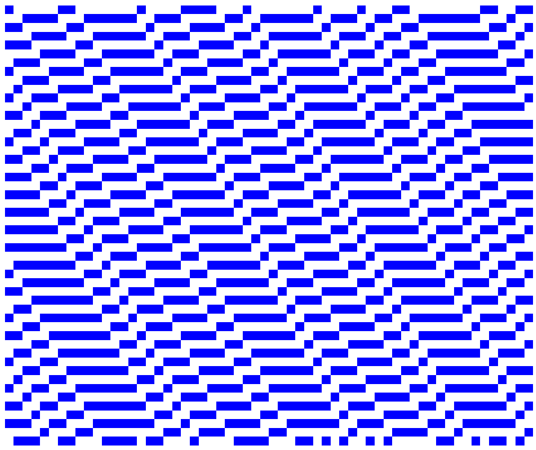
(a) $p = 0.25$



(b) $p = 0.5$



(c) $p = 0.75$



(d) $p = 1$

Diagrammes espace-temps pour différentes valeurs du paramètre p
(simulations effectuées avec le logiciel FiatLux).

Figure 1: Exemple d'ACP sur l'ensemble de cellules $E = \mathbb{Z}$ et l'alphabet $\mathcal{A} = \{0, 1\}$.

de convergence qui apparaissent dans la définition de l'ergodicité, en introduisant les notions de rigidité et de randomisation. Un AC est *rigide* s'il a une unique mesure invariante qui n'est pas dégénérée (dans un sens à préciser), cette mesure étant la mesure produit uniforme. La *randomisation* correspond à la convergence vers la mesure uniforme à partir d'une grande classe de mesures initiales (qui doit aussi être précisée).

Le problème de la classification de la densité consiste à concevoir un ACP ayant un comportement donné. Précisément, les symboles sont binaires, et l'ACP doit converger vers la configuration contenant uniquement l'élément majoritaire, à partir de n'importe quelle mesure de Bernoulli de paramètre différent de 1/2.

Au-delà des ACP, des mesures particulièrement intéressantes sur des espaces symboliques surgissent lorsqu'on étudie le comportement asymptotique d'autres dynamiques à temps discret. Nous introduisons la notion de marches aléatoires sur des produits libres de groupes. La position du marcheur peut être représentée par un mot écrit sous forme normale, et la direction prise par le marcheur dans sa fuite vers l'infini est décrite par une mesure sur les mots infinis. Cette mesure, connue sous le nom de mesure harmonique, a une propriété markovienne particulière : elle est Markov-multiplicative. Les mots écrits sous forme normale constituent un exemple de sous-décalage de type fini (désignés aussi sous le nom de sous-shifts de type fini, ou SFT). Un SFT est l'ensemble des configurations sur un espace symbolique qui ne contiennent pas certains motifs finis. La notion de mesure Markov-multiplicative prend toute son importance lorsqu'on étudie les mesures d'entropie maximale de SFT. Ces mesures, qui sont en un sens les mesures les plus uniformes sur les configurations autorisées, peuvent aussi être vues commes des états d'équilibres particuliers d'ACP.

Nous verrons que l'ACP de la Figure 1, qui a une mesure markovienne invariante, intervient dans l'énumération des animaux dirigés, et est aussi étroitement relié au SFT de Fibonacci, ainsi qu'au modèle de sphères dures utilisé en mécanique statistique. Voilà une illustration des nombreuses connections, parfois inattendues, que les ACP permettent d'établir entre la combinatoire, la mécanique statistique, et la dynamique symbolique. Par ailleurs, tandis que l'ergodicité de cet ACP est facile à prouver pour de petites valeurs du paramètre p, en utilisant un couplage avec un modèle de percolation, la question de l'ergodicité est un problème non résolu pour p proche de 1.

Dans ce contexte, cette thèse commence par une approche générale des ACP, avec un aperçu de différents domaines dans lesquels ils interviennent. Nous présentons la question de l'ergodicité et proposons un algorithme de simulation parfaite pour échantillonner l'unique mesure invariante d'un ACP ergodique.

Nous étudions ensuite des familles spécifiques d'ACP, comme les ACP ayant des mesures de Bernoulli ou des mesures markoviennes invariantes. Nous explorons également le problème de classification de la densité.

Dans la troisième partie, nous nous éloignons un peu des ACP pour nous intéresser à des marches aléatoires sur des produits libres. Mais les ACP jouent à nouveau un rôle fondamental lors de l'analyse des SFT et de leurs mesures d'entropie maximale, qui conclut cette thèse.

Contributions principales

En utilisant la terminologie des chaînes de Markov, un ACP est ergodique s'il a une unique mesure invariante qui est attractive. Dans le cas des AC déterministes, nous prouvons que l'ergodicité est équivalente à la nilpotence (Chap. 3). En corollaire, on obtient que l'ergodicité d'un AC unidimensionnel est indécidable. Cela répond à un problème ouvert proposé par Toom en 2001.

Alors que la mesure invariante d'un AC ergodique est triviale, la mesure invariante d'un ACP ergodique peut être très complexe. Nous proposons un algorithme permettant d'échantillonner parfaitement cette mesure dans certains cas (Chap. 3). Il repose sur l'introduction d'un *ACP enveloppe*, contenant un caractère de remplacement indiquant les états qui ne sont pas encore déterminés. Ce nouvel ACP s'avère être un outil conceptuel et pratique très utile.

Nous présentons une analyse détaillée de l'ACP *majorité-flip*, ainsi que des résultats expérimentaux suggérant une possible transition de phase pour une certaine valeur du paramètre.

Nous présentons une manière alternative de caractériser les ACP élémentaires ayant une mesure de Bernoulli invariante et étudions en détail les propriétés particulières de leurs diagrammes espace-temps (Chap. 4). Nous montrons que les états le long de n'importe quelle ligne droite, à l'exception d'une direction, sont distribués selon la même distribution de Bernoulli, et que l'ACP apparaît dans une seconde direction. À notre connaissance, c'est la première fois que de telles propriétés spatiales sont mises en évidence. La classe d'ACP pour lesquelles elles sont satisfaites apparaît comme l'analogue probabiliste des AC permutatifs (Chap. 4 et 5).

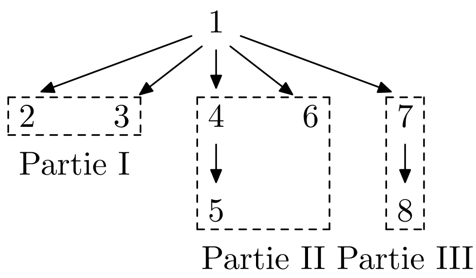
Nous explorons les AC déterministes ayant plusieurs mesures de Bernoulli invariantes, ainsi que les AC *rigides*, pour lesquels la distribution uniforme est essentiellement l'unique mesure invariante (Chap. 5). Nous étendons aux AC qui sont la composition d'une fonction affine et d'une permutation, un théorème de 2003 de Host, Maass, et Martínez portant sur les AC affines.

Nous introduisons le problème de classification de la densité sur des réseaux infinis et des arbres (Chap. 6). En particulier, nous prouvons en utilisant un argument de percolation que l'AC de Toom classifie la densité sur $\mathbb{Z}^2$. Des candidats sont également proposés dans le cas unidimensionnel.

Nous nous intéressons ensuite aux marches aléatoires sur des produits libres de groupes (Chap. 7). Nous présentons un cadre combinatoire permettant de décrire la mesure harmonique, qui fournit la direction prise par le marcheur dans sa fuite vers l'infini. Nous mettons également en évidence le fait que les mesures d'entropie maximale de SFT sur $\mathbb{Z}$ ont cette même propriété d'être Markov-multiplicatives (Chap. 8). Nous étendons cette notion aux SFT sur des arbres infinis et établissons un lien avec la notion de *f-invariant*, introduite par Bowen in 2010. Nous donnons aussi plusieurs manières d'engendrer la mesure d'entropie maximale d'un SFT, et prouvons que c'est la mesure invariante d'un ACP bien choisi.

Structure et contenu de la thèse

Nous présentons ci-dessous plus en détail le contenu des différents chapitres. Les dépendances entre eux sont représentées sur le diagramme suivant.



Comme indiqué sur le diagramme, les différents chapitres sont largement indépendants.

Dans la Section 5.1.3, nous analysons la spécialisation des conditions du Théorème 4.3 aux AC déterministes. Mais le reste du Chapitre 5 est indépendant du Chapitre 4.

Le Chapitre 8 est relié au Chapitre 7 à travers la notion de mesure Markov-multiplicative (Définition 7.1).

Chapitre 1. Cadre mathématique. Nous présentons les principales définitions et notations qui sont utilisées tout au long de cette thèse. Nous introduisons d'abord les espaces symboliques et la dynamique de l'application de décalage (shift) sur les configurations. Nous définissons également les mesures de Bernoulli et les mesures markoviennes sur des espaces symboliques, qui sont des objets centraux de cette thèse. Notre attention se porte ensuite sur les Automates Cellulaires Probabilistes (ACP). Après avoir défini les ACP et leurs mesures invariantes, et introduit la notion d'ergodicité, nous présentons deux spécialisations diamétralement opposées : les ACP à taux strictement positifs (ce sont les ACP n'ayant aucune composante déterministe), et les automates cellulaires déterministes, connus simplement sous le nom d'automates cellulaires. Les diagrammes espace-temps représentent des trajectoires d'ACP. On dit qu'ils sont stationnaires si la trajectoire a pour point de départ une configuration initiale distribuée selon une mesure invariante de l'ACP. Les diagrammes espace-temps d'ACP à taux strictement positifs sont des champs markoviens, tandis que les diagrammes espace-temps d'AC déterministes sont des sous-décalages de type fini. Pour finir, nous présentons des outils issus de la mécanique statistique permettant d'étudier les mesures invariantes d'ACP.

Partie I
Une approche générale des automates cellulaires probabilistes et de leurs mesures invariantes

Cette partie introduit des outils généraux pour étudier les mesures invariantes d'ACP, et explorer leur ergodicité. La présentation est illustrée par différents exemples.

Chapitre 2. Différents points de vue sur les ACP. Nous commençons par commenter notre définition des ACP et nous la comparons avec une définition alternative, pour laquelle l'hypothèse d'indépendance des mises à jour est légèrement assouplie. Cela nous conduit à introduire le modèle du TASEP, qui est étroitement relié à un système de files d'attente. Nous montrons également un autre lien entre les ACP et la combinatoire, qui concerne l'énumération des animaux dirigés. Puis, nous présentons deux spécialisations des ACP qui fournissent des exemples particulièrement intéressants. Toutes deux consistent à considérer un AC déterministe et à "perturber" sa règle locale, soit en effectuant des erreurs aléatoires, soit en introduisant de l'asynchronisme dans l'évolution. En utilisant une approche de mécanique statistique, nous donnons un exemple classique d'ACP de dimension deux à taux strictement positifs qui n'est pas ergodique. La dernière partie illustre avec un modèle de formation d'essaims que les ACP peuvent être utilisés en sciences de la vie comme un outil de modélisation, et que les modèles impliqués soulèvent aussi des questions théoriques passionnantes.

Ce chapitre est essentiellement bibliographique. Le contenu est cependant présenté selon une perspective personnelle, et la dernière partie s'est enrichie de discussions avec Nazim Fatès et Pierre-Yves Louis.

Chapitre 3. Ergodicité et simulation parfaite. Nous revenons à la notion d'ergodicité. Pour les AC déterministes, nous démontrons que l'ergodicité est équivalente à la nilpotence. Cela fournit une preuve de l'indécidabilité de l'ergodicité pour les AC déterministes, ainsi qu'une nouvelle preuve de l'indécidabilité de l'ergodicité pour les ACP. Même dans le cas ergodique, on ne dispose pas d'outil général pour décrire la mesure invariante d'un ACP. Et les simulations doivent être menées avec précaution. En effet, lorsqu'on étudie le comportement à l'équilibre d'un ACP, il y a deux sortes d'infini à prendre en compte : le nombre infini de cellules, et le temps infini, qui correspond au comportement asymptotique de l'ACP. Dans ce contexte, nous avons développé une procédure de simulation parfaite, qui permet, étant donné un ACP ergodique, d'échantilloner son unique mesure invariante (sous certaines conditions). Cette procédure est basée sur une implémentation de l'algorithme de "couplage arrière", en utilisant un processus d'encadrement qui est lui-même un ACP, et que nous appelons l'ACP enveloppe. L'ACP enveloppe est non seulement utile en tant qu'outil pratique pour simuler la mesure invariante d'un ACP ergodique, mais s'avère également être un outil théorique pertinent. Nous illustrons l'intérêt de notre algorithme de simulation parfaite en l'utilisant sur une famille d'ACP à un paramètre, appelé ACP majorité-flip, qui semble présenter une transition de phase à partir d'une certaine valeur critique du paramètre. Nous montrons que cet ACP est relié à la fois à un modèle de percolation et à une marche aléatoire doublement branchante.

Ce chapitre repose sur un travail en collaboration avec Ana Bušić et Jean Mairesse, qui a donné lieu à une publication dans les actes de la conférence *STACS 2011* [BMM11] et à un article plus long accepté pour publication dans le journal *Advances in Applied Probability.*

Partie II
Randomisation, conservation, classification

Cette partie est consacrée à l'étude de différents comportements spécifiques d'ACP. Le point commun de ces trois chapitres est de traiter un problème inverse : nous considérons un certain comportement spécifique, et essayons de trouver un ACP, ou bien l'ensemble des ACP, ayant ce comportement.

Chapitre 4. ACP ayant des mesures de Bernoulli ou des mesures markoviennes invariantes et champs aléatoires avec directions i.i.d. Nous étudions les ACP ayant des mesures produit de Bernoulli invariantes. Lorsque l'alphabet et le voisinage sont tous deux de taille 2, on connaît la condition nécessaire et suffisante sur les valeurs des quatre paramètres définissant l'ACP, sous laquelle l'ACP possède une mesure produit de Bernoulli invariante. Nous présentons une preuve nouvelle et simple de cette caractérisation. Nous explorons ensuite les diagrammes espace-temps stationnaires de tels ACP. Ils peuvent être représentés sur un réseau triangulaire, et ils définissent des champs aléatoires non triviaux ayant de très faibles corrélations. En particulier, des lignes de différentes directions du diagramme espace-temps sont constituées de variables aléatoires i.i.d. Les outils utilisés pour caractériser les ACP ayant des mesures de Bernoulli invariantes permettent également d'étudier les ACP ayant des mesures markoviennes invariantes. Certains de ces ACP interviennent dans l'énumération des animaux dirigés et présentent donc un intérêt particulier. Finalement, nous étendons nos résultats à des alphabets et à des voisinages généraux, et donnons des conditions suffisantes sur les paramètres d'un ACP pour qu'il ait une mesure de Bernoulli invariante.

Ce chapitre repose sur un travail en collaboration avec Jean Mairesse, accepté pour publication aux *Annales de l'Institut Henri Poincaré. Probabilités et statistiques.*

Chapitre 5. Randomisation versus conservation pour les AC unidimensionnels. Nous nous concentrons sur les AC déterministes. Un résultat bien connu est que la mesure uniforme est invariante si et seulement si l'AC est surjectif. Plus généralement, les conditions sous lesquelles un AC déterministe a une mesure de Bernoulli invariante peuvent être écrites sous la forme d'une loi de conservation. En conséquence, les AC pour lesquels toutes les mesures de Bernoulli sont invariantes sont exactement les AC surjectifs et conservatifs, ce qui s'avère très restrictif. À l'opposé, les AC permutatifs apparaissent comme de bons candidats pour la randomisation, c'est-à-dire la convergence (au moins en moyenne de Cesáro) vers la mesure uniforme depuis une grande classe de mesures initiales. Nous introduisons une classe d'AC permutatifs dont la fonction de transition est définie comme la permutation d'une règle affine et prouvons qu'ils sont rigides, au sens où leur seule mesure invariante d'entropie positive est la mesure uniforme.

Ce travail a été initié au cours d'un séjour de recherche avec Alejandro Maass au *Center for Mathematical Modeling* (Universidad de Chile), donnant aussi l'occasion de travailler avec Alexis Ballier. Il a été poursuivi en France avec Benjamin Hellouin de Menibus et Mathieu Sablik.

Chapitre 6. Classification de la densité. Nous explorons le problème de la classification de la densité sur des réseaux infinis et des arbres. Ce problème a d'abord été introduit sur des anneaux finis. Il consiste alors à concevoir un AC (ou un ACP) capable de décider (au moins avec une grande probabilité) si une configuration initiale sur l'alphabet binaire contient plus de 0 ou de 1, en convergeant vers la configuration contenant uniquement l'élément majoritaire. Sur un réseau infini, nous étendons ce problème en demandant à ce que l'AC(P) converge vers la configuration contenant uniquement des 0 à partir d'une mesure produit de Bernoulli de paramètre strictement inférieur à 1/2, et vers la configuration contenant uniquement des 1 à partir d'une mesure de Bernoulli de paramètre strictement supérieur à 1/2. Sur $\mathbb{Z}^2$, nous démontrons que l'AC de Toom classifie la densité. Sur $\mathbb{Z}$, le problème demeure ouvert, et apparaît comme un véritable défi. Nous proposons plusieurs candidats, pour lesquels des résultats expérimentaux suggèrent qu'ils pourraient classifier la densité.

Ce chapitre repose sur un travail en collaboration avec Ana Bušić, Nazim Fatès et Jean Mairesse, qui a donné lieu à une publication dans les actes de la conférence *LATIN 2012* [BFMM12] et à un article plus long publié à l'*Electronic Journal of Probability* [BFMM13].

Partie III
Marches aléatoires et mesures d'entropie maximale

Dans cette partie, nous travaillons sur des mesures spécifiques sur des espaces symboliques, possédant une propriété markovienne. En particulier, les mesures Markov-multiplicatives jouent un rôle fondamental. Une interprétation de ces mesures en termes d'ACP est présentée à la fin du dernier chapitre.

Chapitre 7. Marches aléatoires et mesures Markov-multiplicatives. Nous étudions les marches aléatoires sur les groupes de type produits libres. Ce sont des marches aléatoires sur des graphes réguliers particuliers, à savoir les graphes de Cayley de ces groupes. Elles peuvent également être interprétées comme des empilements aléatoires de pièces. D'un point de vue symbolique, la marche correspond à l'écriture successive de lettres d'un mot sur l'alphabet constitué par les éléments des différents groupes intervenant dans le produit libre. Sous des hypothèses peu restrictives, la marche est transiente, et le mot converge vers un mot infini de forme normale, représentant la direction prise par le marcheur dans sa fuite vers l'infini. Nous étudions la distribution de ce mot infini, appelée la mesure harmonique de la marche aléatoire. Les mesures harmoniques ont la propriété d'être Markov-multiplicative, ce qui en fait en un sens les mesures les plus indépendantes parmi les mesures sur les mots de forme normale. Nous présentons un cadre général permettant d'obtenir une description combinatoire de la mesure harmonique, et illustrons notre méthode sur l'exemple du produit libre $\mathbb{Z}^2 * \mathbb{Z}$, pour lequel nous calculons la valeur de la vitesse de fuite, qui représente la vitesse à laquelle le marcheur s'éloigne vers l'infini.

Ce chapitre repose sur un travail en commun avec Jean Mairesse.

Chapitre 8. Mesures d'entropie maximale de sous-décalages de type fini. Nous considérons d'abord des sous-décalages de type fini (SFT) sur $\mathbb{Z}$. Un résultat bien connu est que la mesure d'entropie maximale d'un SFT est une mesure markovienne, qui peut être décrite via les propriétés de la matrice définissant le SFT (que l'on suppose irréductible). Cette mesure markovienne, désignée sous le nom de mesure de Parry du SFT, a la propriété d'être Markov-multiplicative. Nous présentons des constructions alternatives de cette mesure au moyen de variables aléatoires i.i.d. et d'ACP. Nous considérons ensuite des SFT définis sur des arbres réguliers infinis, et construisons des mesures markoviennes ayant la propriété d'être uniforme sur tous les motifs autorisés, conditionnellement à n'importe quelle valeur du voisinage. Ces mesures, que nous appelons des mesures d-Parry, sont des généralisations naturelles de la mesure de Parry. Nous établissons un lien entre les mesures d-Parry et le f-invariant de Bowen, qui généralise la notion d'entropie aux actions de groupes libres. Précisément, nous prouvons que les mesures maximisant le f-invariant sont les mesures d-Parry. Finalement, nous montrons que les mesures d'entropie maximale sont des mesures réversibles d'ACP.

Le travail sur les mesures de Parry sur $\mathbb{Z}$ est issu de discussions avec Jean Mairesse. L'exploration des SFT définis sur les arbres est un travail en cours avec Vincent Delecroix.

Introduction

This thesis deals with probabilistic cellular automata and specific measures on symbolic spaces.

Symbolic spaces are sets of the form $\mathcal{A}^E$, where $\mathcal{A}$ is a finite set of symbols, and E a countable set, called the set of cells. They appear in various contexts, and in particular when modelling physical and biological phenomena. For example, in the Ising model, which is a mathematical model of ferromagnetism used in statistical mechanics, a material is represented by different spins arranged in a graph (usually, a lattice), each of them being in one of two states, $+1$ or -1. In biology, symbolic spaces can for example be used to model a set of cells that can be in different states (e.g. infected/healthy). Beyond modelling, configurations on symbolic spaces are the way the information is encoded in computing and electronic devices: a digital image consists of a set of pixels arranged in a two-dimensional grid, to which are allocated colors, among a finite set of possible colors.

We are interested in specific probability measures on symbolic spaces. By specific, we mean measures that present original characteristics, as opposed to generic. The specificities of the measures we consider are twofold.

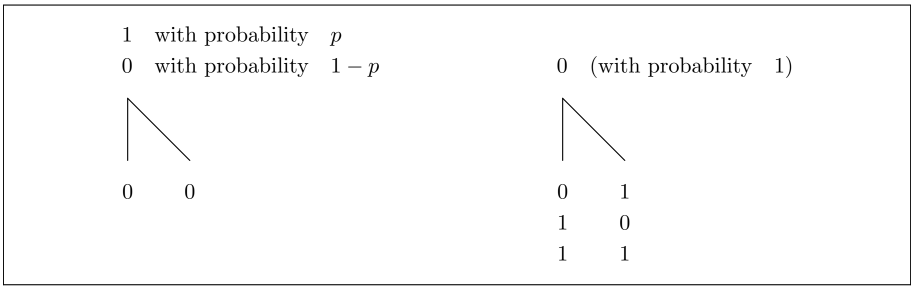
On the one hand, these measures have some intrinsic properties making them special: they have a particular combinatorial structure, involving the topology of the set of cells on which they live. In particular, Markov measures will play a fundamental role.

On the other hand, these measures correspond to some particular equilibrium of stochastic processes, such as random walks or probabilistic cellular automata (PCA).

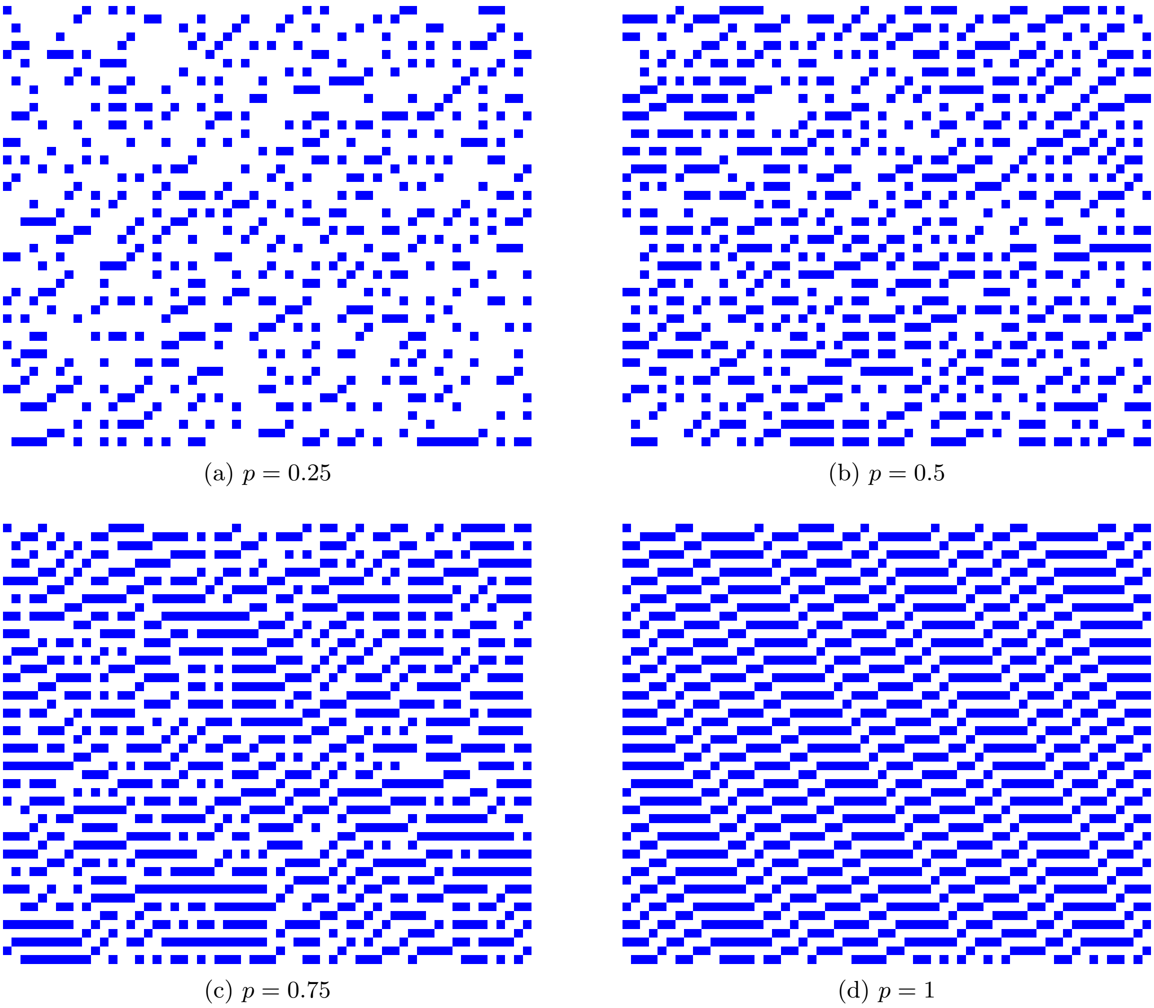
A PCA is a Markov chain on a symbolic space. Time is discrete, and all the cells evolve synchronously: for each cell, the new content is randomly chosen, independently of the others, according to a distribution given by the states in a finite neighbourhood of the cell. Due to the amazing gap between the simplicity of the definition and the intricacy of the generated behaviours, PCA are good candidates for modelling complex systems appearing in physical and biological processes. They are also used to investigate fault-tolerant computational models. Finally, they appear in different contexts in probability theory and in combinatorics.

Consider the specific case of the set of cells $E = \mathbb{Z}$, the alphabet $\mathcal{A} = \{0, 1\}$, and the neighbourhood consisting of the cell itself and its right neighbour (or, the left neighbour and the cell itself). Then, a PCA is entirely determined by the four parameters $(\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11})$, where $\theta_{ij} \in [0, 1]$ is the probability that a cell is updated to 1 if its neighbourhood is ij. Consider for example the PCA defined by the parameters $(p, 0, 0, 0)$ for some $p \in [0, 1]$, see Fig. 2. This PCA can be described as follows: if the neighbourhood of a cell is in state 00, then, with probability p the cell is updated to 1, and with probability $1-p$ the cell is updated to 0. Otherwise, the cell is updated to 0 (with probability 1). For $p = 0$ and $p = 1$, we obtain a deterministic cellular automaton.

The trajectories of a PCA are represented by space-time diagrams, living themselves on a symbolic space, with an additional dimension corresponding to time. In Fig. 2, the cells



Local rule of the PCA.



(a) $p = 0.25$ (b) $p = 0.5$

(c) $p = 0.75$ (d) $p = 1$

Space-time diagrams for different values of the parameter p
(simulations made with the software FiatLux).

Figure 2: Example of PCA on the set of cells $E = \mathbb{Z}$ and the alphabet $\mathcal{A} = \{0, 1\}$.

containing a 0, resp. a 1, are represented in white, resp. blue. The bottom line is the initial condition, here choosen at random, and the next lines, from bottom to top, are the successive updates of the cells.

The equilibrium behaviour of a PCA is studied via the invariant measures of the Markov chain on the symbolic space on which it is defined. Several questions arise. A PCA is ergodic if it has a unique and attractive invariant measure. The problem of the ergodicity of PCA is known to be undecidable: there exists no algorithm able to say, when taking in input the parameters of a PCA, if it is ergodic or not. And no general tool is known to describe the invariant measures of PCA.

A challenging problem in this area is the positive rates problem. A PCA is said to have positive rates if for any neighbourhood, the updated content of a cell can be any symbol with a strictly positive probability. There are simple examples of two-dimensional positive-rate PCA that are non-ergodic, but for one-dimensional PCA, the only known example was exhibited in 2001 by Gács (after a first publication in 1986), and it is very complex. If we restrict ourselves to one-dimensional PCA having a neighbourhood of size 2, and defined on a set of symbols of size 2, it is unknown if any positive-rate PCA (that is, PCA such that $\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11} \in (0,1)$) is ergodic.

Tools coming from statistical mechanics have been developed to study positive-rate PCA. In this case, the research of invariant measures is shown to be equivalent to an equilibrium statistical mechanics problem. In our work, we are also interested in PCA having deterministic components. Some tools of statistical mechanics can still be adapted, but the analysis of such PCA needs to be done individually and very carefully.

When explicit computation is not possible, simulation becomes the alternative. Simulating PCA is known to be a challenging task, costly both in time and space. Also, configurations cannot be tracked down one by one when the number of cells is infinite, and may only be observed through some measured parameters. So the crucial point is whether some guarantees can be given upon the results obtained from simulations. By adapting the coupling from the past method of Propp and Wilson, we propose a perfect sampling procedure to estimate the invariant measure of an ergodic PCA, under some suitable conditions. The algorithm is random and returns a configuration (or a portion of configuration) distributed according to the invariant measure, so that by applying the procedure repeatedly, we can estimate the invariant measure with arbitrary precision.

In some very particular cases, it is however possible to foresee theoretically the asymptotic behaviour of a PCA. For example, there is a known characterisation of PCA having a Bernoulli product invariant measure. We show that the stationary space-time diagrams of such PCA define measures with very weak dependence, sharing some special properties.

The case of deterministic CA having Bernoulli invariant measure turns out to be also interesting. Since for deterministic CA, ergodicity is equivalent to nilpotency, it is relevant to relax both the uniqueness of the invariant measure and the convergence property of the definition of the ergodicity, by introducing the notions of rigidity and randomisation. A CA is rigid if its only invariant measure that is non-degenerated (in some sense that has to be specified) is the uniform product measure. The randomisation corresponds to the convergence to the uniform measure from a large class of initial measures (which also needs to be specified).

The density classification problem consists in designing a PCA having a certain behaviour. Precisely, the symbols are binary, and the PCA should converge to the configuration containing only the element in the majority from any Bernoulli product measure of parameter different from 1/2.

Beyond PCA, measures on symbolic spaces of particular interest arise when studying the asymptotic behaviour of other discrete time dynamics. We introduce random walks on free products of groups. The position of the walker can be represented by a normal form word, and the direction taken by the walker in its escape to infinity is described by a measure on infinite words. This measure, known as the harmonic measure, has a particular Markov property: it is Markov-multiplicative. Normal form words are an example of subshift of finite type (SFT). An SFT is the set of configurations on a symbolic space that do not contain some given finite patterns. The notion of Markov-multiplicative measures takes also great importance when studying measures of maximal entropy of SFT. These measures, that are in some sense the most uniform measures on admissible configurations, can also be seen as special equilibrium measures of PCA.

We will see that the PCA of Fig. 2, which has a Markov invariant measure, is involved in the counting of directed animals, and is also tightly related to the Fibonacci SFT and the hardcore lattice gas model. This is one illustration of the many unexpected connections offered by PCA between combinatorics, statistical mechanics, and symbolic dynamics. Furthermore, whereas ergodicity is easy to prove for small values of the parameter p, by a coupling with a percolation model, the question of ergodicity appears to be a difficult problem when p is close to 1.

In that context, this thesis presents first a general approach to PCA, with an insight into different domains in which they are involved. We address the question of ergodicity and propose a perfect sampling algorithm to sample the unique invariant measure of an ergodic PCA.

Second, we study specific families of PCA, such as PCA having Bernoulli and Markov invariant measures. We also explore the density classification problem.

In the third part of the thesis, we leave PCA for a while to consider random walks on free product of groups. But PCA will play again a fundamental role when analysing subshifts of finite type and their measures of maximal entropy.

Main contributions

Using the terminology of Markov chains, a PCA is called ergodic if it has a unique and attractive invariant measure. In the case of deterministic CA, we prove that ergodicity is equivalent to nilpotency (Chap. 3). As a corollary, one obtains that it is undecidable if a given one-dimensional CA is ergodic. This answers an open problem asked by Toom in 2001.

While the invariant measure of an ergodic CA is trivial, the invariant measure of an ergodic PCA can be very complex. We describe an algorithm to perfectly sample this measure in certain cases (Chap. 3). It is based on the introduction of an *envelope PCA*, containing a wildcard state indicating states that are not yet determined. This new PCA turns out to be a powerful conceptual and practical tool.

We present an in-depth analysis of the *majority-flip* PCA as well as experimental results, suggesting a possible phase transition for some value of the parameter.

We present an alternative way to characterise elementary PCA having Bernoulli invariant measure and study in detail the peculiar properties of their space-time diagrams (Chap. 4). The states along any line, with the exception of one direction, are proved to be distributed according to the same Bernoulli distribution, and the original PCA appears in a second direction. To our knowledge, it is the first time that such spatial properties are highlighted. The class of PCA for which they hold appear as the probabilistic counterpart of deterministic permutative CA (Chap. 4 and 5).

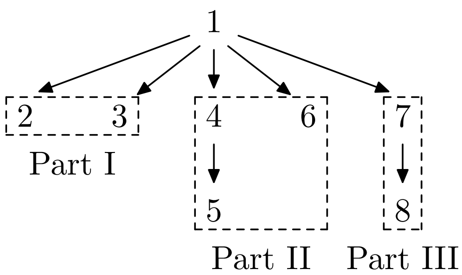
We explore deterministic CA having several Bernoulli measures as well as *rigid* CA, for which the uniform distribution is practically the only invariant measure (Chap. 5). We extend to CA that are compositions of an affine function and a state permutation a theorem on affine CA published in 2003 by Host, Maass, and Martínez.

We explore the density classification problem on infinite lattices and trees (Chap. 6). In particular, it is proved using a percolation argument that the well-known Toom's CA classifies the density on $\mathbb{Z}^2$. Candidates are also proposed for the one-dimensional case.

We then focus on random walks on free products of groups (Chap. 7). We provide a combinatorial description of the harmonic measure, giving the direction taken by the walker in its escape to infinity. We also highlight the fact that measures of maximal entropy of SFT on $\mathbb{Z}$ have this same property of being Markov-multiplicative measures (Chap. 8). We extend this notion to SFT on infinite trees and exhibit a connection with the notion of *f-invariant*, introduced by Bowen in 2010. We also provide different ways to generate the measure of maximal entropy of a SFT, which is shown to be the invariant measure of a well-suited PCA.

Thesis structure and content

We present below more in detail the content of the different chapters. The dependence between them is represented by the following diagram.



As illustrated, the different chapters are largely independent.

In Sec. 5.1.3, we analyse the specialisation of the conditions of Th. 4.3 to deterministic CA. But the rest of Chap. 5 is independent from Chap. 4.

Chap. 8 is related to Chap. 7 through the notion of Markov-multiplicative measure (Def. 7.1).

Chapter 1. Mathematical framework. We present the main definitions and notations that will be used all along the thesis. We first introduce symbolic spaces and the dynamics of the shift on configurations. We also define Bernoulli and Markov measures on symbolic spaces, that are central objects of the present thesis. The focus is then on probabilistic cellular automata (PCA). After having defined PCA and their invariant measures, and introduced the notion of ergodicity, we present two diametrically opposed specialisations: positive-rate PCA (which are PCA having no deterministic components), and deterministic cellular automata, known also simply as cellular automata. Space-time diagrams represent trajectories of PCA. They are called stationary if the trajectory is initiated from an initial configuration distributed according to an invariant measure of the PCA. Space-time diagrams of positive-rate PCA are Markov random fields, whereas space-time diagrams of deterministic CA are subshifts of finite type. We also introduce interacting particle systems, that are the continuous-time analogues of PCA. Finally, we present classical tools coming from statistical mechanics to study the invariant measures of PCA.

Part I
Probabilistic cellular automata and their invariant measures: a general approach

This part introduces general tools to study the invariant measures of PCA, and to explore their ergodicity. The presentation is illustrated by different examples.

Chapter 2. Different viewpoints on PCA. We first discuss our definition of PCA, and compare it with an alternative one, where the assumption of independence of the updates of different cells is slightly relaxed. This leads us to introduce the TASEP model, which is tightly related to a queueing system. We also show another link between PCA and combinatorics, concerning the counting of directed animals. Then, we present two specifications of PCA that provide examples of particular interest. They both consist in considering a deterministic CA and "perturbating" the rule, either by doing random errors or by introducing asynchronism in the evolution. Using the statistical mechanics approach, we also exhibit a two-dimensional positive-rate PCA that is not ergodic. The last section illustrates with a model of swarming, that PCA can also be used in life sciences as a modelling tool, and that the models involved also give rise to exciting theoretical questions.

This chapter is mostly bibliographical. The content is however presented from a personal perspective, and the last section has benefited from discussions with Nazim Fatès and Pierre-Yves Louis.

Chapter 3. Ergodicity and perfect sampling. We come back to the notion of ergodicity. For deterministic CA, we prove that ergodicity is equivalent to nilpotency. This provides a proof of the undecidability of the ergodicity for deterministic CA, as well as a new proof of the undecidability of the ergodicity for PCA. Even in the ergodic case, there are no general tools to describe the invariant measure of a PCA. And simulations have to be taken with care, since when studying the equilibrium behaviour of a PCA, there are two kinds of infinity one has to take into account: the infinite number of cells, and the infinite time, corresponding to the asymptotic behaviour of the PCA. In that context, we have developed a perfect sampling procedure that allows, given an ergodic PCA, to sample its unique invariant measure (under some conditions). This procedure is based on an implementation of the "coupling from the past" algorithm, using a bounding process which is itself a PCA, that we call the envelope PCA. The envelope PCA turns out to be useful not only as a practical tool for sampling the invariant measure of an ergodic PCA, but also as a theoretical tool. We illustrate the use of our perfect sampling algorithm with a one-parameter family of PCA called the majority-flip PCA, that is suspected to present a phase transition from some threshold value of the parameter. We also show that this PCA is related both with a percolation model and with a double branching annihilating random walk.

This chapter is based on a joint work with Ana Bušić and Jean Mairesse, that has given rise to a publication in the proceedings of the conference *STACS 2011* [BMM11] and to a longer article that will be published in the journal of *Advances in Applied Probability.*

Part II
Randomisation, conservation, classification

This part is devoted to the study of different specific behaviours of PCA. The common point of the three chapters is that the approach consists in studying an inverse problem: we consider some specific asymptotic behaviour, and try to find one or all PCA presenting this behaviour.

Chapter 4. PCA having Bernoulli or Markov invariant measures and random fields with i.i.d. directions. We study PCA having Bernoulli product invariant measures. When both alphabet and neighbourhood have size 2, there is a known necessary and sufficient condition on the values of the four parameters defining the PCA under which it has a Bernoulli product invariant measure. We give a new and simple proof of this characterisation. We then explore the stationary space-time diagrams of such PCA. They can be represented on a triangular lattice, and they define non-trivial random fields having very weak correlations. In particuliar, lines in different directions of the space-time diagram are constituted of i.i.d. random variables. The tools used to characterise PCA having Bernoulli invariant measures can also be used to study PCA having Markov invariant measures. Some of these PCA are related to the counting of directed animals and thus present a particular interest. Finally, we extend our results to general alphabet and neighbourhood, and give sufficient conditions on the parameters of a PCA for having a Bernoulli invariant measure.

This chapter is based on a joint work with Jean Mairesse, accepted for publication at the *Annales de l'Institut Henri Poincaré. Probabilités et statistiques.*

Chapter 5. Randomisation versus conservation in one-dimensional CA. We focus on deterministic CA. It is well known that the uniform Bernoulli product measure is invariant if and only if the CA is surjective. More generally, the conditions for a deterministic CA to have a Bernoulli product measure can be written in terms of a conservation law. Consequently, CA for which all the Bernoulli product measures are invariant are exactly surjective and state-conserving CA, and it corresponds to very constraint rules. At the opposite, permutative CA appear to be good candidate for randomising, that is, converging (or at least, converging in Cesáro mean) to the uniform product measure from a large range of initial measures. We introduce a class of permutative CA whose transition function is defined by a permutation of an affine rule, and prove that they are rigid in the sense that their unique invariant measure of positive entropy is the uniform measure.

This work was initiated during a research period with Alejandro Maass at the *Center for Mathematical Modeling* (Universidad de Chile), giving also the opportunity to discuss with Alexis Ballier. It has been carried on in France with Benjamin Hellouin de Menibus and Mathieu Sablik.

Chapter 6. Density classification on infinite lattices and trees. We explore the density classification problem on infinite lattices and trees. This problem was initially considered on finite rings. It then consists in designing a CA (or a PCA) able to decide (at least with a high probability) if an initial configuration on the binary alphabet contains more 0's or 1's, by converging to the configuration containing only the element in the majority. On an infinite lattice, we extend this problem by asking the PCA to converge to the configuration with only 0's from any Bernoulli product measure of parameter smaller than 1/2, and to the configuration with only 1's from a Bernoulli product measure of parameter larger than 1/2. On $\mathbb{Z}^2$, we prove that Toom's CA classifies the density. On infinite trees, we are also able to provide examples of CA that classify the density. The problem is open on $\mathbb{Z}$ and appears as a difficult challenge. We propose some candidates, for which numerical results suggest that they could classify the density.

This chapter is based on a joint work with Ana Bušić, Nazim Fatès and Jean Mairesse, that has lead to a publication in the proceedings of the conference *LATIN 2012* [BFMM12] and to a longer article published at the *Electronic Journal of Probability* [BFMM13].

Part III
Random walks and measures of maximal entropy

In this part, we work on specific measures on symbolic spaces, having a Markov property. In particular, Markov-multiplicative measures play a fundamental role. An interpretation in terms of PCA is presented at the end of the last chapter.

Chapter 7. Random walks and Markov-multiplicative measures. We study random walks on groups of free-product type. They can be thought of as random walks on particular regular graphs, that are the Cayley graphs of these groups, but also as random heaps of pieces. From the symbolic viewpoint, the walk can be seen as the successive writings of letters of a word on the alphabet constituted by the elements of the different groups involved in the free product. Under very weak conditions, the walk is transient, and the word converges to an infinite normal form word, representing the direction taken by the walk in its escape to infinity. We study the distribution of that infinite word, which is the so-called harmonic measure. Harmonic measures have the property to be Markov-multiplicative. This makes them in some sense the most independent measures among the measures on normal form words. We present a general framework allowing to obtain a combinatorial description of the harmonic measure, and illustrate it in the case of the free product $\mathbb{Z}^2 * \mathbb{Z}$, for which we also compute the value of the drift, which represents the speed of escape to infinity of the walk.

This chapter is based on a joint work with Jean Mairesse.

Chapter 8. Measures of maximal entropy of subshifts of finite type. We first consider subshifts of finite type on $\mathbb{Z}$. It is well known that the measure of maximal entropy of a SFT is a Markov measure, which can be described through the properties of the matrix defining the SFT (which we assume to be irreducible). This Markov measure, which we refer to as the Parry measure of the SFT, happens to be Markov-multiplicative. We present alternative constructions of that measure with the mean of i.i.d. random variables and PCA. We then consider subshifts of finite-type defined on infinite regular trees, and design Markov measures having the property to be uniform on all allowed patterns conditionally to any fixed value of the neighbourhood. These measures, that we call d-Parry measures, are natural generalisations of the Parry measure. We relate d-Parry measures with the f-invariant of Bowen, generalising the notion of entropy to free group actions. Precisely, we prove that the measures maximising the f-invariant are d-Parry measures. Finally, we present an interpretation of measures of maximal entropy as reversible measures of PCA.

The work on Parry measures on $\mathbb{Z}$ stemmed from discussions with Jean Mairesse. The exploration of SFT defined on trees is a work in progress with Vincent Delecroix.

Chapter 1

Mathematical background

Cómo se llama una flor que vuela de pájaro en pájaro?

– Pablo Neruda, *El Libro De Las Preguntas*

Contents

This chapter presents the main concepts and notations that will be used through this thesis.

1.1 Shift spaces

Let $\mathcal{A}$ be a finite set called the *alphabet*, whose elements are referred to as *letters* or *symbols*, and let E be a countable set of *cells*. We consider the *symbolic space* $\mathcal{X} = \mathcal{A}^E$. An element $(x_k)_{k\in E}$ of $\mathcal{X}$ is a *configuration*.

To go forward, we need some additional structure on the set E. For simplicity, we assume in this section that E is equal to $\mathbb{Z}^d$, for some $d \geq 1$, but most of the notions that follow can be extended to general discrete groups.

For a finite set $K \subset E$, a *cylinder* of *base* K is a subset of $\mathcal{X}$ having the form

$$[y_K] = \{x \in \mathcal{X}\,;\ \forall k \in K,\, x_k = y_k\}$$

for some element $y = (y_k)_{k\in K} \in \mathcal{A}^K$. We denote by $\mathcal{C}(K)$ the set of cylinders of base K.

For $x \in \mathcal{A}^K$ and $\alpha \in \mathcal{A}$, we denote by $|x|_\alpha$ the number of occurrences of the letter α in x, that is,

$$|x|_\alpha = \text{Card}\ \{k \in K; x_k = \alpha\}.$$

We equip $\mathcal{A}$ with the discrete topology. The product topology on $\mathcal{X}$ can be described as the topology generated by cylinders. With this topology, $\mathcal{X}$ is a compact metric space. A distance on $\mathcal{X}$ can be given by:

$$d(x,y) = 2^{-r}, \text{ where: } r = \max\{r \in \mathbb{N}; \forall \|k\| \leq r, x_k = y_k\},$$

where $\|\cdot\|$ can for example denote the 1-norm $\|k\|_1 = \sum_{i=1}^{d} |k_i|$. The distance d on $\mathcal{X}$ reflects the combinatorics of configurations: two configurations are close from each other if they coincide on a large pattern around the origin.

For $n \in \mathbb{Z}^d$, the *shift* σ^n is the homeomorphism defined by:

$$\begin{array}{rrcl} \sigma^n : & \mathcal{X} & \to & \mathcal{X} \\ & x = (x_k)_{k\in\mathbb{Z}^d} & \mapsto & \sigma^n(x) = (x_{n+k})_{k\in\mathbb{Z}^d}. \end{array} \tag{1.1}$$

A set $X \subset \mathcal{X}$ is said *shift-invariant* if $\sigma^n(X) = X$ for any $n \in \mathbb{Z}^d$. A *subshift* is a closed shift-invariant subset X of $\mathcal{X}$. The set $\mathcal{X}$ is referred as the *full shift*.

For any non-empty set $F \subset \mathbb{Z}^d$, we define the map π_F as the projection restricting each element $x \in \mathcal{X}$ to the window F, that is:

$$\begin{array}{rrcl} \pi_F : & \mathcal{X} & \to & \mathcal{A}^F \\ & (x_k)_{k\in\mathbb{Z}^d} & \mapsto & (x_k)_{k\in F}. \end{array}$$

Definition 1.1 (Subshifts of finite type)**.** A subshift $X \subset \mathcal{X}$ is a *subshift of finite type (SFT)* if there exists a finite set $F \subset \mathbb{Z}^d$ and a set of patterns $P \subset \mathcal{A}^F$ such that:

$$X = \{x \in \mathcal{X}; \forall n \in \mathbb{Z}^d, \pi_F \circ \sigma^n(x) \in P\}.$$

The set P is then known as the set of *allowed patterns*.

1.2 Bernoulli and Markov measures

We still consider a space $\mathcal{X} = \mathcal{A}^E$, with $E = \mathbb{Z}^d$ for some $d \geq 1$. Let us denote by $\mathcal{M}(\mathcal{A})$ the set of probability measures on the alphabet $\mathcal{A}$, and by $\mathcal{M}(\mathcal{X})$ the set of probability measures on $\mathcal{X}$ for the Borel σ-algebra.

Let $p = (p_i)_{i\in\mathcal{A}} \in [0,1]^{\mathcal{A}}$ be a vector satisfying $\sum_{i\in\mathcal{A}} p_i = 1$. We denote by $\mathcal{B}(p)$ the corresponding probability measure on $\mathcal{A}$, called the *Bernoulli measure* of parameter p.

The *Bernoulli product measure* induced by p on $\mathcal{X}$ is the measure $\mu_p = \mathcal{B}_p^{\otimes\mathbb{Z}^d}$. Thus, for any cylinder set $[y]_K$, we have

$$\mu_p([y_K]) = \prod_{k\in K} p_{y_k} = \prod_{i\in\mathcal{A}} p_i^{|y|_i}.$$

The *uniform measure* on $\mathcal{X}$ is the Bernoulli product measure induced by the uniform Bernoulli measure on $\mathcal{A}$. We will denote it by λ.

Definition 1.2 (Markov random fields)**.** A measure $\mu \in \mathcal{M}(\mathcal{X})$ is a *Markov random field* if it satisfies:

$$\mu\Big([a_F] \;\Big|\; [b_{\partial F}] \cap [c_G]\Big) = \mu\Big([a_F] \;\Big|\; [b_{\partial F}]\Big)$$

whenever F and G are finite subsets of $\mathbb{Z}^d$, $F \cap G = \emptyset$, and $\mu([b_{\partial F}] \cap [c_G]) > 0$, where ∂F stands for the *boundary* of F:

$$\partial F = \{k \in \mathbb{Z}^d \setminus F \; ; \; \exists n \in F, \|n-k\|_1 = 1\}.$$

The definition can naturally be extended to other graphs than $\mathbb{Z}^d$, as well as to other definitions of the boundary.

Bernoulli product measures are examples of Markov random fields.

For $d = 1$, Markov random fields coincide with finite state stationary *Markov measures*. A *Markov chain* is described by a *transition matrix* $Q \in [0,1]^{\mathcal{A}^2}$, such that

$$\forall i \in \mathcal{A}, \sum_{j \in \mathcal{A}} Q_{i,j} = 1.$$

An *invariant measure* of Q is a probability $(\pi_i)_{i \in \mathcal{A}} \in [0,1]^{\mathcal{A}}$ satisfying:

$$\forall j \in \mathcal{A}, \ \pi_j = \sum_{i \in \mathcal{A}} \pi_i Q_{i,j}.$$

The Markov measure μ induced on $\mathcal{A}^{\mathbb{Z}}$ is defined on cylinders by:

$$\mu([x_m, x_{m+1}, \ldots, x_n]) = \pi_{x_m} \prod_{i=m}^{n-1} Q_{x_i, x_{i+1}}.$$

On $\mathbb{Z}^d (d \geq 1)$, Markov fields are equivalent to Gibbs measures with nearest neighbour potentials [Spi71].

The terminology used in the next definition is proper to this thesis.

Definition 1.3 (Markov-uniform measure)**.** We say that a Markov random field $\mu \in \mathcal{M}(\mathcal{X})$ is *Markov-uniform* if the quantity $\mu([a_F] \mid [b_{\partial F}])$ does not depend on the cylinder $[a_F]$ of base F such that $\mu([a_F] \cap [b_{\partial F}]) > 0$. That is, conditionally to any fixed value of the neighbourhood, the measure μ is uniform on all patterns of positive probability.

1.3 Probabilistic cellular automata (PCA)

In this section, we define and consider probabilistic cellular automata on the set of cells $E = \mathbb{Z}^d$. All the definitions still make sense if one replaces $\mathbb{Z}^d$ by $\mathbb{Z}_{m_1} \times \ldots \mathbb{Z}_{m_d}$, where $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$. This corresponds to the restriction of a PCA defined on $\mathbb{Z}$ to a finite window of size $m_1 \times \ldots \times m_d$, with periodic boundary conditions. The definitions can also be adapted to more general discrete groups with very slight changes. This extension will be used in particular in Chap. 6.

1.3.1 Definition

Definition 1.4 (Probabilistic cellular automata)**.** Let $\mathcal{N} \subset \mathbb{Z}^d$ be a finite set, called the *neighbourhood*. A *(local) transition function* of neighbourhood $\mathcal{N}$ is a function

$$f : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A}).$$

The *probabilistic cellular automaton (PCA)* of transition function f is the map

$$\begin{array}{rccc} F: & \mathcal{M}(\mathcal{X}) & \to & \mathcal{M}(\mathcal{X}) \\ & \mu & \mapsto & \mu F \end{array}$$

defined on cylinders by:

$$\mu F([y_K]) = \sum_{[x_{K+\mathcal{N}}] \in \mathcal{C}(K+\mathcal{N})} \mu([x_{K+\mathcal{N}}]) \prod_{k \in K} f((x_{k+v})_{v \in \mathcal{N}})(y_k).$$

A PCA is a Markov chain on the state space $\mathcal{X}$. Consider a realisation $(X^n)_{n\in\mathbb{N}}$ of that Markov chain. If X^0 is distributed according to $\mu \in \mathcal{M}(\mathcal{X})$, then X^n is distributed according to μF^n.

Let us assume that the initial measure is concentrated on some configuration $x \in \mathcal{X}$. Then by application of F, the content of the cell $k \in \mathbb{Z}^d$ is updated to the letter $\alpha \in \mathcal{A}$ with probability $f((x_{k+v})_{v\in\mathcal{N}})(\alpha)$, choices being independent for different cells. The real number $f((x_{k+v})_{v\in\mathcal{N}})(\alpha) \in [0,1]$ is thus to be thought as the conditional probability, that, after application of F, the k-th cell will be in the state α if, before its application, the neighbourhood of k was in the state $(x_{k+v})_{v\in\mathcal{N}}$.

In other words, if δ_x denotes the Dirac measure concentrated on the configuration x, its image $\delta_x F$ by the PCA F is a product measure. In particular, if x is the *monochromatic configuration* $\alpha^{\mathbb{Z}^d}$, which means that $x_k = \alpha$ for all $k \in \mathbb{Z}^d$, then the measure $\delta_x F$ is the Bernoulli measure μ_p induced by the probability $p = f(\alpha^{\mathcal{N}})$.

Definition 1.5. A measure $\mu \in \mathcal{M}(\mathcal{X})$ is an *invariant measure* of the PCA F if $\mu F = F$. A PCA F is *ergodic* if it has exactly one invariant measure $\pi \in \mathcal{M}(\mathcal{X})$ which is *attractive*, that is, for any measure $\mu \in \mathcal{M}(\mathcal{X})$, the sequence μF^n converges weakly to π, *i.e.* for any cylinder $[x_K]$,

$$\lim_{n\to+\infty} \mu F^n([x_K]) = \pi([x_K]).$$

An important point is that any probabilistic cellular automaton has at least one invariant measure. The proof of that proposition is based on the observation that the set $\mathcal{M}(\mathcal{X})$ of measures on $\mathcal{X}$ is compact for the weak topology [DKT90]. Since the application $\mu \mapsto \mu F$ is continuous for this topology, Schauder-Tychonoff fixed point theorem gives the result. An other way to conclude the proof is to observe that for every measure $\mu \in \mathcal{M}(\mathcal{X})$, the sequence of Cesàro sums

$$\frac{\mu + \mu F + \ldots + \mu F^{n-1}}{n}$$

has some accumulation point, which is an invariant measure of the PCA.

One can even prove a stronger statement: any PCA has at least one invariant measure which is shift-invariant.

The existence of several invariant measures obviously implies the non-ergodicity of the system, but the reverse is not true, as it will be evocated in Chap. 3.

1.3.2 Positive-rate PCA and deterministic cellular automata

PCA having no deterministic components are said to be *positive-rate PCA*.

Definition 1.6 (Positive-rate PCA)**.** A PCA has *positive rates*, if

$$\forall (x_v)_{v\in\mathcal{N}} \in \mathcal{A}^{\mathcal{N}}, \forall \alpha \in \mathcal{A},\ f((x_v)_{v\in\mathcal{N}})(\alpha) > 0.$$

At the opposite, classical cellular automata are another specialisation of PCA, for which the transition function is deterministic.

Definition 1.7 (Deterministic cellular automata)**.** A PCA is a *deterministic cellular automaton (CA)* if for each sequence $(x_v)_{v\in\mathcal{N}} \in \mathcal{A}^{\mathcal{N}}$, the measure $f((x_v)_{v\in\mathcal{N}})$ is concentrated on a single letter of the alphabet.

The transition function can thus be seen as a function $f : \mathcal{A}^{\mathcal{N}} \mapsto \mathcal{A}$, and the CA as a deterministic function $F : \mathcal{X} \to \mathcal{X}$.

Deterministic cellular automata have been widely studied, in particular on the set of cells $E = \mathbb{Z}$. They are classical and relevant mathematical objects: by Curtis-Hedlund-Lyndon theorem [Hed69], deterministic cellular automata are precisely the mappings from $\mathcal{A}^{\mathbb{Z}^d}$ to $\mathcal{A}^{\mathbb{Z}^d}$ which are continuous (with respect to the product topology) and commute with the shift.

1.3.3 Space-time diagrams and update functions

As an extension of the usual notion of space-time diagrams in the deterministic context, we introduce the following definition.

Definition 1.8 (Space-time diagram)**.** A *space-time diagram* is a trajectory $(X^n)_{n\in\mathbb{N}} = (X^n_k)_{k\in\mathbb{Z}^d,n\in\mathbb{N}}$ of a PCA, from the (random) initial configuration X^0. The variable X^n_k of the random field $(X^n)_{n\in\mathbb{N}} = (X^n_k)_{k\in\mathbb{Z},n\in\mathbb{N}}$ is indexed by its *space-coordinate* k, and its *time-coordinate* n.

Consider a PCA F on the set of cells $\mathbb{Z}^d$ and let μ be an invariant measure of F. It is possible to start the evolution of the PCA F from an initial configuration distributed according to μ at instant $-N$, instead of 0. By invariance of μ, the laws of the space-time diagrams obtained are consistent. It follows from Kolmogorov extension theorem that there exists a uniquely defined law in $\mathcal{M}(\mathcal{A}^{\mathbb{Z}^d\times\mathbb{Z}})$ whose restrictions are the laws of these space-time diagrams. Roughly, it is the law of a space-time diagram starting at time $-\infty$.

We will see in Sec. 1.4.2 that if a PCA has positive rates, then for any of its invariant measures, the stationary space-time diagram obtained defines a Markov random field on $\mathcal{A}^{\mathbb{Z}^{d+1}}$.

For a deterministic CA, any initial configuration defines a unique space-time diagram. The set of bi-infinite space-time diagrams of a given deterministic CA defines a subshift of finite type on $\mathcal{A}^{\mathbb{Z}^{d+1}}$. Let $F = \{(v,0)\,;v\in\mathcal{N}\}\cup\{(0,1)\}$. The set of allowed patterns is: $P = \{x_F\in\mathcal{A}^F;\, x^1_0 = f((x^0_v)_{v\in\mathcal{N}})\}$.

Let τ be the uniform measure on $[0,1]$. We define the product measure $\mathcal{U} = \tau^{\otimes\mathbb{Z}^d}$ on $[0,1]^{\mathbb{Z}^d}$. Space-time diagrams of PCA can be generated using an *update function* that takes in input a configuration and a sample in $[0,1]^{\mathbb{Z}^d}$, and returns a new configuration according to the right probability.

Definition 1.9 (Update function)**.** An *update function* of the PCA F is a deterministic function $\phi:\mathcal{A}^{\mathbb{Z}^d}\times[0,1]^{\mathbb{Z}^d}\to\mathcal{A}^{\mathbb{Z}^d}$, satisfying for each $x\in\mathcal{A}^{\mathbb{Z}^d}$, and each cylinder y_K,

$$\mathcal{U}(\{r\in[0,1]^{\mathbb{Z}^d};\phi(x,r)\in[y_K]\}) = \prod_{k\in K} f((x_{k+v})_{v\in\mathcal{N}})(y_k).$$

In practice, it is always possible to define an update function ϕ for which the value of $\phi(x,r)_k$ only depends on $(x_{k+v})_{v\in\mathcal{N}}$ and on r_k. For example, if the alphabet is $\mathcal{A} = \{\alpha_1,\ldots,\alpha_n\}$, one can set

$$\phi(x,r)_k = \begin{cases} \alpha_1 & \text{if } \ 0\le r_k < f((x_{k+v})_{v\in\mathcal{N}})(\alpha_1) \\ \alpha_2 & \text{if } \ f((x_{k+v})_{v\in\mathcal{N}})(\alpha_1)\le r_k < f((x_{k+v})_{v\in\mathcal{N}})(\{\alpha_1,\alpha_2\}) \\ \vdots & \\ \alpha_n & \text{if } \ f((x_{k+v})_{v\in\mathcal{N}})(\{\alpha_1,\alpha_2,\ldots,\alpha_{n-1}\})\le r_k\le 1. \end{cases} \tag{1.2}$$

For an initial configuration $X^0\in\mathcal{A}^{\mathbb{Z}^d}$, and a sequence of independent samples $(r^t)_{t\in\mathbb{N}}$, $r^t\in[0,1]^{\mathbb{Z}^d}$ distributed according to $\mathcal{U}$, we can then define recursively a space-time diagram by

$$X^{t+1} = \phi(X^t,r^t).$$

In Chap. 2, we will give examples of PCA as well as some representations of space-time diagrams.

1.3.4 Interacting particle systems

The analogue of PCA in continuous time are (finite-range) *interacting particle systems (IPS)* [Lig05]. IPS are also characterised by a finite neighbourhood $\mathcal{N} \subset \mathbb{Z}^d$, and a transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A})$. We attach random and independent clocks to the cells of $\mathbb{Z}^d$. For a given cell, the instants of $\mathbb{R}_+$ at which the clock rings form a Poisson process of parameter 1. Let x^t be the configuration at time $t \geq 0$ of the process. If the clock at cell k rings at instant t, the state of the cell k is updated according to the probability measure $f((x^t_{k+v})_{v\in\mathcal{N}})$. This defines a transition semigroup $F = (F^t)_{t\in\mathbb{R}_+}$, with $F^t : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{X})$. If the initial measure is μ, the distribution of the process at time t is given by μF^t. A measure μ is an *invariant measure* if $\mu F^t = \mu$ for all $t \in \mathbb{R}_+$.

Observe that PCA are discrete-time Markov chains, while IPS are continuous-time Markov processes. In a PCA, all cells are updated at each time step, in a "synchronous" way. On the other hand, for an IPS, the updating is "fully asynchronous". Indeed, the probability of having two clocks ringing at the same instant is 0.

1.4 Statistical mechanics of PCA

The connection with equilibrium statistical mechanics is essential to understand PCA. Here, the results are presented without proofs. In addition to the references given through this section, we refer to the article of Lebowitz, Maes and Speer [LMS90]. A comprehensive survey in french can also be found in the thesis of Louis [Lou02].

1.4.1 Gibbs measures

We first introduce some background of statistical mechanics. Let Γ be the set of vertices of a non-oriented graph, locally finite (in Sec. 1.4.2, we will consider a graph of set of vertices $\mathbb{Z}^{d+1}$, but in Sec. 1.4.3, we will also introduce a graph of vertices indexed by $\mathbb{Z}^d \times \{0,1\}$). We use the notation $C \Subset \Gamma$ to specify that C is a finite subset of Γ. For $K \subset \Gamma$, $x \in \mathcal{A}^\Gamma$, let $x_K \in \mathcal{A}^K$ be the restriction of x to K.

Definition 1.10 (Gibbs potential)**.** A *Gibbs potential* on Γ is a family $\varphi = (\varphi_C)_{C\Subset\Gamma}$ of functions $\varphi_C : \mathcal{A}^C \to \mathbb{R}$.

By convention, for $x \in \mathcal{A}^\Gamma$, we set $\varphi_C(x) = \varphi_C(x_C)$.

The potential φ has a *finite range* if there exists $L \in \mathbb{N}$ such that $\varphi_C \equiv 0$ as soon as the set C contains two elements at distance larger than L in the graph. In the following, we will consider only finite range potentials. For a set $K \subset \Gamma$, we define $\mathcal{V}(K)$ as the union of the sets $C \Subset \Gamma$ such that $C \cap K \neq \emptyset$ and $\varphi_C \not\equiv 0$. We also define $\partial K = \mathcal{V}(K) \setminus K$.

Definition 1.11 (Gibbs measure)**.** A measure μ on $\mathcal{A}^\Gamma$ is a *Gibbs measure* with potential φ if for any finite sets J and K such that $\mathcal{V}(K) \subset J$,

$$\mu\Big([x_K] \;\Big|\; [x_{J\setminus K}]\Big) = \frac{1}{Z(x_{\partial K})} \exp\Big(-\sum_{C\cap K\neq\emptyset} \varphi_C(x)\Big),$$

as soon as $\mu([x_{J\setminus K}]) > 0$, where $Z(x_{\partial K})$ is a normalising factor depending only on $x_{\partial K}$.

In the classical approach, the Kolmogorov extension theorem defines a probability measure given a family of consistent finite-dimensional distributions. Here, the marginals are specified through their *conditional* distributions. This is referred to as the DLR approach, in tribute to Dobrushin, Lanford and Ruelle. One can prove that for each potential φ, there exists always at least one associated Gibbs measure. But there can be several ones. We denote by $\mathcal{G}(\varphi)$ the set of Gibbs measure with potential φ. It is a non-empty, convex and compact set of $\mathcal{M}(\mathcal{A}^\Gamma)$. We say that there is a *phase transition* if there are several Gibbs measures associated to the same potential. An important result is that on the graph $\Gamma = \mathbb{Z}$, there are no phase transitions [Geo11].

Gibbs measures are Markov random fields. As already mentioned, there is in fact an equivalence between Markov random fields and Gibbs measures with finite range potentials [Spi71, Geo11].

1.4.2 PCA and equilibrium statistical mechanics

In this section, we present the correspondence between the stationary space-time diagrams of a PCA defined on $\mathbb{Z}^d$, and the Gibbs measures corresponding to a related potential defined on the graph $\mathbb{Z}^{d+1}$.

Let us consider a positive-rate PCA F on $\mathbb{Z}^d$ of neighbourhood $\mathcal{N}$ and local function f. We define Γ as the graph of vertices $\mathbb{Z}^{d+1}$, with edges between $(k, n+1)$ and $(k+v, n)$ for any $(k, n, v) \in \mathbb{Z}^d \times \mathbb{Z} \times \mathcal{N}$. For $(k, n) \in \mathbb{Z}^d \times \mathbb{Z}$, we define the set $F(k, n) = \{(k+v, n)\,; v \in \mathcal{N}\} \cup \{(k, n+1)\}$. We define the potential φ on $\mathbb{Z}^{d+1}$ by:

$$\varphi_{F(k,n)}(x) = -\log f((x^n_{k+v})_{v\in\mathcal{N}})(x^{n+1}_k),$$

and $\varphi_C \equiv 0$ if there are no $(k, n) \in \mathbb{Z}^d \times \mathbb{Z}$ such that $C = F(k, n)$.

This potential is invariant under temporal and spatial translations in the space-time diagram.

Proposition 1.1 ([GKLM89]). *The translation-invariant Gibbs measures for φ correspond exactly to the translation-invariant space-time diagrams for F.*

Corollary 1.1. *There is a phase transition for φ if and only if F has several invariant measures.*

In Prop. 1.1, the difficulty consists in showing that any translation-invariant Gibbs measures correspond to the invariant space-time diagrams for F. The proof uses conditional entropy and the variational principle. The other direction, stating that an invariant space-time diagram for F is a Gibbs measure for φ, is easier and remains true without the positive rates assumption.

Let us consider a PCA on $\mathbb{Z}$, of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$. We consider a portion of space-time diagram as in Fig. 1.1, with time going up. A consequence of Prop. 1.1 is that if the space-time diagram is translation-invariant, the conditional distribution of the central cell knowing all the other values of the space-time diagram is equal to its conditional distribution knowing the states of the 10 neighbouring cells represented on the figure. Conditionally to the values of these cells, the central cell takes the value σ with probability:

$$\frac{1}{Z}\, f(a_1, a_2, a_3)(\sigma)\; f(b_1, b_2, \sigma)(c_1)\; f(b_2, \sigma, b_3)(c_2)\; f(\sigma, b_3, b_4)(c_3),$$

where

$$Z = \sum_{\alpha\in\mathcal{A}} f(a_1, a_2, a_3)(\alpha)\; f(b_1, b_2, \alpha)(c_1)\; f(b_2, \alpha, b_3)(c_2)\; f(\alpha, b_3, b_4)(c_3).$$

It follows from Prop. 1.1 that any invariant measure of a PCA on $\mathbb{Z}^d$ which is translation-invariant is the projection of a Gibbs measure defined on $\mathbb{Z}^{d+1}$. But in general, projections of Gibbs measures are not Gibbs measures.

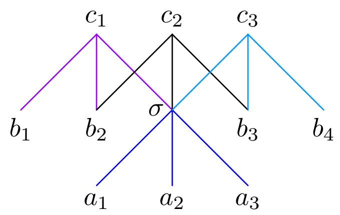


Figure 1.1: Illustration of Prop. 1.1.

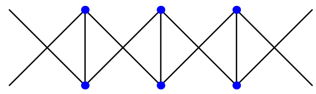


Figure 1.2: Doubling graph of a PCA on $\mathbb{Z}$ of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$.

1.4.3 Reversibility

From any finite-range potential φ, one can define an interacting particle system such that the set of Gibbs measures $\mathcal{G}(\varphi)$ is equal to the set of reversible invariant measures of the dynamics [Lig05]. This is not true for PCA [Daw75, DKT90].

We say that an invariant measure μ of a PCA F is *reversible* if a stationary space-time diagram $(X_k^n)_{n\in\mathbb{Z}}$ associated to μ has the same distribution as the random field $(X_k^{-n})_{n\in\mathbb{Z}}$ obtained when reversing the direction of time, or equivalently, if (X^0, X^1) has the same law as (X^1, X^0) when X^0 is distributed according to μ and X^1 obtained by one iteration of the PCA from X^0. This means that both the transitions, from $t = 0$ to $t = 1$ and from $t = 1$ to $t = 0$, are described by the same PCA F. The measure μ is said to be a *quasi-reversible* measure of F if the transition from $t = 1$ to $t = 0$ is described by a PCA, possibly different from F.

Let us consider a PCA F on $\mathbb{Z}^d$, of neighbourhood $\mathcal{N}$. The *doubling graph* of F is the undirected graph of set of vertices $\mathbb{Z}^d \times \{0, 1\}$, with edges between (k, t) and $(k + v, 1 - t)$ for any $(k, v, t) \in \mathbb{Z}^d \times \mathcal{N} \times \{0, 1\}$. Fig. 1.2 represents the doubling graph for a PCA on $\mathbb{Z}$ of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$.

To a measure μ on $\mathbb{Z}^d$, we can associate a measure $\overline{\overline{\mu}}$ on $\Gamma = \mathbb{Z}^d \times \{0, 1\}$, corresponding to the distribution of (X^0, X^1) when X^0 is distributed according to μ, and X^1 obtained from X^0 with the PCA.

One can prove that quasi-reversibility is equivalent to the condition that $\overline{\overline{\mu}}$ is a Markov random field on Γ. The following property follows from this observation.

Proposition 1.2 ([Vas78]). *An invariant measure μ of a positive-rate PCA F is quasi-reversible if and only if the corresponding measure $\overline{\overline{\mu}}$ on $\Gamma = \mathbb{Z}^d \times \{0, 1\}$ is a Gibbs measure.*

A potential φ on Γ is a *pair potential* if $\varphi_C \equiv 0$ as soon as C is not constituted of a single site or of two adjacent vertices. In the context of Prop. 1.2, the Gibbs measure $\overline{\overline{\mu}}$ on $\Gamma = \mathbb{Z}^d \times \{0, 1\}$ can be shown to be defined by a pair potential. If we want the measure μ to be not only quasi-reversible, but reversible, this pair potential has to satisfy symmetry conditions.

Let $\mathcal{N}$ be a symmetric neighbourhood, and let us consider a function $\phi : \mathcal{A} \to \mathbb{R}$, as well as functions $\phi_v : \mathcal{A}^2 \to \mathbb{R}$, for $v \in \mathcal{N}$, such that $\phi_v(a, b) = \phi_{-v}(b, a)$. We define a symmetric pair potential on $\Gamma = \mathbb{Z}^d \times \{0, 1\}$ by: $\varphi_{\{s\}} = \phi$ for any $s \in \Gamma$, $\varphi_{\{(j,0),(k,1)\}}(x_j, y_k) = \phi_v(x_j, y_k)$ for $j, k \in \mathbb{Z}^d$ satisfying $v = j - k \in \mathcal{N}$, and $\varphi_C \equiv 0$ otherwise. By definition, $\varphi_{\{(k,0),(j,1)\}}(y_k, x_j) = \varphi_{\{(j,0),(k,1)\}}(x_j, y_k)$.

As a corollary of Prop. 1.2, we get the next proposition.

Proposition 1.3 ([Vas78, KV80, DKT90]). *If F is a positive-rate PCA, a necessary and sufficient condition for the existence of at least one reversible invariant measure is the existence of a symmetric pair potential φ such that the local function f can be represented by:*

$$f((x_v)_{v\in\mathcal{N}})(\alpha) = \frac{1}{Z(x_{\mathcal{N}})} \exp\Big(-\phi(\alpha) - \sum_{v\in\mathcal{N}} \phi_v(x_v,\alpha)\Big), \tag{1.3}$$

where $Z(x_{\mathcal{N}}) = \sum_{\gamma\in\mathcal{A}} \exp\Big(-\phi(\gamma) - \sum_{v\in\mathcal{N}} \phi_v(x_v,\gamma)\Big)$. Under this condition, the reversible measures are exactly the projections on $\mathbb{Z}^d$ of the Gibbs measures on $\Gamma = \mathbb{Z}^d \times \{0,1\}$ of potential φ derived from ϕ, that are equal on both copies of $\mathbb{Z}^d$. They are themselves Gibbs measures on $\mathbb{Z}^d$, of potential $\hat{\varphi}$ defined by:

$$\hat{\varphi}_{\{k\}} = \phi \ \text{ for any } \ k\in\mathbb{Z}, \quad \hat{\varphi}_{k+\mathcal{N}}((x_{k+v})_{v\in\mathcal{N}}) = -\log Z(x_{k+\mathcal{N}}), \quad \text{and } \ \hat{\varphi}_C \equiv 0 \ \text{ otherwise.}$$

We can associate this proposition to the following theorem, proved in the thesis of Dai Pra [DP92, DPLR02], to obtain Corollary 1.2.

Theorem 1.1. *If there exists a potential $\hat{\varphi}$ on $\mathcal{A}^{\mathbb{Z}^d}$ such that at least one translation-invariant measure μ, invariant for the PCA F, is a Gibbs measure with respect to the potential $\hat{\varphi}$, then all the translation-invariant measures that are invariant for F are Gibbs measures of potential $\hat{\varphi}$.*

Corollary 1.2. *If a PCA F has a reversible measure, then all its invariant measures are reversible, and the set of its invariant measures is equal to the set of Gibbs measures of potential $\hat{\varphi}$ that are left invariant by F. In particular, if there is no phase transition for $\hat{\varphi}$, then F is ergodic.*

Example 1.1. [[DKT90]] Consider the set of cells $\mathbb{Z}$, the alphabet $\mathcal{A} = \{0,1\}$, and the neighbourhood $\mathcal{N} = \{-1,0,1\}$. The family of positive rates reversible PCA can be described by three parameters $c_1, c_2, c_3 > 0$, such that for $x,y,z\in\mathcal{A}$,

$$f(x,y,z)(0) = \frac{1}{1+c_1c_2^{x+z}c_3^y} \qquad f(x,y,z)(1) = \frac{c_1c_2^{x+z}c_3^y}{1+c_1c_2^{x+z}c_3^y}.$$

Since there are no phase transitions for one-dimensional Gibbs potentials, these PCA are ergodic, and their unique invariant measures are 2-Markov.

When the neighbourhood is asymmetric, it can be relevant to modify the representation of the space-time diagram in order to recover a symmetric neighbourhood. For example, if $E = \mathbb{Z}$, $\mathcal{N}_1 = \{0,1\}$, it appears natural to shift by 1/2 the image X^1 of the initial configuration X^0, which amounts to consider that the neighbourhood is in fact $\mathcal{N}_1' = \{-1/2, 1/2\}$. In the same way, if $E = \mathbb{Z}^2$, $\mathcal{N}_2 = \{(0,0),(0,1),(1,0),(1,1)\}$, one can shift the image of a configuration by the vector $(1/2,1/2)$, which corresponds to the choice of a symmetric neighbourhood $\mathcal{N}_2' = \{(\pm 1/2, \pm 1/2)\}$. The respective doubling graphs are then represented as in Fig. 1.3. The notion of reversibility and all the results of this section can be extended to this context.

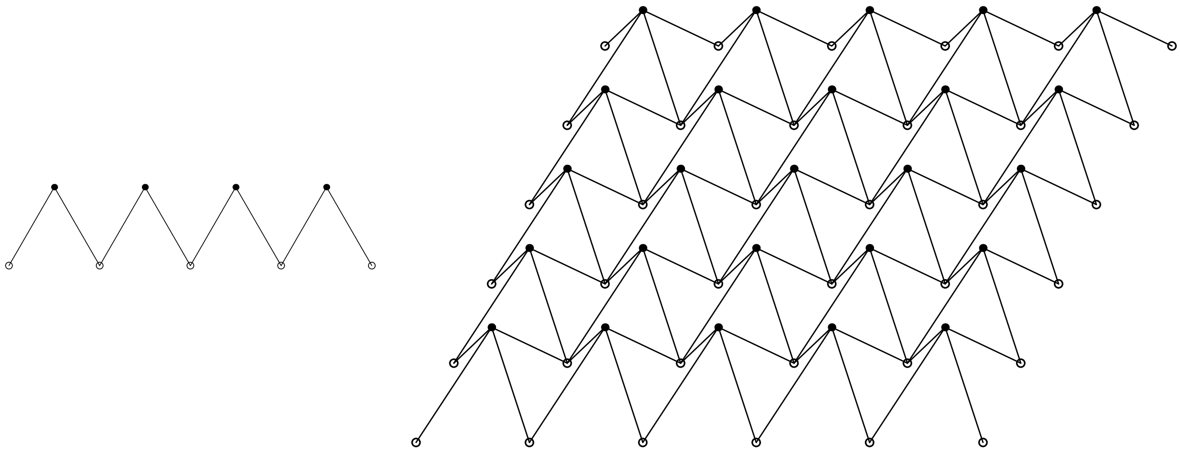


Figure 1.3: Symmetric doubling graphs associated to the neighbourhoods $\mathcal{N}_1$ and $\mathcal{N}_2$.

Part I

PCA and their invariant measures: a general approach

Chapter 2

Different viewpoints on PCA

Then I thought of something, all of a sudden. "Hey, listen," I said. "You know those ducks in that lagoon right near Central Park South? That little lake? By any chance, do you happen to know where they go, the ducks, when it gets all frozen over? Do you happen to know, by any chance?" I realized it was only one chance in a million.

– J. D. Salinger, *The Catcher in the Rye*

Contents

In this chapter, we first discuss the definition of PCA given in Chap. 1, and then present different viewpoints on PCA, coming from statistical mechanics, computation theory, and biology modelling.

2.1 Discussion of the definition of PCA

In the definition of PCA given in Chap. 1, Sec. 1.3, we have made the assumption that the updatings of different cells were independent conditionally to the value of their neighbours. Depending on the modelling context, it might be more convenient to weaken this assumption. For instance, it is practical to define the discrete TASEP model, see below, as a generalised PCA.

Example 2.1 (Discrete TASEP)**.** TASEP stands for *Totally Asymetric Simple Exclusion Process.* Here, we consider a discrete version of the model. The continuous-time version is a standard and widely studied model of interacting particle systems.

The alphabet is $\mathcal{A} = \{0, 1\}$, a 1 standing for a particle and a 0 for an empty space, and at each time step, if its right neighbour is empty, a particle jumps to the right with some probability $p \in (0, 1)$.

Strictly speaking, this model is not a PCA since the updates of two adjacent cells are dependent: if a cell is in state 1 and the neighbouring right cell in state 0, either the states of both cell will change (probability p) or none of them (probability $1 - p$).

To include such models, alternative definitions of PCA have been proposed, and we now present one of them [AST13]. The set of cells is still $E = \mathbb{Z}^d$, and we have a finite neighbourhood $\mathcal{N}$. But in addition to the finite alphabet $\mathcal{A}$, we assume that we are given a finite set $\mathcal{R}$ called the random symbols, and the transition function is now a (deterministic) function: $\varphi : \mathcal{A}^{\mathcal{N}} \times \mathcal{R}^{\mathcal{N}} \to \mathcal{A}$. At each time step, starting from some configuration in $\mathcal{A}^{\mathbb{Z}^d}$, a configuration of random symbols is chosen in $\mathcal{R}^{\mathbb{Z}^d}$ according to a given Bernoulli product measure, and then, the transition function φ is applied to obtain a new configuration in $\mathcal{A}^{\mathbb{Z}^d}$. This also defines a Markov chain on $\mathcal{A}^{\mathbb{Z}^d}$.

With this setting, the previous TASEP model can be implemented by introducing a set $\mathcal{R} = \{a, i\}$. Each cell is allocated independently the random symbol a with probability p ("active" cell) and i with probability $1-p$ ("inactive" cell), the neighbourhood is $\mathcal{N} = \{-1, 0, 1\}$ and the function φ is defined, for $x_{-1}, x_0, x_1 \in \mathcal{A}$ and $r_{-1}, r_0, r_1 \in \mathcal{R}$, by: $\varphi((x_{-1}, x_0, x_1), (r_{-1}, r_0, r_1)) =$

$$\begin{cases} 0 & \text{if} \quad [x_{-1} = x_0 = 0] \vee [x_{-1} = 1, x_0 = 0, r_{-1} = i] \vee [x_0 = 1, x_1 = 0, r_0 = a] \\ 1 & \text{if} \quad [x_0 = x_1 = 1] \vee [x_{-1} = 1, x_0 = 0, r_{-1} = a] \vee [x_0 = 1, x_1 = 0, r_0 = i] \end{cases}.$$

This definition allows more flexibility and can be favoured in some cases, for practical purposes. But, in essence, it is not really different from the one given in Chap. 1. Indeed, given a set $\mathcal{R}$ of symbols that are to be sampled according to a Bernoulli product measure μ_p, and a transition function $\varphi : \mathcal{A}^{\mathcal{N}} \times \mathcal{R}^{\mathcal{N}} \to \mathcal{A}$, we can define a PCA F (in the sense of Def. 1.4), on the extended alphabet $\mathcal{B} = \mathcal{A} \times \mathcal{R}$ that presents the same behaviour. Its transition function is defined for $(a_v, r_v)_{v \in \mathcal{N}} \in \mathcal{B}^{\mathcal{N}}$ by:

$$f((x_v, r_v)_{v \in \mathcal{N}})(a, b) = \begin{cases} p_b & \text{if } \phi((x_v, r_v)_{v \in \mathcal{N}}) = a, \\ 0 & \text{otherwise.} \end{cases}$$

The difference is that one is now interested in the projections of the trajectories on the first coordinates, giving configurations in $\mathcal{A}^{\mathbb{Z}^d}$. In fact, the PCA F can be viewed as operating on two tapes, one with the $\mathcal{A}$-symbols and one with the $\mathcal{R}$-symbols. At each time step, the $\mathcal{A}$-tape is updated by applying φ, and the $\mathcal{R}$-tape is updated by choosing brand new random symbols, independently of everything. Concentrating on the $\mathcal{A}$-tape of the classical PCA F, we recover exactly the behaviour of the "generalised" PCA.

If we do not assume anymore that the elements of the random symbols are chosen according to a Bernoulli product measure but from any probability measure on $\mathcal{R}^{\mathbb{Z}^d}$, the models obtained are no more directly included in our definition of PCA.

2.2 Traffic models and queues

Some particular PCA are tightly related to queues models. Exhibiting a correspondence between a PCA and a queue system can be an efficient tool to prove properties of the invariant measure, by transposing to PCA well-known results of queueing theory.

Let us come back to the discrete TASEP model. Despite its simplicity, it shows a rich behaviour, and appears in very different contexts, such as random growth models (last-passage percolation on $\mathbb{Z}^2$), and random domino tilings [CEP96, JPS98]. But first, there is an important connection between the TASEP and tandem queues, that is described in Fig. 2.1.

Precisely, a configuration of $\{0, 1\}^{\mathbb{Z}}$ is interpreted as a bi-infinite sequence of queues: each 0 corresponds to a queue with an infinite capacity buffer and the the consecutive 1's on its left (if any) correspond to the customers waiting in line at the queue. The dynamics of the TASEP PCA translates as follows for the queueing model: at a given queue, customers are served one by one in their order of arrival, their service time is positive geometric of parameter $1 - p$, and upon being served a customer joins instantaneously the next queue to its right.

This is a standard model in queueing theory, which enjoys remarkable properties. We can backtrack the results obtained for the queueing model to the TASEP PCA, to get next result.

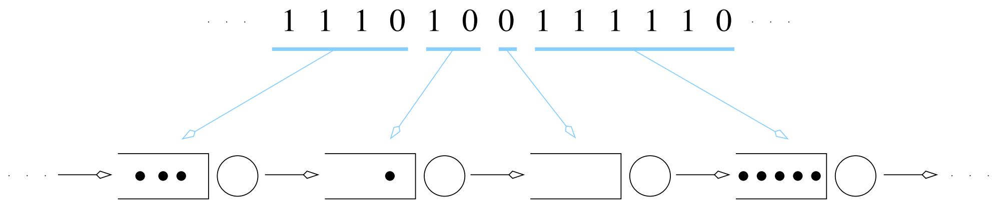


Figure 2.1: From the TASEP to tandem queues.

For any $q \in (0, p)$, define the Markov measure ν_q of transition matrix:

$$Q = \begin{pmatrix} \frac{p-q}{p} & \frac{q}{p} \\ \frac{p-q}{p(1-q)} & \frac{q(1-p)}{p(1-q)} \end{pmatrix},$$

so that

$$\nu_q([0]) = \frac{p-q}{p-q^2}, \qquad \text{and} \qquad \nu_q([1]) = \frac{q(1-q)}{p-q^2}.$$

Proposition 2.1. *Consider the TASEP PCA T_p for $p \in (0,1)$. For any $q \in (0,p)$, the Markovian measure ν_q is an invariant measure of T_p.*

Proof. To prove that ν_q is an invariant measure, a first way to proceed is simply to check "by hand" that it is left invariant by the dynamics [JPS98, Sec. 4.1]. The verification has to be made for all cylinders. However, such a proof is not very informative, hiding in particular how the right invariant measures were guessed. Let us sketch instead a queueing theoretic argument.

Consider a single queue with "positive gometric" services of parameter p, that is, at each time step, if the buffer is non-empty, there is a departure with probability p, independently of the past. Asume that the arrivals to the queue are distributed according to $\mathcal{B}_q^{\otimes \mathbb{Z}}$, $q \in (0, p)$, that is, at each time step, there is an arrival with probability q, independently of the past. Then, it can be checked immediately that the equilibrium queue-length process (which is a birth-and-death Markov chain) is distributed according to π defined by:

$$\pi_0 = \frac{p-q}{p}, \qquad \forall n \geq 1, \ \pi_n = \left(\frac{q(1-p)}{p(1-q)}\right)^{n-1} \frac{q}{p(1-q)} \frac{p-q}{p}. \tag{2.1}$$

Furthermore, the equilibrium departure process from the queue is distributed according to $\mathcal{B}_q^{\otimes \mathbb{Z}}$. This last result is known as a "Burke-type" theorem, since an analogous result was first proved by Burke [Bur56] in the context of continuous-time queues with exponential services, with a later and transparent proof by Reich [Rei57]. For the present setting, we also refer to the article of Draief, Mairesse and O'Connel [DMO05].

Consider now several of the above queues in tandem. The departure process from a queue is the arrival process to the next queue. According to the Burke-type theorem, the arrival process to each queue is distributed according to $\mathcal{B}_q^{\otimes \mathbb{Z}}$. An even stronger result holds:

in equilibrium, the different queue-lengths are distributed according to π and *independent*. This last result is known as a "product-form" theorem.

By letting the number of queues go to infinity, and by using the translation of Fig. 2.1, one can retrieve the desired result. In particular, the matrix Q can be derived from (2.1). For instance, $Q_{0,0}$, the probability to have a 0 followed by a 0 in the TASEP, is equal to $\pi_0 = (p-q)/p$, the probability to have an empty queue in the queueing model (see Fig. 2.1). □

The discrete TASEP thus admits a whole family of Markov invariant measures. One can prove that the invariant measures of T_p which are translation-invariant are precisely the convex combinations of ν_q, $q \in (0,p)$, $\delta_{0^{\mathbb{Z}}}$, and $\delta_{1^{\mathbb{Z}}}$.

As discussed in the previous section, the discrete TASEP is not a PCA in the strict sense, but to obtain a PCA model, it is sufficient to extend the alphabet by adding a one-bit information telling in advance if a particle is likely to jump at the next step of time or not.

We present now a dual model which is striclty speaking a PCA on the alphabet $\mathcal{A} = \{0,1\}$. Its neighbourhood is $\mathcal{N} = \{-1,0,1\}$ and the local rule is given by:

$$f(0,0,1) = p\delta_1 + (1-p)\delta_0, \quad f(1,1,0) = (1-p)\delta_1 + p\delta_0, \text{ and } f(x,y,z) = \delta_y \text{ otherwise.}$$

The interpretation as a queue system is yet as follows: a sequence of n consecutive 0 (resp. 1) represents a queue of $n-1$ customers. The queues are in tandem and at each step of time, the first customer of a queue is served with probability p and then jumps to the next queue. Once again, results of queueing theory allow to exhibit Markov invariant measures for that PCA.

2.3 Directed animals

For $p \in (0,1)$, let us introduce the PCA F_p on $\mathbb{Z}$ of alphabet $\mathcal{A} = \{0,1\}$, and neighbourhood $\mathcal{N} = \{-1,0\}$, defined by the local function

$$f(x,y)(1) = \begin{cases} p & \text{if } x = y = 0, \\ 0 & \text{otherwise,} \end{cases}$$

see Fig. 2 of the introductory chapter.

As a consequence of Th. 4.2 of Chap. 4, this PCA has a Markov invariant measure, of transition matrix

$$Q = \begin{pmatrix} 1-a & a \\ 1-b & b \end{pmatrix},$$

with parameters given by:

$$a = \frac{2p^2 - p - 1 + \sqrt{1+2p-3p^2}}{2p^2}, \qquad b = \frac{1+p-\sqrt{1+2p-3p^2}}{2p}. \tag{2.2}$$

We denote by ν_p this Markov invariant measure of F_p.

We will see that this PCA plays an unexpected role for the enumeration of *directed animals*. So called "animals" are classical objects in combinatorics. They are related to (site) percolation models. The ultimate goal is to count the number of animals of a given size. There exist two variants: classical and directed animals. Here we consider only directed animals which are simpler to study. We will see that a particular PCA can be introduced for the enumeration of directed animals.

Consider the directed infinite graph:

$$(\mathbb{Z}\times\mathbb{N}, A), \qquad A = \big\{(i,j)\to(i+v,j+1) \mid (i,j)\in\mathbb{Z}\times\mathbb{N}, v\in\{0,1\}\big\}\,.$$

Let C be a non-empty finite subset of $\mathbb{Z}$. A *directed animal of base* C is a finite subset E of $\mathbb{Z}\times\mathbb{N}$ such that:

- $E\cap(\mathbb{Z}\times\{0\}) = C\times\{0\}$;
- $\forall x\in E,\ \exists x_0\in C\times\{0\}, x_1,\dots,x_{n-1}\in E, x_n = x,\ \forall i,\ x_i\to x_{i+1}$.

A *directed animal* is a directed animal of base $\{0\}$, see Fig. 2.2.

Figure 2.2: A directed animal (left), *not* a directed animal (right).

It is customary in combinatorics to count objects according to their size, and to encapsulate all the information in a formal series. The *counting series* of directed animals of base C, respectively of directed animals, is the formal series defined by:

$$S_C(x) = \sum_{E:\ \text{DA base}\,C} x^{|E|}, \qquad S(x) = S_{\{0\}}(x)\,. \tag{2.3}$$

The coefficient of x^n in $S(x)$ is the number of directed animals of size n.

The goal of the section is to give a sketch of the proof of next theorem, which is based on a connection with the PCA F_p.

Theorem 2.1 ([Dha83])**.** *The counting series of directed animals is given by:*

$$S(x) = \frac{1}{2}\left(\frac{\sqrt{1-2x-3x^2}}{1-3x} - 1\right). \tag{2.4}$$

From a combinatorics point of view, this is an ideal result, since $S(x)$ is algebraic and defined in an explicit way. By performing a Taylor expansion around 0 of $S(x)$, we get the first terms of the counting series.

Proof. Removing the bottom line of a directed animal provides either the empty set or a new animal on the lines $\{1,2,\dots\}$. This simple observation provides a recurrence relation on counting series:

$$S_C(x) = x^{|C|}\Big(\sum_{D\subset C+\mathcal{N}} S_D(x)\Big), \tag{2.5}$$

with the convention $S_\emptyset(x) = 1$.

Recall that ν_p is the Markov invariant measure of F_p, defined above. For a finite subset C of $\mathbb{Z}$, set

$$s_C(p) = \nu_p([1\cdots 1]_C).$$

By definition, it is the probability, under the measure ν_p, that all of the sites of C are in state 1. Consider a sequence $(X_i)_{i\in\mathbb{Z}}$ distributed according to ν_p, and let $(Y_i)_{i\in\mathbb{Z}}$ be a realisation of the image of $(X_i)_{i\in\mathbb{Z}}$ by the PCA F_p. It implies that $(Y_i)_{i\in\mathbb{Z}}$ is also distributed according to ν_p. We have:

$$s_C(p) = \mathbb{P}\big(\forall i \in C,\ Y_i = 1\big) = \mathbb{P}\big(\forall i \in C + \mathcal{N},\ X_i = 0\big)\, p^{|C|}.$$

According to the inclusion-exclusion principle, we get:

$$\mathbb{P}\big(\forall i \in C + \mathcal{N},\ X_i = 0\big) = \sum_{D\subset C+\mathcal{N}} (-1)^{|D|} P\big(\forall i \in D,\ X_i = 1\big) = \sum_{D\subset C+\mathcal{N}} (-1)^{|D|} s_D(p),$$

with the convention $s_\emptyset(p) = 1$. So, we have:

$$s_C(p) = p^{|C|} \sum_{D\subset C+\mathcal{N}} (-1)^{|D|} s_D(p). \tag{2.6}$$

By comparing (2.3) and (2.6), we get that

$$S_C(-p) = (-1)^{|C|} s_C(p), \qquad S(-p) = -\nu_p([1]), \tag{2.7}$$

are possible solutions for the recurrence equations (2.5). This provides an unexpected relation between two a priori unrelated models.

Now we use the fact that we have an exact expression for the invariant measure ν_p, see (2.2). We obtain immediately $\nu_p([1]) = a/(1-b+a) = (\sqrt{1+2p-3p^2} - 1 - 3p)/(2+6p)$. By evaluating S formally according to (2.7), we obtain (2.4).

The last step consists in argueing that S is indeed the counting series. This requires an argument since the recurrence relations (2.5) may admit several families of solutions, with only one of them defining the counting series. □

Directed animals can be defined on other infinite regular graphs, and the connection with a PCA model still holds [Alb09]. In all cases where the counting series can be explicitly computed, it is done by using the PCA connection. The problem is that the invariant measure of the associated PCA cannot always be explicited.

2.4 From CA to PCA: noisy CA and α-asynchronous CA

We present two specifications of PCA that provide examples of particular interest. They both consist in considering a deterministic CA and "perturbating" the rule, either by doing random errors or by introducing asynchronism in the evolution.

Noisy CA. Let F be a deterministic CA on $\mathcal{A}^{\mathbb{Z}^d}$, of transition function f, and let $\varepsilon \in (0,1)$. We assume that with probability ε, when updating the value of a cell, a letter is chosen uniformly in $\mathcal{A}$, instead of applying the deterministic function f. If the probability of doing such errors is independent for different cells, this defines a PCA of transition function φ given by:

$$\varphi((x_v)_{v\in\mathcal{N}}) = (1-\varepsilon)\delta_{f((x_v)_{v\in\mathcal{N}})} + \varepsilon\, \mathtt{Unif},$$

where `Unif` denotes the uniform measure on $\mathcal{A}$. Starting from a deterministic CA, we thus define a positive-rate PCA. One can also consider other variants of faulty CA, by assuming that the noise is not uniformly distributed in $\mathcal{A}$.

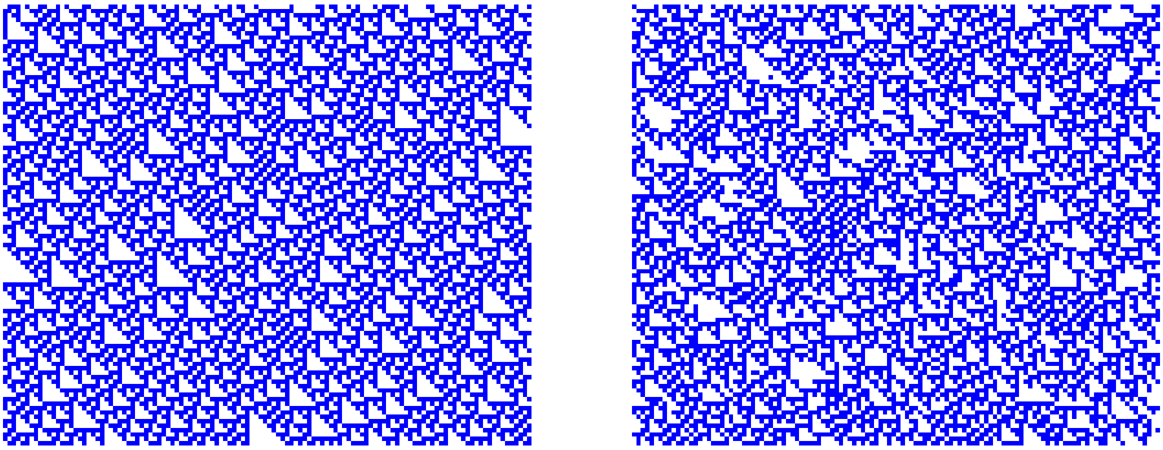


Figure 2.3: Space-time diagrams of the PCA of Ex. 2.2, for $\varepsilon = 0$, and $\varepsilon = 0.1$, starting from a uniform Bernoulli measure.

Example 2.2. Let $\mathcal{A} = \{0,1\}$, $E = \mathbb{Z}$, and $\mathcal{N} = \{0,1\}$. For some $\varepsilon \in (0,1)$, consider the local function

$$f(x,y) = (1-\varepsilon)\,\delta_{x+y} + \varepsilon\,\delta_{x+y+1},$$

where the sums $x+y$ and $x+y+1$ are computed modulo 2. For $\varepsilon = 0$, we obtain the deterministic additive CA defined by: $F(x)_i = x_i + x_{i+1}$ for any $i \in \mathbb{Z}$. For small values of $\varepsilon > 0$, the PCA can be interpreted as a perturbation of that deterministic CA, with some errors occurring in each cell independently. In Fig. 2.3, we represent two space-time diagrams, for respectively $\varepsilon = 0$ and $\varepsilon = 0.1$.

One can prove [DKT90, Chap. 16 and 17] that for any value of $\varepsilon \in (0,1)$, the PCA is ergodic and that its unique invariant measure is the uniform mesure λ, that is, the Bernoulli product measures of parameter $1/2$, defined by $\lambda = \mathcal{B}_{1/2}^{\otimes\mathbb{Z}}$. In Chap. 4, we will study criteria for having a Bernoulli product invariant measure. However, we can here give an elementary proof of the fact that λ is an invariant measure.

Let us look at the possible antecedents of a word $v \in \mathcal{A}^n$. If we know for each cell if its value has been computed using the sum rule (probability $1-\varepsilon$) or by the sum plus 1 (probability ε, event corresponding to an "error"), there are exactly two possible sequences $u \in \mathcal{A}^{n+1}$ whose outcome is v: the first term u_1 can be chosen to be either 0 or 1, and then, the rest of the sequence u is entirely determined. And these two words have the same probability $2^{-(n+1)}$ with respect to the measure λ. We have thus:

$$\lambda F([v]) = \sum_{\vartheta \in \{0,1\}^n} 2(1-\varepsilon)^{|\vartheta|_0}\varepsilon^{|\vartheta|_1}2^{-(n+1)},$$

where $\vartheta_i = 1$ corresponds to an "error" at cell i (event of probability ε). It follows that: $\lambda F([v]) = 2^{-n} = \lambda([v])$.

Noisy CA have been introduced in relation with the question of reliable computation. In the previous example, even for an arbitrarily small probability of noise ε, when iterating the CA, all the information of the initial configuration is lost: whatever the initial configuration is, the iterates of the PCA converge to the uniform measure. In Sec. 2.5, we will present a two-dimensional positive-rate PCA having several invariant measures. In that case, "something" can thus be remembered forever about the initial configuration. In Chap. 6, we will also come back to this question of reliable computation, which is related to the positive rates problem. In particular, we refer to that chapter for the bibliographical references.

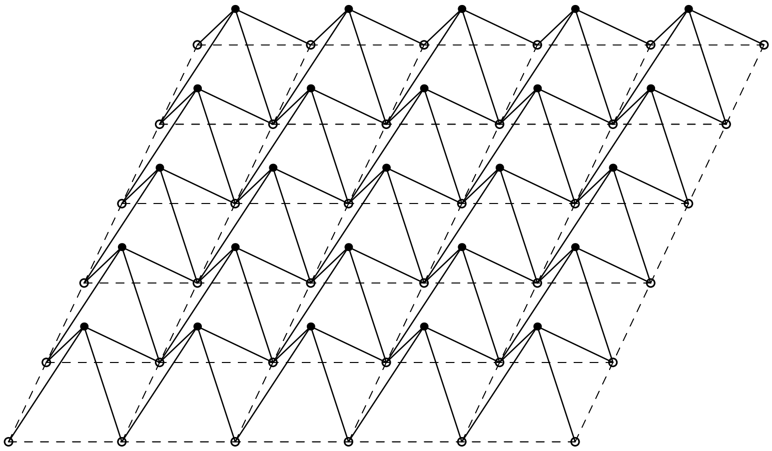


Figure 2.4: Symmetric doubling graph (continuous lines) of the PCA, defined on $\mathbb{Z}^2$ (dashed lines).

α-asynchronous CA. Let F be a deterministic CA on $\mathcal{A}^{\mathbb{Z}^d}$, of neighbourhood $\mathcal{N}$ containing the origin 0 of $\mathbb{Z}^d$, and of transition function f. Let $\alpha \in (0,1)$. We now assume that with probability $1-\alpha$, when updating the value of a cell, its current value is kept, instead of applying the deterministic function f. This defines a PCA of transition function φ given by:

$$\varphi((x_v)_{v\in\mathcal{N}}) = \alpha\, \delta_{f((x_v)_{v\in\mathcal{N}})} + (1-\alpha)\, \delta_{x_0}.$$

The α-asynchronous dynamics was studied experimentally and it was shown that the 256 elementary cellular automata produce various qualitative responses to asynchronism [FM05]. For some particular rules, when varying continuously the rate α, there appears a critical value at which the behaviour of the PCA presents a qualitative change [Fat09].

2.5 A two-dimensional non-ergodic PCA with positive rates

Using the statistical mechanics approach of Chap. 1, Sec. 1.4, one can make use of the Ising model to design a two-dimensional PCA with positive rates that is not ergodic [Vas78, KV80, DKT90].

We introduce the alphabet $\mathcal{A} = \{-1,+1\}$. We will define a non-ergodic positive-rate PCA on $\mathcal{A}^{\mathbb{Z}^2}$, of neighbourhood $\mathcal{N} = \{(0,0),(0,1),(1,0),(1,1)\}$. But as mentionned at the end of Sec. 1.4.3, we will prefer a symmetric representation of that neighbourhood.

Let us consider the doubling graph represented on the right part of Fig. 1.3. This graph is in fact isomorphic to $\mathbb{Z}^2$. We define in an analogous way as in Sec. 1.4.3 a symmetrical pair potential φ on that graph by setting $\phi \equiv 0$ (no-contribution from self-interaction) and $\phi_v(a,b) = -\beta ab$ for any edge v. This potential corresponds to the classical Ising model. It is known that for β large enough, this potential presents a phase transition: there exist at least two translation invariant Gibbs measures, of density of $+1$ respectively strictly larger and strictly smaller than $1/2$.

Let us set $\varepsilon = \exp(-4\beta)$. Like in (1.3), we introduce the PCA F on $\mathbb{Z}^2$ of neighbourhood $\mathcal{N} = \{(0,0),(0,1),(1,0),(1,1)\}$ defined by the transition function:

$$f(x,y,z,t)(+1) = \frac{\varepsilon^2}{1+\varepsilon^2},\ \frac{\varepsilon}{1+\varepsilon},\ \frac{1}{2},\ \frac{1}{1+\varepsilon},\ \frac{1}{1+\varepsilon^2}, \tag{2.8}$$

if there are respectively $0,1,2,3$ or 4 times the state $+1$ among x,y,z,t. This defines a positive-rate PCA.

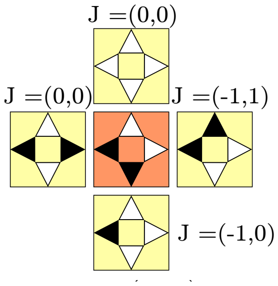


Figure 2.5: Example of neighbourhood configuration for the swarming model [BFC11].

By Prop. 1.3, any Gibbs measure $\overline{\overline{\mu}}$ of potential φ on the doubling graph provides an invariant measure μ for the PCA when projecting it on the lattice on which the PCA is defined (this lattice is the dashed grid on Fig. 2.4, whereas the doubling graph is represented with continuous lines).

As a consequence, if β is large enough (corresponding to small values of ε), this PCA has at least two different invariant measures of density of $+1$ respectively strictly larger than $1/2$ and strictly smaller. Summarising the above, we obtain next result.

Proposition 2.2. *Consider the positive-rate PCA defined on the set of cells $\mathbb{Z}^2$, with alphabet $\mathcal{A} = \{-1, +1\}$, neighbourhood $\mathcal{N} = \{(0,0), (0,1), (1,0), (1,1)\}$, and local function f defined by (2.8). For ε small enough, this PCA is non-ergodic with at least two invariant measures.*

In Sec. 6.2.3, we will describe another example of non-ergodic two-dimensional PCA with positive rates. This PCA, proposed by Toom, has the additional property of *robustness*: a small perturbation of its probability transitions still preserves non-ergodicity.

2.6 PCA as a modelling tool: example of the swarming model

PCA have been widely used to model some physical and biological phenomena, following the advice of Mark Kac: "Be wise, discretise!". In this section, we present a model of swarming, that has been introduced in order to understand how a collective motion can emerge from a decentralised organisation, as observed for flock of birds. Beyond being interesting for modelling, this PCA is also very interesting from a mathematical point of view: when varying the parameters, different behaviours appear. The figures of this section, as well as the qualitative comments on this model, are extracted from the works of Bouré, Chevrier and Fatès [BFC11, BFC13].

The set of cells is $E = \mathbb{Z}^2$, and we divide each cell into four sites, represented by triangles pointing east, west, north or south, each one being potentially occupied or empty, as represented in Fig. 2.5. This corresponds to working with the alphabet $\mathcal{A} = \{0,1\}^4$ of size $2^4 = 16$, with for $c \in \mathcal{A}$, $c_1 = 1$ (resp. c_2, c_3, c_4) if there is a particle pointing to the east (resp. west, north, south). We denote the number of particles by $|c| = c_1 + c_2 + c_3 + c_4$, it can be any integer between 0 and 4.

We define $n_1 = (1,0), n_2 = (-1,0), n_3 = (0,1), n_4 = (0,-1)$, and we set $\mathcal{N} = \{n_1, n_2, n_3, n_4\}$. We also introduce $\mathcal{N}_0 = \mathcal{N} \cup \{0,0\}$, which is known as the von Neumann neighbourhood. The *local flux* of $c \in \mathcal{A}$, is defined by

$$J(c) = \sum_{i=1}^{4} c_i n_i.$$

The dynamics of the model consists in the successive applications of two transition rules, applied to all cells synchronoulsy:

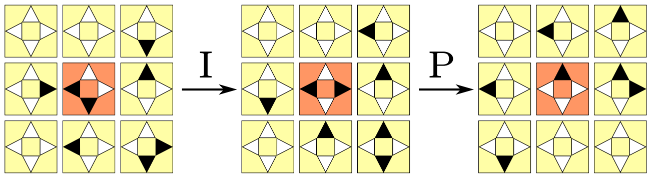


Figure 2.6: The interaction and the propagation rules [BFC11].

- the **interaction** rule reorganises the particles within each cell,
- the **propagation** rule moves all particles according to the direction they occupy.

Fig. 2.6 presents an example of composition of these two rules.

The propagation rule is deterministic. If we see each cell as being divided into four parts (E, W, N, S) that can be occupied or not, it corresponds to four shifts: all the particles located in the E position of a cell are shifted to the east (and after the move, still occupy an E position in their new cell), and the same for W, N, and S particles.

The interaction rule is a PCA that conserves the number of particles in each cell. Let us describe this step. Let $x \in \mathcal{A}^{\mathbb{Z}^2}$ be some configuration, and let $k \in \mathbb{Z}^2$. After the interaction step, the cell k will be in state $c \in \mathcal{A}$ with probability 0 if $|x_k| \neq |c|$, and if $|x_k| = |c|$, with probability:

$$\frac{1}{Z(x_{k+\mathcal{N}_0})} \exp\Big(\alpha \sum_{i=1}^{4} J(c) \cdot J(x_{k+n_i})\Big),$$

where $Z_k(x)$ is the normalisation factor, defined by:

$$Z(x_{k+\mathcal{N}_0}) = \sum_{|c|=|x_k|} \exp\Big(\alpha \sum_{i=1}^{4} J(c) \cdot J(x_{k+n_i})\Big).$$

This PCA has some deterministic components: if $x_k = (0,0,0,0)$ (resp. $(1,1,1,1)$), then with probability 1, cell k will be in the same state after the interaction step. The *alignment sensitivity* α is a parameter controling the intensity of the dependence on the neighbours.

The composition of the two steps (interaction and propagation) can also be defined as a strict PCA, by enlarging the neighbourhood and the alphabet. But we prefer to keep the description of the model as the composition of these two simple steps. Experimental studies (on finite lattices with periodic boundary conditions) show that depending on the initial density of particles and the value of α, either disordered configurations are observed, or particles find a consensus to move together in one or more directions. We do not enter into the details, but present in Fig. 2.7 the diagram in which Bouré, Chevrier and Fatès sum up the different phases observed [BFC11].

For $\alpha = 0$, the interaction step gives an equal probability to each state $c \in \mathcal{A}$ such that $x_k = c$, regardless of the configuration of the neighbourhood of cell k, so that the behaviour is disordered. In terms of invariant measure, the Bernoulli product measure on $\mathbb{Z}^2$ that consists in occupying each elementary triangle independly with probability p, so that the probability of a state c is given by: $\prod_{i=1}^{4} p_i^{c_i}(1-p)^{1-c_i}$ provides an invariant measure of the model. Indeed, this Bernoulli product measure is clearly invariant by the interaction step, and it is also invariant for the propagation step since it defines four independent Bernoulli measures on E, W, N, and S parts of the cells, and each of them is shift-invariant.

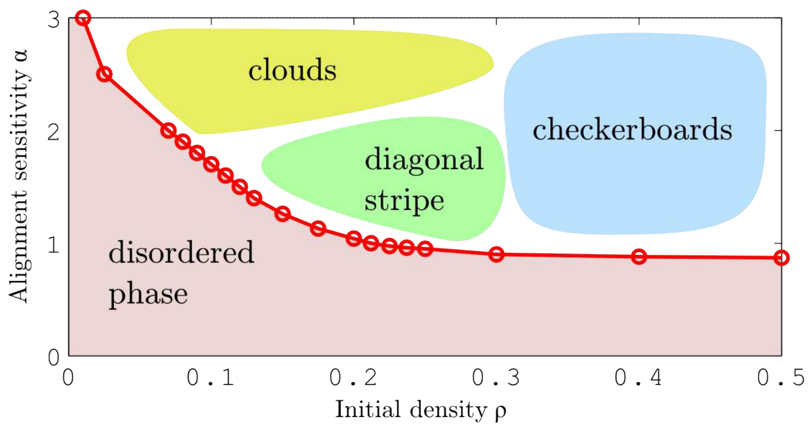


Figure 2.7: The different phases observed for the swarming model [BFC11].

Let us now also consider the case $\alpha \neq 0$. For the time being, let us consider only the PCA defining the interaction step, that is interesting in itself. We work on the finite lattice $E = (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$. For a given measure ν on $\{0,1,2,3,4\}^E$, we define a measure π on $\mathcal{A}^E$, by:

$$\pi(x) = \nu(|x|) \prod_{k \in E} Z(x_{k+\mathcal{N}_0}),$$

where we write $|x|$ for $(|x_k|)_{k \in E}$. This measure is an invariant measure with respect to the interaction step. Let us denote by P the Markov chain on $\mathcal{A}^E$ corresponding to the interaction step. The measure π is a reversible measure for P, since for $x, y \in \mathcal{A}^E$, we have $P(x,y) = 0$ if $|x| \neq |y|$, and if $|x| = |y|$, then:

$$\begin{aligned}
\pi(x)P(x,y) &= \nu(|x|) \prod_{k \in E} Z(x_{k+\mathcal{N}_0}) \prod_{k \in E} \frac{1}{Z(x_{k+\mathcal{N}_0})} \exp\Big(\alpha \sum_{i=1}^{4} J(y_k) \cdot J(x_{k+n_i})\Big) \\
&= \nu(|x|) \prod_{k \in E} \exp\Big(\alpha \sum_{i=1}^{4} J(y_k) \cdot J(x_{k+n_i})\Big) \\
&= \nu(|x|) \ \exp\Big(\alpha \sum_{k \in E} \sum_{i=1}^{4} J(y_k) \cdot J(x_{k+n_i})\Big) \\
&= \nu(|y|) \ \exp\Big(\alpha \sum_{k \in E} \sum_{i=1}^{4} J(x_k) \cdot J(y_{k+n_i})\Big) \\
&= \pi(y)P(y,x).
\end{aligned}$$

In order to understand the diagram of Fig. 2.7, a direction could be to study the values of the parameters for which there exists a probability measure ν inducing a measure π not only invariant with respect to the interaction step, but also with respect to the propagation step.

Chapter 3

Ergodicity and perfect sampling

Assis à sa caisse, il regardait la grande horloge fixée au-dessus du magasin de meussieu Poucier, et il suivait la marche de la grande aiguille. Il réussissait à la voir sauter une fois, deux fois, trois fois, puis tout à coup il se retrouvait un quart d'heure plus tard et la grosse aiguille elle-même en avait profité pour bouger sans qu'il s'en aperçût. Où était-il allé pendant ce temps là ?

– Raymond Queneau, *Le dimanche de la vie*

Contents

The equilibrium of a PCA is studied via its invariant measures. A PCA is *ergodic* if it has a unique and attractive invariant measure. Finding conditions to ensure ergodicity is a difficult problem which has been thoroughly investigated [DKT90, Gác01]. When a PCA is ergodic, it is usually impossible to determine the invariant measure explicitly, and simulation becomes the alternative. Simulating PCA is known to be a challenging task, costly both in time and space. Also, configurations cannot be tracked down one by one (there is an infinite number of them when the set of cells is infinite) and may only be observed through some measured parameters. The point is to have guarantees upon the results obtained from simulations.

In this context, our contributions are as follows. First, we prove that the ergodicity of a CA on $\mathbb{Z}$ is undecidable. This was mentioned an unsolved problem [Too01]. Since a CA is a special case of a PCA, it also provides a new proof of the undecidability of the ergodicity of a PCA [DKT90, Too00]. Second, we propose an efficient perfect sampling algorithm for ergodic PCA. Recall that a *perfect sampling* procedure is a random algorithm which returns a configuration distributed according to the invariant measure. By applying the procedure repeatedly, we can estimate the invariant measure with arbitrary precision. We propose such an algorithm for PCA by adapting the *coupling from the past* method of Propp and Wilson [PW96]. When the set of cells is finite, a PCA is a finite state space Markov chain. Therefore, coupling from the past from all possible initial configurations provides a basic perfect sampling procedure, but a very inefficient one since the number of configurations is exponential in the number of cells. Here, the contribution consists in an important simplification of the procedure. We define a new PCA on an extended alphabet, called the *envelope PCA* (EPCA). We obtain a perfect sampling procedure for the original PCA by running the EPCA on a single initial configuration. When the set of cells is infinite, a PCA is a Markov chain on an uncountable state space. So there is no basic perfect sampling procedure anymore. We prove the following: If the PCA is ergodic, then the EPCA may or may not be ergodic. If it is ergodic, then we can use the EPCA to design an efficient perfect sampling procedure (the result of the algorithm is the finite restriction of a configuration with the right invariant distribution). The EPCA can be viewed as a systematic treatment of ideas already used by Toom for *percolation PCA* [Too01, Sec. 2].

The perfect sampling procedure can also be run on a PCA whose ergodicity is unknown, with the purpose of testing it. We illustrate this approach on the *majority-flip PCA*, prototype of a PCA whose equilibrium behaviour is not well understood.

3.1 Ergodicity of PCA

3.1.1 Invariant measures and ergodicity

A PCA has at least one invariant measure, and the set of invariant measures is convex and compact. This is a standard fact, based on the observation that the set $\mathcal{M}(\mathcal{X})$ of measures on $\mathcal{X}$ is compact for the weak topology [DKT90]. Therefore, there are three possible situations for a PCA:

(*i*) several invariant measures;

(*ii*) a unique invariant measure which is not attractive;

(*iii*) a unique invariant measure which is attractive (ergodic case).

Example 3.1. Let $\mathcal{A} = \{0,1\}$, $E = \mathbb{Z}^d$, and let $\mathcal{N}$ be a finite subset of $\mathbb{Z}^d$. Consider $0 < \gamma < 1$ and the local function:

$$f((x_v)_{v\in V}) = \gamma\,\delta_{\max(x_v,\ v\in V)} + (1-\gamma)\,\delta_0.$$

The corresponding PCA is called the *percolation PCA* associated with $\mathcal{N}$ and γ. The particular case of the space $E = \mathbb{Z}$ and the neighbourhood $\mathcal{N} = \{0,1\}$ is called the *Stavskaya PCA*. Observe that the Dirac measure $\delta_{0^{\mathbb{Z}^d}}$ is an invariant measure. Using a coupling with a site percolation model, one can prove the following [Too01, Sec. 2]. There exists $\gamma^* \in (0,1)$ such that:

$$\begin{aligned} \gamma < \gamma^* \quad &\Longrightarrow \quad (iii):\ \text{ergodicity} \\ \gamma > \gamma^* \quad &\Longrightarrow \quad (i):\ \text{several invariant measures.} \end{aligned}$$

The exact value of γ^* is not known but it satisfies $(1/\text{Card}\,\mathcal{N}) \le \gamma^* \le 53/54$. In Fig. 3.1, we represent three space-time diagrams of the percolation operator of neighbourhood $\mathcal{N} = \{-1,0,1\}$, from the configuration with only 1's (full squares), for respectively $\gamma = 0.45$, $\gamma = 0.50$ and $\gamma = 0.55$.

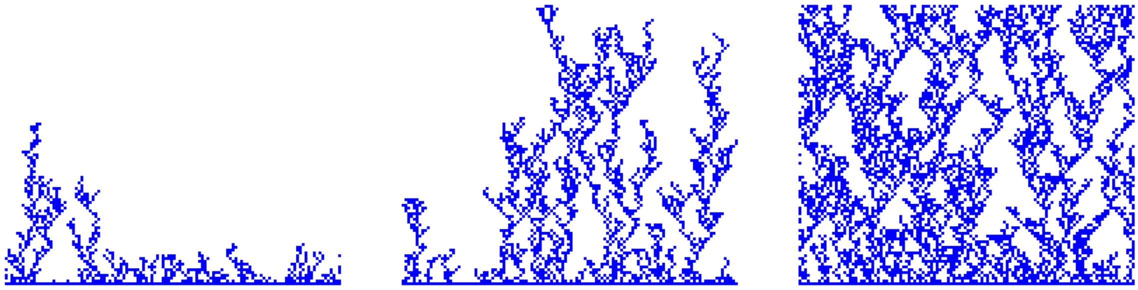


Figure 3.1: Space-time diagrams of the percolation operator of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$, for $\gamma = 0.45$, $\gamma = 0.50$ and $\gamma = 0.55$.

The existence of a PCA corresponding to situation (*ii*) was mentioned as an open problem by Toom [Too01]. It was proved by Chassaing and Mairesse [CM11] that situation (*ii*) occurs for the PCA on $\{0,1\}^{\mathbb{Z}}$ of neighbourhood $\mathcal{N} = \{-1, 0\}$, and local function f defined by

$$f(0,0)(1) = 1/2, \quad f(0,1)(1) = 0, \quad f(1,0)(1) = 1, \quad f(1,1)(1) = 1/2.$$

The unique invariant measure of that PCA is the mixture of two measures, concentrated on $(01)^{\mathbb{Z}}$ and $(10)^{\mathbb{Z}}$ respectively, that are not shift-invariant, and cycle between each other. It would be interesting to find other examples. In particular, it is still unknown if there are also positive-rate PCA for which situation (*ii*) occurs.

The PCA of Ex. 3.1 exhibits a phase transition between the situations (*i*) and (*iii*). In Sec. 3.3, we study a PCA that may have a phase transition between the situations (*ii*) and (*iii*). It would provide the first example of this type.

3.1.2 Undecidability of the ergodicity

Deterministic cellular automata (CA) form the simplest class of PCA, it is therefore natural to study their ergodicity. In this section, we prove the undecidability of ergodicity for CA (Th. 3.1). This also gives a new proof of the undecidability of the ergodicity for PCA.

Remark. In the context of CA, the terminology of Def. 1.5 might be confusing. Indeed a CA F can be viewed in two different ways:

(*i*) a (degenerated) Markov chain; (*ii*) a symbolic dynamical system.

In the dynamical system terminology, F is *uniquely ergodic* if:

$$\exists! \mu, \ \mu F = \mu.$$

In the Markov chain terminology (that we adopt), F is *ergodic* if:

$$\exists! \mu, \ \mu F = \mu \text{ and } \forall \nu, \ \nu F^n \xrightarrow{w} \mu,$$

where $\xrightarrow{w}$ stands for the weak convergence. Knowing if the unique ergodicity (of symbolic dynamics) implies the ergodicity (of the Markovian theory) is an open question for CA.

Let F be a CA on $\mathcal{X} = \mathcal{A}^E$, with $E = \mathbb{Z}^d$ for some $d \geq 1$. The *limit set* of F is defined by

$$LS(F) = \bigcap_{n \in \mathbb{N}} F^n(\mathcal{X})\,.$$

In words, a configuration belongs to LS if it may occur after an arbitrarily long evolution of the cellular automaton.

Observe that $LS(F)$ is non-empty since it is the decreasing limit of non-empty closed sets. A constructive way to show that $LS(F)$ is non-empty is as follows. Let us recall that a *monochromatic configuration* is a configuration of the type α^E, for some letter $\alpha \in \mathcal{A}$. The image by F of a monochromatic configuration is monochromatic. In particular, since $\mathcal{A}$ is a finite set, there exists a monochromatic periodic orbit for F:

$$\alpha_0^E \to \alpha_1^E \to \cdots \to \alpha_{k-1}^E \to \alpha_0^E .$$

This implies that $\{\alpha_0^E, \alpha_1^E, \dots, \alpha_{k-1}^E\} \subset LS(F)$.

Recall that δ_x denotes the probability measure concentrated on the configuration x. The periodic orbit $(\alpha_0^E, \dots, \alpha_{k-1}^E)$ provides an invariant measure given by $(\delta_{\alpha_0^E} + \dots + \delta_{\alpha_{k-1}^E})/k$. More generally, the support of any invariant measure is included in the limit set.

Definition 3.1 (Nilpotent CA)**.** A CA is *nilpotent* if its limit set is a singleton.

Using the above observation on monochromatic periodic orbits, we see that a CA F is nilpotent if and only if $LS(F) = \{\alpha^E\}$ for some letter $\alpha \in \mathcal{A}$. The following stronger statement [CPY89] is proved using a compactness argument:

$$[\ F \ \textit{nilpotent}\ ] \quad \Longleftrightarrow \quad [\ \exists \alpha \in \mathcal{A},\ \exists N \in \mathbb{N},\ F^N(\mathcal{A}^E) = \{\alpha^E\}\ ] .$$

We obtain the next proposition as a corollary.

Proposition 3.1. *Consider a CA F. We have:*

$$[\ F \ \textit{nilpotent}\ ] \quad \Longrightarrow \quad [\ F \ \textit{ergodic}\ ] .$$

Proof. Let $\alpha \in \mathcal{A}$ and $N \in \mathbb{N}$ be such that $F^N(\mathcal{A}^E) = \{\alpha^E\}$. For any probability measure μ on $\mathcal{A}^E$, we have $\mu F^N = \delta_{\alpha^E}$. Therefore, F is ergodic with unique invariant measure δ_{α^E}. □

We also have the converse statement.

Theorem 3.1. *Consider a CA F on the set of cells $\mathbb{Z}^d$. We have:*

$$[\ F \ \textit{nilpotent}\ ] \quad \Longleftrightarrow \quad [\ F \ \textit{ergodic}\ ] .$$

Proof. Let F be an ergodic CA. Assume that there exists a monochromatic periodic orbit $(\alpha_0^E, \dots, \alpha_{k-1}^E)$ with $k \geq 2$. Then $\mu = (\delta_{\alpha_0^E} + \cdots + \delta_{\alpha_{k-1}^E})/k$ is the unique invariant measure. The sequence $\delta_{\alpha_0^{\mathbb{Z}}} F^n$ does not converge weakly to μ, which is a contradiction. Therefore, there exists a monochromatic fixed point: $F(\alpha^E) = \alpha^E$, and δ_{α^E} is the unique invariant measure.

Define the cylinder $[\alpha_K] = \{x \in \mathcal{A}^E; \forall i \in K,\ x_i = \alpha\}$, where K is some finite subset of E. For any initial configuration $x \in \mathcal{A}^E$, using the ergodicity of P, we have:

$$\delta_x F^n([\alpha_K]) \underset{n \to +\infty}{\longrightarrow} \delta_{\alpha^E}([\alpha_K]) = 1 .$$

But $\delta_x F^n$ is a Dirac measure, so $\delta_x F^n([\alpha_K])$ is equal to 0 or 1. Consequently, we have $\delta_x F^n([\alpha_K]) = 1$ for n large enough, that is,

$$\exists N \in \mathbb{N}, \forall n \geq N, \forall i \in K, \quad F^n(x)_i = \alpha .$$

In words, in any space-time diagram of F, any finite column of base K becomes eventually equal to α^K. Using the terminology of Guillon and Richard, the CA F has a *weakly nilpotent trace*. These two authors have proved that for one-dimensional CA, the weak nilpotency of the trace implies the nilpotency of the CA [GR08]. A recent result [Sal12] proves that it is still true in larger dimensions.

This completes the proof. □

Kari proved that the nilpotency of a CA on $\mathbb{Z}$ is undecidable [Kar92]. (For CA on $\mathbb{Z}^d$, $d \geq 2$, the proof was published a few years before [CPY89].) By coupling Kari's result with Th. 3.1, we obtain the following result.

Corollary 3.1. *Consider a CA F on the set of cells $\mathbb{Z}$. The ergodicity of F is undecidable.*

The undecidability of the ergodicity of a PCA was a known result, proved by Kurdyumov [DKT90] and Toom [Too00]. Kurdyumov's and Toom's proofs use a non-deterministic PCA of dimension 1 and a reduction of the halting problem of a Turing machine.

Corollary 3.1 is a stronger statement. In fact, the (un)decidability of the ergodicity of a CA was mentioned by Toom as an unsolved problem [Too01]. We point out that Corollary 3.1 can also be obtained without Th. 3.1, by directly adapting Kari's proof to show the undecidability of the ergodicity of the CA associated with a North-West deterministic tile set.

3.2 Sampling the invariant measure of an ergodic PCA

Generally, the invariant measure(s) of a PCA cannot be described explicitly. Numerical simulations are consequently very useful to get an idea of the behaviour of a PCA. Given an ergodic PCA, we propose a *perfect sampling* algorithm which generates configurations *exactly* according to the invariant measure.

A perfect sampling procedure for finite Markov chains has been proposed by Propp and Wilson [PW96] using a *coupling from the past* scheme. Perfect sampling procedures have been developed since in various contexts. We mention below some works directly linked to the present article. For more information see the annotated bibliography: *Perfectly Random Sampling with Markov Chains*, `http://dimacs.rutgers.edu/~dbwilson/exact.html/`.

The complexity of the algorithm depends on the number of all possible initial conditions, which is prohibitive for PCA. Various techniques have been developed to reduce the number of trajectories that need to be considered in the coupling from the past procedure. A first crucial observation already appears in the work of Propp and Wilson [PW96]: for a monotone Markov chain, one has to consider two trajectories corresponding to minimal and maximal states of the system. For anti-monotone systems, an analogous technique has been developed by Häggström and Nelander [HN98] that also considers only extremal initial conditions. To cope with more general situations, Huber [Hub04] introduced the idea of a bounding chain for determining when coupling has occurred. The construction of these bounding chains is model-dependent and in general not straightforward.

Our contribution is to show that the bounding chain ideas can be given in a particularly simple and convenient form in the context of PCA via the introduction of the *envelope PCA*.

3.2.1 Basic coupling from the past for PCA

Finite set of cells

Consider an ergodic PCA F on the alphabet $\mathcal{A}$ and on a finite set of cells E, for example $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$. Let π be the invariant measure on $\mathcal{X} = \mathcal{A}^E$. A *perfect sampling* procedure is a random algorithm which returns a configuration $x \in \mathcal{X}$ with probability $\pi(x)$. Let us present

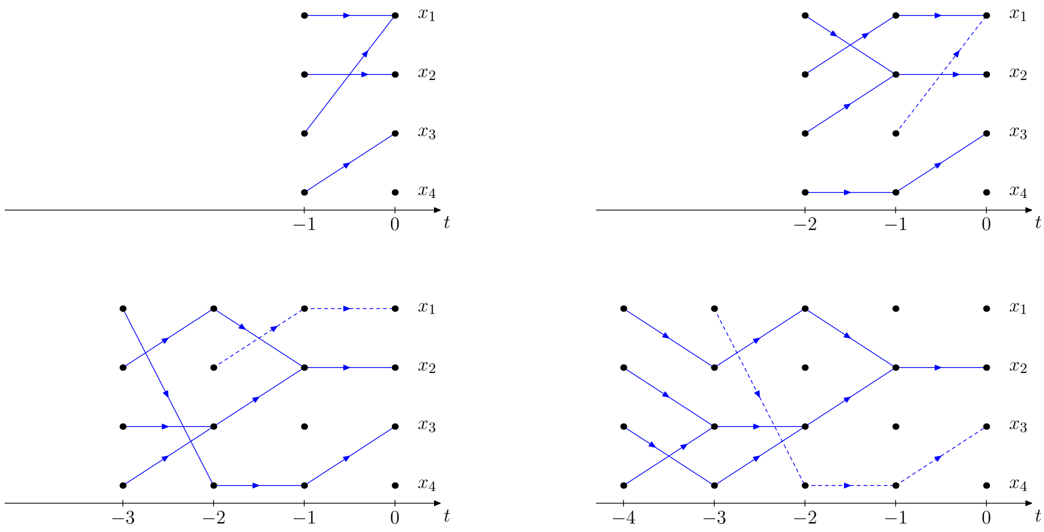


Figure 3.2: Coupling from the past.

coupling from the past (CFTP), perfect sampling procedure.

Algorithm 1: Basic CFTP algorithm for a finite set of cells

Data: An update function $\phi : \mathcal{X} \times [0,1]^E \to \mathcal{X}$ of a PCA. A family $(r_k^{-n})_{(k,n)\in E\times\mathbb{N}}$ of i.i.d. r.v. with uniform distribution in $[0,1]$.
Result: a state of $\mathcal{A}^E$ distributed according to the invariant distribution of the PCA.
begin
  $t = 1$;
  repeat
    $R_{-t} = \mathcal{X}$;
    for $j = -t$ **to** -1 **do**
      $R_{j+1} = \{\phi(x, (r_i^j)_{i\in E}) \; ; \; x \in R_j\}$
    $t = t + 1$
  until $|R_0| = 1$;
  return *the unique element of* R_0
end

The good way to implement this algorithm is to keep track of the partial couplings of trajectories. This allows to consider only one-step transitions.

Proposition 3.2 ([PW96]). *If the procedure stops almost surely, then the PCA is ergodic and the output is distributed according to the invariant measure.*

The converse statement is not true in general: even for ergodic PCA, there exist choices of ϕ for which the procedure does not stop. Nevertheless, for PCA having positive rates (see Def. 1.6), the algorithm stops almost surely in finite time if the update function is chosen according to (1.2).

In Fig. 3.2, we illustrate the algorithm on the toy example of a PCA on the alphabet $\{0,1\}$ and the set of cells $\mathbb{Z}_2$. The state space is thus $\mathcal{X} = \{x_1 = 00, x_2 = 01, x_3 = 10, x_4 = 11\}$. On this sample, the algorithm returns x_2.

A sketch of the proof of Prop. 3.2 can be given using Fig. 3.2. On the last of the four pictures, the Markov chain is run from time -4 onwards and its value is x_2 at time 0. If we had run the Markov chain from time $-\infty$ to 0, then the result would obviously still be x_2.

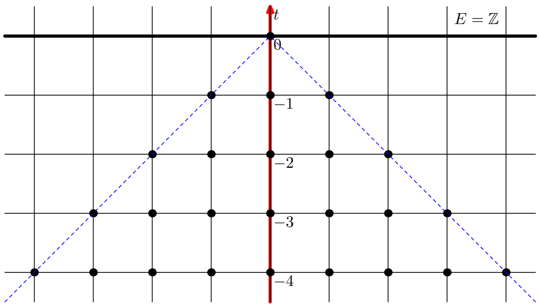


Figure 3.3: Dependence cone of a cell.

But if we started from time $-\infty$, then the Markov chain would have reached equilibrium by time 0.

Infinite set of cells

Assume that the set of cells E is infinite $(E = \mathbb{Z}^d, d \geq 1)$. Then a PCA defines a Markov chain on the infinite state space $\mathcal{X} = \mathcal{A}^E$, so the above procedure is not effective anymore. However, it is possible to use the locality of the updating rule of a PCA to still define a perfect sampling procedure. (This observation was already mentionned by van den Berg and Steif [vdBS99].)

Let F be an ergodic PCA and denote by π its invariant distribution. In this context, a *perfect sampling* procedure is a random algorithm taking as input a finite subset K of E and returning an element $x_K \in \mathcal{A}^K$ with probability $\pi([x_K])$.

To get such a procedure, we use the following fact: if the PCA is run from time $-k$ onwards, then to compute the content of the cells in K at time 0, it is enough to consider the cells in the finite dependence cone of K. This is illustrated in Figure 3.3 for the set of cells $E = \mathbb{Z}$ and the neighbourhood $\mathcal{N} = \{-1, 0, 1\}$, with the choice $K = \{0\}$. Observe that the orientation has changed with respect to Fig. 3.2 in order to be consistent with the convention used for space-time diagrams.

Let us define this formally. Let $\mathcal{N}$ be the neighbourhood of the PCA. Given a subset K of E, the *backward dependence cone* of K corresponds to the family $(V_{-t}(K))_{t\geq 0}$ of subsets of E defined recursively by $V_0(K) = K$ and $V_{-t}(K) = \mathcal{N} + V_{-t+1}(K)$. Let $\phi : \mathcal{X} \times [0, 1]^E \to \mathcal{X}$ be an update function, for instance the one defined according to (1.2). For a given subset K of E, we denote $\phi_{-t} : \mathcal{A}^{V_{-t}(K)} \times [0, 1]^{V_{-t+1}(K)} \to \mathcal{A}^{V_{-t+1}(K)}$ the corresponding restriction of ϕ.

With these notations, the algorithm can be written as follows.

Algorithm 2: Basic CFTP algorithm for an infinite set of cells

Data: An update function $\phi : \mathcal{X} \times [0,1]^E \to \mathcal{X}$ of a PCA. A family $(r_k^{-n})_{(k,n)\in E\times\mathbb{N}}$ of i.i.d. r.v. with uniform distribution in $[0,1]$. A finite subset K of E.
Result: a state of $\mathcal{A}^K$ distributed according to the invariant distribution of the PCA.
begin
  $V_0(K) = K$;
  $t = 1$;
  repeat
    $V_{-t}(K) = \mathcal{N} + V_{-t+1}(K)$;
    $R_{-t} = \mathcal{A}^{V_{-t}(K)}$;
    for $j = -t$ **to** -1 **do**
      $R_{j+1} = \{\phi_j(x, (r_i^j)_{i\in V_{j+1}(K)}) \ ; \ x \in R_j\} \subset \mathcal{A}^{V_{j+1}(K)}$
    $t = t + 1$
  until $|R_0| = 1$;
  return *the unique element of* R_0
end

Next proposition is an easy extension of Prop. 3.2.

Proposition 3.3. *If the procedure stops almost surely, then the PCA is ergodic and the output is distributed according to the marginal of the invariant measure.*

The converse statement is not true in general. It would be interesting to know if it holds true for the update function (1.2) and for PCA having positive rates (possibly under additional hypothesis).

3.2.2 Envelope probabilistic cellular automata (EPCA)

The CFTP algorithm is inefficient when the state space is large. This is the case for PCA: when E is finite, the set $\mathcal{A}^E$ is very large, and when E is infinite, it is the number of configurations living in the dependence cone described above which is very large. We cope with this difficulty by introducing the *envelope* PCA.

To begin with, let us assume that F is a PCA on the alphabet $\mathcal{A} = \{0, 1\}$ (as previously, the set of cells is denoted by E, the neighbourhood by $\mathcal{N} \subset E$, and the local function by f). The case of a general alphabet is treated in Sec. 3.2.6.

Definition of the EPCA

Let us introduce a new alphabet:
$$\mathcal{B} = \{\mathbf{0}, \mathbf{1}, ?\}.$$
A word on $\mathcal{B}$ is to be thought as a word on $\mathcal{A}$ in which the letters corresponding to some positions are not known, and are thus replaced by the symbol "?". Formally we identify $\mathcal{B}$ with $2^{\mathcal{A}} - \emptyset$ as follows: $\mathbf{0} = \{0\}$, $\mathbf{1} = \{1\}$, and $? = \{0, 1\}$. So each letter of $\mathcal{B}$ is a set of possible letters of $\mathcal{A}$. With this interpretation, we view a word on $\mathcal{B}$ as a set of words on $\mathcal{A}$. For instance,
$$?\mathbf{1}? = \{010, 011, 110, 111\}.$$

We will associate to the PCA F a new PCA on the alphabet $\mathcal{B}$, that we call the *envelope probabilistic cellular automaton* of F.

Definition 3.2 (Envelope PCA)**.** The *envelope probabilistic cellular automaton (EPCA)* of F, is the PCA $\mathrm{env}(F)$ of alphabet $\mathcal{B}$, defined on the set of cells E, with the same neighbourhood $\mathcal{N}$ as for F, and a local function $\mathrm{env}(f) : \mathcal{B}^{\mathcal{N}} \to \mathcal{M}(\mathcal{B})$ defined for each $y \in \mathcal{B}^{\mathcal{N}}$ by

$$\mathrm{env}(f)(y)(\mathbf{0}) = \min_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(0)$$

$$\mathrm{env}(f)(y)(\mathbf{1}) = \min_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(1)$$

$$\mathrm{env}(f)(y)(?) = 1 - \min_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(0) - \min_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(1).$$

We point out that $\min_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(1) + \max_{x \in \mathcal{A}^{\mathcal{N}},\, x \in y} f(x)(0) = 1$, so that the last quantity $\mathrm{env}(f)(y)(?)$ is non-negative.

Moreover, $\mathrm{env}(F)$ acts like F on configurations which do not contain the letter "?". More precisely,

$$\forall y \in \mathcal{A}^{\mathcal{N}}, \qquad \mathrm{env}(f)(y)(\mathbf{0}) = f(y)(0), \quad \mathrm{env}(f)(y)(\mathbf{1}) = f(y)(1), \quad \mathrm{env}(f)(y)(?) = 0\,. \tag{3.1}$$

In particular, we get the following.

Proposition 3.4. *If the EPCA* $\mathrm{env}(F)$ *is ergodic then the PCA* F *is ergodic.*

Proof. According to (3.1), any invariant measure of F corresponds to an invariant measure of $\mathrm{env}(F)$. Therefore, if F has several invariant measures, so does $\mathrm{env}(F)$. Assume that F has a unique invariant measure μ which is non-ergodic. Let μ_0 be such that $\mu_0 F^n$ does not converge to μ. Then $\mu_0\ \mathrm{env}(F)^n$ does not converge either, see (3.1). To summarise, we have proved that F non-ergodic implies $\mathrm{env}(F)$ non-ergodic. □

The converse of Prop. 3.4 is not true and counter-examples will be given in Sec. 3.2.4.

Construction of an update function for the EPCA.

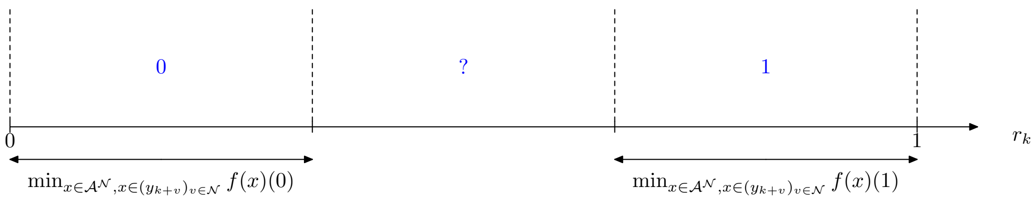
Let us define the update function

$$\tilde{\phi} : \mathcal{B}^E \times [0,1]^E \to \mathcal{B}^E$$

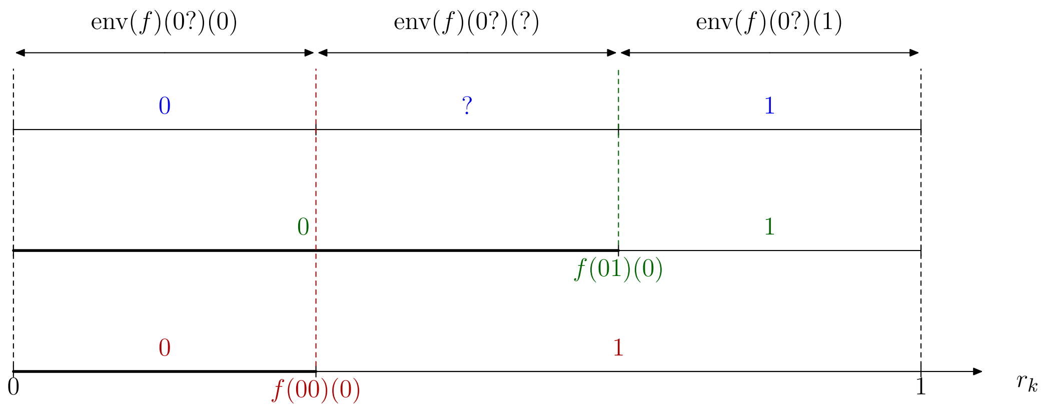
of the PCA $\mathrm{env}(F)$, by:

$$\tilde{\phi}(y,r)_k = \begin{cases} \mathbf{0} \text{ if } 0 \leq r_k < \mathrm{env}(f)((y_{k+v})_{v \in \mathcal{N}})(\mathbf{0}) \\ \mathbf{1} \text{ if } 1 - \mathrm{env}(f)((y_{k+v})_{v \in \mathcal{N}})(\mathbf{1}) \leq r_k \leq 1 \\ ? \text{ otherwise.} \end{cases} \tag{3.2}$$

The value of $\tilde{\phi}(y,r)_k$ as a function of r_k can thus be represented as follows.



For a PCA of neighbourhood $\mathcal{N} = \{0,1\}$, we represent below the construction of the updates of the EPCA when the value of the neighbourhood is $\mathbf{0}?$.



Let ϕ be the natural update function for the PCA P defined as in (1.2). Observe that the function $\tilde{\phi}$ coincides with ϕ on configurations which do not contain the letter "?". Furthermore, we have:

$$\forall r \in [0,1]^E,\ \forall x \in \mathcal{A}^E,\ \forall y \in \mathcal{B}^E, \qquad x \in y \implies \phi(x,r) \in \tilde{\phi}(y,r)\,. \tag{3.3}$$

3.2.3 Perfect sampling using EPCA

We propose two perfect sampling algorithms, for a finite and for an infinite number of cells. We show that in both cases, the algorithm stops almost surely if and only if the EPCA is ergodic (Th. 3.2). The ergodicity of the EPCA implies the ergodicity of the PCA but the converse is not true: we provide a counterexample for each case, finite and infinite (Sec. 3.2.4).We also give conditions of ergodicity of the EPCA (Prop. 3.5 and 3.6).

Finite set of cells

The idea is to consider only one trajectory of the EPCA - the one that starts from the initial configuration $?^E$ (coding the set of all configurations of the PCA). The algorithm stops when at time 0, this trajectory hits the set $\mathcal{A}^E$.

Algorithm 3: Perfect sampling using the EPCA for a finite set of cells

Data: The pre-computed update function $\tilde{\phi}$. A family $(r_k^{-n})_{(k,n)\in E\times\mathbb{N}}$ of i.i.d. r.v. with uniform distribution in $[0,1]$.
Result: a state of $\mathcal{A}^E$ distributed according to the invariant distribution of the PCA.
begin
  $t = 1$;
  repeat
    $c = ?^E$;
    for $j = -t$ **to** -1 **do**
      $c = \tilde{\phi}(c, (r_i^j)_{i\in E})$
    $t = t + 1$
  until $c \in \mathcal{A}^E$;
  return c
end

Infinite set of cells

As already mentioned in Sec. 3.2.1, when the set of cells E is infinite, one is no more interested in generating a complete configuration of $\{0,1\}^E$ according to the invariant measure π of F, but rather in simulating finite-dimensional marginals of π. Once again, we consider only one trajectory of the EPCA. Let K be a finite set of cells from E. We propose the following algorithm to simulate the marginals of π corresponding to these cells.

Algorithm 4: Perfect sampling using the EPCA for an infinite set of cells

Data: The pre-computed update function $\tilde{\phi}$. A family $(r_k^{-n})_{(k,n)\in E\times\mathbb{N}}$ of i.i.d. r.v. with uniform distribution in $[0,1]$. A finite subset K of E.
Result: a state of $\mathcal{A}^K$ distributed according to the invariant distribution of the PCA.

```
begin
    V_0(K) = K ;
    t = 1 ;
    repeat
        V_{-t}(K) = 𝒩 + V_{-t+1}(K) ;
        c = ?^{V_{-t}(K)} ;
        for j = -t to -1 do
            c = φ̃_j(c, (r_i^j)_{i∈V_{j+1}(K)}) ∈ ℬ^{V_{j+1}(K)}
        t = t + 1
    until c ∈ 𝒜^K ;
    return c
end
```

Theorem 3.2. *Algorithm 3, resp. 4, stops almost surely if and only if the EPCA is ergodic. In that case, the output of the algorithm is distributed according to the unique invariant measure of the PCA.*

Proof. The argument is the same in the finite and infinite cases. We give it for the finite case. Assume first that Algorithm 3 stops almost surely. By construction, it implies that for all μ_0, the measure $\mu_0\ \mathrm{env}(F)^n$ is asymptotically supported by $\mathcal{A}^E$. Therefore, we can strengthen the result in Prop. 3.4: the invariant measures of $\mathrm{env}(F)$ coincide with the invariant measures of F. In that case, $\mathrm{env}(F)$ is ergodic iff F is ergodic. Using (3.3), the halting of Algorithm 3 implies the halting of Algorithm 1. Furthermore, if we use the same samples $(r_k^{-n})_{(k,n)\in E\times\mathbb{N}}$, Algorithms 3 and 1 will have the same output. According to Prop. 3.2, this output is distributed according to the unique invariant measure of P. In particular, F is ergodic. So $\mathrm{env}(F)$ is ergodic.

Assume now that the EPCA is ergodic. The unique invariant measure π of $\mathrm{env}(F)$ has to be supported by $\mathcal{A}^E$. Also, by ergodicity, we have $\delta_{?^E}\ \mathrm{env}(F)^n \xrightarrow{w} \pi$. This means precisely that the Algorithm 3 stops a.s. □

3.2.4 Criteria of ergodicity for the EPCA

Finite set of cells

In the next proposition, we give a necessary and sufficient condition for the EPCA to be ergodic. In particular, this condition is satisfied if the PCA has positive rates (see Def. 1.6).

Proposition 3.5. *The EPCA* $\mathrm{env}(F)$ *is ergodic if and only if* $\mathrm{env}(f)(?^{\mathcal{N}})(?) < 1$. *This condition can also be written as:*

$$\min_{x\in\mathcal{A}^{\mathcal{N}}} f(x)(0) + \min_{x\in\mathcal{A}^{\mathcal{N}}} f(x)(1) > 0. \tag{3.4}$$

Proof. If $\mathrm{env}(f)(?^{\mathcal{N}})(?) = 1$, then for almost any $r \in [0,1]^E$, we have $\tilde{\phi}(?^E, r) = ?^E$, so that at each step of the algorithm, the value of c is $?^E$ with probability 1.

Conversely, if we assume for example that $p = \min_{x\in\mathcal{A}^{\mathcal{N}}} f(x)(0) > 0$, then for any configuration $d \in \mathcal{B}^E$, the probability to have $\tilde{\phi}(x,r) = \mathbf{0}^E$ is greater than $p^{|E|}$, so that the algorithm stops almost surely, and the expectation of the running time can be roughly bounded by $1/p^{|E|}$. □

Infinite set of cells

For an infinite set of cells the situation is more complex. The condition of Prop. 3.5 is not sufficient to ensure the ergodicity of the EPCA. A counter-example is given in Sec. 3.2.4. First, we propose a rough sufficient condition of convergence for Algorithm 4.

Proposition 3.6. *Let $\gamma^* \in (0,1)$ be the critical probability of the percolation PCA with neighbourhood $\mathcal{N}$, see Ex. 3.1 and 3.1. The EPCA* env(F) *is ergodic if*

$$\mathrm{env}(f)(?^{\mathcal{N}})(?) \;<\; \gamma^* \tag{3.5}$$

and non-ergodic if

$$\min_{x\in\mathcal{B}^{\mathcal{N}}-\mathcal{A}^{\mathcal{N}}} \mathrm{env}(f)(x)(?) \;>\; \gamma^*. \tag{3.6}$$

Proof. Recall that $\mathcal{B} = \{\mathbf{0},\mathbf{1},?\}$. Define $\mathcal{C} = \{\mathbf{d},?\}$, with $\mathbf{d} = \{\mathbf{0},\mathbf{1}\}$. A word over $\mathcal{C}$ is interpreted as a set of words over $\mathcal{B}$, for instance, $\mathbf{d}? = \{\mathbf{0}?,\mathbf{1}?\}$. The symbol $\mathbf{d}$ stands for **d**etermined letter, as opposed to ? which represents an unknown letter.

We define a new PCA G on the alphabet $\mathcal{C}$, with the same neighbourhood $\mathcal{N}$ as F and env(F), and with the transition function $g : \mathcal{C}^{\mathcal{N}} \to \mathcal{M}(\mathcal{C})$ defined by:

$$g(\mathbf{d}^{\mathcal{N}}) = \delta_{\mathbf{d}}, \quad \text{and} \quad \forall u \in \mathcal{C}^{\mathcal{N}} - \{\mathbf{d}^{\mathcal{N}}\}, \quad g(u) = \alpha\,\delta_? + (1-\alpha)\,\delta_{\mathbf{d}}\;,$$

for $\alpha = \max_{x\in\mathcal{B}^{\mathcal{N}}} \mathrm{env}(f)(x)(?) = \mathrm{env}(f)(?^{\mathcal{N}})(?)$.

Observe that $\delta_{\mathbf{d}^E}$ is an invariant measure of G. Recall that $\tilde{\phi}$ is an update function of env(F), see (3.2). Given the way G is defined, we can construct an update function ϕ_G of G such that

$$\forall x \in \mathcal{B}^E, \forall y \in \mathcal{C}^E, \forall r \in [0,1]^E, \qquad x \in y \implies \tilde{\phi}(x,r) \in \phi_G(y,r)\;. \tag{3.7}$$

In particular, assume that G is ergodic. Then $\delta_{?^E}\, G^n \xrightarrow{w} \delta_{\mathbf{d}^E}$. Using (3.7), it implies that Algorithm 4 stops almost surely, and env(F) is ergodic according to Th. 3.2. To summarise, the ergodicity of G implies the ergodicity of env(F).

Observe that the PCA G is a percolation PCA as defined in Ex. 3.1 (here, $\mathbf{d}$ plays the role of 0 and ? plays the role of 1). Let $\gamma^* \in (0,1)$ be the critical probability of the percolation PCA with neighbourhood $\mathcal{N}$, see Ex. 3.1. For $\alpha < \gamma^*$, the percolation PCA G is ergodic. This completes the proof of (3.5).

Define a PCA H on the alphabet $\mathcal{C}$, with neighbourhood $\mathcal{N}$, and with the transition function:

$$h(\mathbf{d}^{\mathcal{N}}) = \delta_{\mathbf{d}}, \quad \text{and} \quad \forall u \in \mathcal{C}^{\mathcal{N}} - \{\mathbf{d}^{\mathcal{N}}\}, \quad h(u) = \beta\,\delta_? + (1-\beta)\,\delta_{\mathbf{d}}\;,$$

for $\beta = \min_{x\in\mathcal{B}^{\mathcal{N}}-\mathcal{A}^{\mathcal{N}}} \mathrm{env}(f)(x)(?)$. Given the way H is defined, we can construct an update function ϕ_H of H such that

$$\forall x \in \mathcal{B}^E, \forall y \in \mathcal{C}^E, \forall r \in [0,1]^E, \forall k \in E, \quad [x \in y,\ \phi_H(y,r)_k = ?] \implies \tilde{\phi}(x,r)_k = ?\;.$$

Therefore, the ergodicity of *env*(F) implies the ergodicity of H. Equivalently, the non-ergodicity of H implies the non-ergodicity of *env*(F). Observe that the PCA H is a percolation PCA. Therefore, for $\beta > \gamma^*$, the percolation PCA H is non-ergodic. This completes the proof of (3.6). □

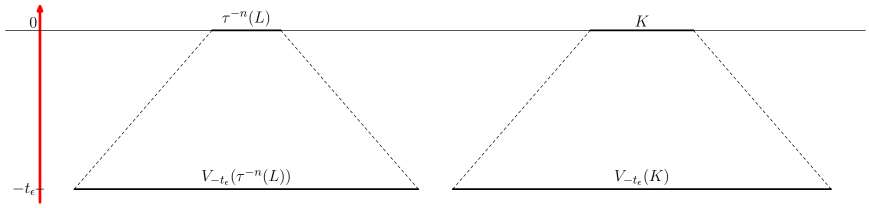


Figure 3.4: Illustration of the proof of Prop. 3.7.

Counter-examples

Recall Prop. 3.4: [EPCA ergodic] $\implies$ [PCA ergodic]. We now show that the converse is not true.

Example 3.2. Consider the PCA with alphabet $\mathcal{A} = \{0,1\}$, neighbourhood $\mathcal{N} = \{-1,0,1\}$, set of cells $E = \mathbb{Z}/n\mathbb{Z}$, and transition function

$$f(x,y,z) = \begin{cases} \delta_{1-y} & \text{if } xyz \in \{101, 010\} \\ \alpha\delta_y + (1-\alpha)\delta_{1-y} & \text{otherwise ,} \end{cases}$$

for a parameter $\alpha \in (0,1)$. This is the majority-flip PCA studied in Sec. 3.3. For n odd, it is easy to check that the PCA is ergodic. However the associated EPCA satisfies $\text{env}(f)(???) = \delta_?$. According to Prop. 3.5, the EPCA is not ergodic.

On the other hand, on a finite set of cells, if the PCA P is ergodic and has positive rates, then Prop. 3.5 let us conclude that the EPCA is ergodic. This is not true anymore for an infinite set of cells as emphasized by next example.

Example 3.3. Consider the noisy additive PCA of Ex. 2.2. This PCA has positive rates, in particular, it satisfies (3.4). So the EPCA is ergodic on a finite set of cells. Now let the set of cells be $\mathbb{Z}$. The PCA is ergodic for $\varepsilon \in (0,1)$, see Ex. 2.2. Consider the associated EPCA $\text{env}(F)$. Assume for instance that $\varepsilon \in (0,1/2)$. We have

$$\text{env}(f)(u) = \begin{cases} f(u) & \text{if } u \in \{\mathbf{0},\mathbf{1}\}^{\mathcal{N}} \\ \varepsilon\delta_{\mathbf{0}} + \varepsilon\delta_{\mathbf{1}} + (1-2\varepsilon)\delta_? & \text{otherwise .} \end{cases}$$

By applying Prop. 3.6, $\text{env}(F)$ is non-ergodic if $1 - 2\varepsilon > \gamma^*$.

3.2.5 Decay of correlations

In what follows, the set of cells is $E = \mathbb{Z}^d$, $d \geq 1$. It is easy to prove that the invariant measure of an ergodic PCA is shift-invariant. Using the coupling from the past tool, we give conditions for the invariant measure of an ergodic PCA to be shift-mixing. We recall that for $n \in \mathbb{Z}^d$, the shift σ^n is the homeomorphism defined by (1.1).

Definition 3.3 (Shift-mixing measure)**.** A measure μ on $\mathcal{X} = \mathcal{A}^{\mathbb{Z}^d}$ is *shift-mixing* if for any cylinder sets A, B of $\mathcal{X}$,

$$\lim_{\|n\| \to +\infty} \mu(A \cap \sigma^{-n}(B)) = \mu(A)\mu(B). \tag{3.8}$$

The proof of the following proposition is inspired from the proof of the validity of the coupling from the past method [PW96, HN98].

Proposition 3.7. *If Algorithm 2 stops almost surely, then the unique invariant measure of the PCA is shift-mixing. It is in particular the case under condition (3.5).*

Proof. Assume that F is an ergodic PCA, and denote by π its unique invariant measure. Let K and L be two finite subsets of E, and denote by $[x_K]$ and $[y_L]$ some cylinders corresponding to these subsets. Since the perfect sampling algorithm stops almost surely, for each $\varepsilon > 0$, there exists an integer t_ε such that with probability greater than $1-\varepsilon$, the algorithm stops before reaching the time $-t_\varepsilon$ when it is run for the set of cells K or for the set of cells L. If $n \in \mathbb{Z}^d$ is such that $\|n\|$ is large enough, the backward dependence cones corresponding to K and $\sigma^{-n}(L)$ are disjoint if they are considered only after time $-t_\varepsilon$, that is: $V_{-t_\varepsilon}(K) \cap V_{-t_\varepsilon}(\sigma^{-n}(L)) = \varnothing$ (see Fig. 3.4).

Let Z be the output of the algorithm if it is asked to sample the marginals of π corresponding to the cells of $K \cup \sigma^{-n}(L)$.

Imagine running the PCA from time $-t_\varepsilon$ and set of cells $V_{-t_\varepsilon}(K) \cup V_{-t_\varepsilon}(\sigma^{-n}(L))$ up to time 0, using the same update variables as the ones used to get Z. Choose the initial condition at time $-t_\varepsilon$ as follows: independently on $V_{-t_\varepsilon}(K)$ and $V_{-t_\varepsilon}(\sigma^{-n}(L))$, and according to the relevant marginals of π. Let X, resp. Y, be the output at time 0 on the set of cells K, resp. $\sigma^{-n}(L)$. Observe that X and Y are distributed according to the marginals of π. Furthermore, X and Y are independent since the dependence cones of K and $\sigma^{-n}(L)$ originating at time $-t_\varepsilon$ are disjoint.

We therefore obtain:

$$\begin{aligned}
\pi([x_K] \cap \sigma^{-n}([y_L])) - \pi([x_K])\pi([y_L]) &= \mathbb{P}(Z_K = x_K, Z_{\sigma^{-n}(L)} = y_L) - \mathbb{P}(X = x_K)\mathbb{P}(Y = y_L) \\
&= \mathbb{P}(Z_K = x_K, Z_{\sigma^{-n}(L)} = y_L) - \mathbb{P}(X = x_K, Y = y_L) \\
&\leq \mathbb{P}((Z_K, Z_{\sigma^{-n}(L)}) = (x_K, y_L) \text{ and } (X,Y) \neq (x_K, y_L)) \\
&\leq \mathbb{P}((Z_K, Z_{\sigma^{-n}(L)}) \neq (X,Y)) \ \leq \ 2\varepsilon \,.
\end{aligned}$$

In the same way, we get $\pi([x_K])\pi([y_L]) - \pi([x_K] \cap \sigma^{-n}([y_L])) \leq 2\varepsilon$. It completes the proof. □

In Prop. 3.7, the coupling from the past method is not used as a sampling tool but as a way to get theoretical results. Knowing if there exists an ergodic PCA having an invariant measure which is not shift-mixing is an open question [CT10].

3.2.6 Extensions

Non-homogeneous probabilistic cellular automata (NH-PCA)

In a PCA, the dynamic is homogeneous in space. It is possible to get rid of this characteristic by defining non-homogeneous PCA, for which the neighbourhood and the transition function depend on the position of the cell. The definition below is to be compared with Def. 1.4. The configuration space $\mathcal{X} = \mathcal{A}^E$ is unchanged.

Definition 3.4. For each $k \in E$, denote by $\mathcal{N}_k \subset E$ the (finite) neighbourhood of the cell k, and by $f_k : \mathcal{A}^{\mathcal{N}_k} \to \mathcal{M}(\mathcal{A})$ the transition function associated to k. Set $\mathcal{N}(K) = \cup_{k \in K} \mathcal{N}_k$. The *non-homogeneous PCA (NH-PCA)* of transition functions $(f_k)_{k \in E}$ is the application $F : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{X}),\ \mu \mapsto \mu F$, defined on cylinders by

$$\mu F[y_K] = \sum_{[x_{\mathcal{N}(K)}] \in \mathcal{C}(\mathcal{N}(K))} \mu[x_{\mathcal{N}(K)}] \prod_{k \in K} f_k((x_v)_{v \in \mathcal{N}_k})(y_k) \,.$$

Observe that it is not necessary for E to be equipped with a semigroup structure anymore. We use this below to define the finite restriction of a PCA.

It is quite straightforward to adapt the coupling from the past algorithms to NH-PCA. More precisely, given a NH-PCA, we define the associated NH-EPCA by considering Def. 3.2 and replacing $\mathcal{N}$ and $\text{env}(f)$ by $\mathcal{N}_k$ and $\text{env}(f)_k$ for each $k \in E$. The algorithms of Sec. 3.2.1 and 3.2.3 are then unchanged, and Prop. 3.4 and Th. 3.2 still hold in the non-homogeneous setting.

In Sec. 3.3, we study the majority-flip PCA by approximating it by a sequence of NH-PCA. Let us explain the construction in a general setting.

Let F be a PCA on the infinite set of cells $E = \mathbb{Z}^d$, with neighbourhood $\mathcal{N}$ and transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A})$. Let D be a finite subset of E. Define the *boundary* of D by

$$\partial D = (D + \mathcal{N}) \setminus D,$$

Fix a probability measure ν on $\mathcal{A}$. The *restriction* of F associated with ν and D is the NH-PCA $F(\nu, D)$ with set of cells $(D + \mathcal{N}) \cup D$ and neighbourhoods:

$$\forall k \in D,\ \mathcal{N}_k = \{k\} + \mathcal{N}, \qquad \forall k \in \partial D,\ \mathcal{N}_k = \varnothing\ ;$$

and transition functions:

$$\forall k \in D,\ f_k = f, \qquad \forall k \in \partial D,\ f_k(\cdot) = \nu\ .$$

In words, the boundary cells are i.i.d. of law ν and the cells of D are updated according to F.

If μ is a probability measure on $\mathcal{A}^S$, where S is a finite subset of E, we define its extension $\tilde{\mu}$ on $\mathcal{A}^E$ by setting, for a fixed letter $\alpha \in \mathcal{A}$:

$$\forall x \in \mathcal{A}^E,\ \tilde{\mu}(x) = \begin{cases} \mu((x_k)_{k \in S}) & \text{if } \forall i \in E \setminus S,\ x_i = \alpha \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3.1. *Let $(D_i)_{i \in \mathbb{N}}$ be an increasing sequence of finite domains $D_i \subset E$ such that $\cup_{i \in \mathbb{N}} D_i = E$. Let $(\nu_i)_{i \in \mathbb{N}}$ be a sequence of probability measures on $\mathcal{A}$. For each i, let μ_i be an invariant measure of $F(\nu_i, D_i)$. Any accumulation point of the sequence $(\tilde{\mu}_i)_{i \in \mathbb{N}}$ is an invariant measure of the original PCA F defined on E.*

Proof. Upon extracting a subsequence, we may assume that $(\tilde{\mu}_j)_{j \in \mathbb{N}}$ converges to $\tilde{\mu} \in \mathcal{M}(\mathcal{X})$. We need to prove that for any cylinder $[y_K] \in \mathcal{C}(K)$, we have $\tilde{\mu}F([y_K]) = \tilde{\mu}([y_K])$.

By definition, $\mu_j F(\nu_j, D_j) = \mu_j$. Let the subset K of E and the cylinder $y_K \in \mathcal{C}(K)$ be fixed. If j is large enough, $\mu_j([y_K]) = \tilde{\mu}_j([y_K])$, and $F(\nu_j, D_j)$ and F coincide on K. We deduce that $\tilde{\mu}_j F([y_K]) = \tilde{\mu}_j([y_K])$. By taking the limit on both sides, we get $\tilde{\mu}F([y_K]) = \tilde{\mu}([y_K])$. □

Alphabet with more than two elements

The EPCA and the associated algorithms have been defined on a two letters alphabet. It is possible to extend the approach to a general finite alphabet. In this paragraph, we give a sketch of the method that can be employed.

Let $\mathcal{A}$ be the finite alphabet. Let F be a PCA with set of cells E, neighbourhood $\mathcal{N}$, and transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A})$.

Consider the alphabet $\mathcal{B} = 2^{\mathcal{A}} - \{\varnothing\}$, that is, the set of non-empty subsets of $\mathcal{A}$. A word over $\mathcal{B}$ is viewed as a set of words over $\mathcal{A}$.

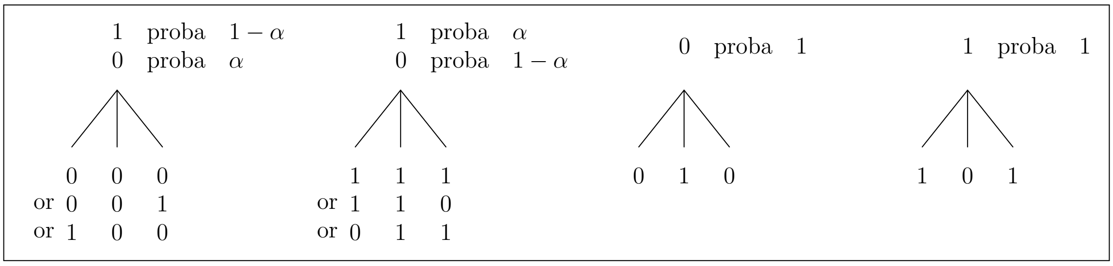


Figure 3.5: The transition function of the majority-flip PCA.

The EPCA $\text{env}(F)$ associated with F is a PCA on the alphabet $\mathcal{B}$ with neighbourhood $\mathcal{N}$ and transition function $\text{env}(f)$ that we now determine. Let us fix some $v \in \mathcal{B}^{\mathcal{N}}$ and define $\rho_S = \min_{u \in v} f(u)(\{S\})$. For a single letter $a \in \mathcal{A}$, we still want to have: $\text{env}(f)(v)(\{a\}) = \min_{u \in v} f(u)(a) = \rho_a$. Now, let us consider some $b \in \mathcal{A}$, $b \neq a$, we will set: $\text{env}(f)(v)(\{a,b\}) = \rho_{a,b} - \rho_a - \rho_b$, and so on.

By the inclusion-exclusion principle, we finally obtain the following formula for the transition function $\text{env}(f)$:

$$\forall v \in \mathcal{B}^{\mathcal{N}},\ \forall y \in \mathcal{B}, \quad \text{env}(f)(v)(y) = \sum_{x \subset y} (-1)^{|y|-|x|} \min_{u \in v} f(u)(x)\,.$$

For instance, $\text{env}(f)(v)(\{0,1,2\}) = \rho_{0,1,2} - \rho_{1,2} - \rho_{0,2} - \rho_{0,1} + \rho_0 + \rho_1 + \rho_2$.

The algorithms of Sec. 3.2.3 are unchanged. Observe however that the construction of an update function is not as natural as in the two-letters alphabet case.

3.3 The majority-flip PCA: a case study

The *majority-flip* PCA is one of the simplest examples of PCA whose behaviour is not well understood. Therefore, it provides a good case study for the sampling algorithms of Sec. 3.2.

3.3.1 Definition of the majority-flip PCA

Given $0 < \alpha < 1$, the PCA majority-flip(α), or simply majority-flip, is the PCA on the alphabet $\mathcal{A} = \{0,1\}$, with set of cells $E = \mathbb{Z}$ (or $\mathbb{Z}/n\mathbb{Z}$), neighbourhood $\mathcal{N} = \{-1,0,1\}$, and transition function

$$f(x,y,z) = \alpha\, \delta_{\text{maj}(x,y,z)} + (1-\alpha)\, \delta_{1-y}\,,$$

where $\text{maj} : \mathcal{A}^3 \to \mathcal{A}$ is the *majority function*: the value of $\text{maj}(x,y,z)$ is 0, resp. 1, if there are two or three 0's, resp 1's, in the sequence x, y, z. The transition function of majority-flip(α) can thus be represented as in Fig. 3.5. It consists in choosing independently for each cell to apply the elementary rule 232 (with probability α) or to flip the value of the cell.

The PCA Minority(α) has also been studied [Sch09]. It is defined by the transition function $g(x,y,z) = f(1-x, 1-y, 1-z)$.

Gray has proved that all one-dimensional positive-rate monotonic two-state nearest-neighbour PCA are ergodic [Gra87]. But here, majority-flip is not a rule with positive rates, and it is not even monotonic, so that we cannot use this result.

Let $x = (01)^{\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}}$ be defined by: $\forall n \in \mathbb{Z},\ x_{2n} = 0,\ x_{2n+1} = 1$. The configuration $(10)^{\mathbb{Z}}$ is defined similarly. Consider the probability measure

$$\mu = (\delta_{(01)^{\mathbb{Z}}} + \delta_{(10)^{\mathbb{Z}}})/2\,. \tag{3.9}$$

Clearly, μ is an invariant measure for the majority-flip PCA. The question is whether other invariant measures exist.

To get some insight on this question, consider the majority-flip PCA on the set of cells $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. This PCA has two completely different behaviours depending on the parity of n. Indeed, a simple analysis of the structure of the transition matrix shows that the Markov chain has a unique invariant measure which is $(\delta_{(01)^{n/2}} + \delta_{(10)^{n/2}})/2$ if n is even, and which is supported on $\{0,1\}^{\mathbb{Z}_n}$ if n is odd.

Let us come back to the majority-flip PCA on $\mathbb{Z}$. The invariant measure μ in (3.9) can be viewed as the "limit" over n of the invariant measures of the PCA on $\mathbb{Z}_{2n}$. What about the "limits" of the invariant measures of the PCA on $\mathbb{Z}_{2n+1}$? Do they define other invariant measures for the PCA on $\mathbb{Z}$?

One of the motivations of our work on perfect sampling algorithms for PCA was to test the following conjecture, which is inspired by the observations made by Regnault [Reg08] and Schabanel [Sch09] on a PCA equivalent to majority-flip. This conjecture concerns the existence of a phase transition phenomenon for the marjority-flip PCA.

Conjecture. There exists $\alpha_c \in (0,1)$ such that majority-flip(α) has a unique invariant measure if $\alpha < \alpha_c$, and several invariant measures if $\alpha > \alpha_c$.

In the next subsection, we give some rigorous (but partial) results about the invariant measures of majority-flip(α). We first introduce a related PCA and use it to prove that if α is large enough, majority-flip(α) has indeed non-trivial invariant measures; we then present a dual model that could be used to provide some information for small values of α. The last subsection is devoted to the experimental study of majority-flip(α) using the perfect sampling tools developed in the previous section.

3.3.2 Theoretical study

A related model: the "flip-if-not-all-equal" PCA

Let us define as in the work of Regnault [Reg08], the PCA FINAE(α) of neighbourhood $\mathcal{N} = \{-1,0,1\}$ and transition function $g : \{0,1\}^{\mathcal{N}} \to \mathcal{M}(\{0,1\})$ given by

$$g(x,y,z) = \alpha\, \delta_{\text{flip-if-not-all-equal}(x,y,z)} + (1-\alpha)\, \delta_y \,,$$

where the function flip-if-not-all-equal (FINAE), corresponding to the elementary cellular automaton 178, is defined by

$$\text{flip-if-not-all-equal}(x,y,z) = \begin{cases} y & \text{if } x = y = z \\ 1-y & \text{otherwise.} \end{cases}$$

Clearly, $\delta_{0^{\mathbb{Z}}}$ and $\delta_{1^{\mathbb{Z}}}$ are invariant measures of the PCA. Let us define flip-odd : $\{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ and flip-even : $\{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ by, for $x = (x_i)_{i\in\mathbb{Z}}$,

$$\text{flip-odd}(x)_i = \begin{cases} x_i & \text{if } i \text{ is even} \\ 1-x_i & \text{if } i \text{ is odd} \end{cases}, \qquad \text{flip-even}(x)_i = \begin{cases} 1-x_i & \text{if } i \text{ is even} \\ x_i & \text{if } i \text{ is odd.} \end{cases}$$

If we extend flip-odd and flip-even to mappings on $\mathcal{M}(\{0,1\}^{\mathbb{Z}})$, we have

$$\text{majority-flip}(\alpha) = \text{flip-odd} \circ \text{FINAE}(\alpha) \circ \text{flip-even}\,.$$

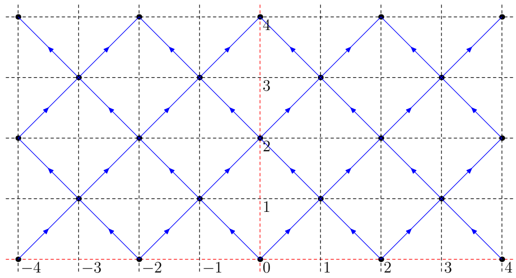


Figure 3.6: The graph G.

This equality can be checked on the local functions of the PCA majority-flip(α) and FINAE(α). One thus obtains that if π is an invariant measure for FINAE(α), then

$$(\text{flip-odd}(\pi) + \text{flip-even}(\pi))/2$$

is an invariant measure for majority-flip(α). The invariant measures $\delta_{0^{\mathbb{Z}}}$ and $\delta_{1^{\mathbb{Z}}}$ of FINAE(α) correspond to the invariant measure μ in (3.9) for majority-flip(α), and the existence of a non-trivial invariant measure for FINAE(α) corresponds to the existence of a second invariant measure for majority-flip(α).

Validity of the conjecture for large values of α

The partial result of Prop. 3.8 relies on ideas from Regnault [Reg08].

Proposition 3.8. *Let p_c be the percolation threshold of directed bond-percolation in $\mathbb{N}^2$. If $\alpha \geq \sqrt[3]{1-(1-p_c)^4}$, then majority-flip($\alpha$) has several invariant measures (resp. FINAE(α) has other invariant measures than the combinations of $\delta_{0^{\mathbb{Z}}}$ and $\delta_{1^{\mathbb{Z}}}$). It is in particular the case if $\alpha \geq 0.996$.*

Proof. It is known that $0.6298 \leq p_c \leq 2/3$, see for instance the work of Grimmett [Gri99]. This provides the bound $\sqrt[3]{1-(1-p_c)^4} \leq 0.996$. Let us consider the directed graph $G = (N, A)$ such that the set of nodes is $N = 2\mathbb{Z} \times 2\mathbb{N} \cup (2\mathbb{Z}+1) \times (2\mathbb{N}+1)$ and for each $(i,j) \in N$, there is an arc (oriented bond) from (i,j) to $(i-1, j+1)$ and one from (i,j) to $(i+1, j+1)$.

Let S be some subset of $2\mathbb{Z} \times \{0\}$ called the *source*. The oriented bond-percolation on G of parameter p and source S is defined as follows: each bound (edge) is open with probability p and closed with probability $1-p$, independently of the others, and a node of N is said to be *wet* if there is an open path joining it from some node of S. We say that the space-time diagram $(x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ of FINAE(α) and the percolation model satisfy the *correspondence criterion at time t* if for each wet cell (k,t) of height t, we have $x_k^t \neq x_{k+1}^t$ or $x_k^t \neq x_{k-1}^t$. For values of (α, p) satisfying $\alpha \geq \sqrt[3]{1-(1-p)^4}$, Regnault is able to construct a coupling between FINAE(α) and the percolation model such that if the correspondence criterion is true at time t, it is still true at time $t+1$. Let us take for the initial configuration of FINAE(α) the configuration x^0 defined by $x_k^0 = 1$ if k is odd and $x_k^0 = 0$ if n is even. We also choose $S = 2\mathbb{Z} \times \{0\}$ for the percolation model. The correspondence criterion is true at time 0. By the coupling of Regnault [Reg08], the criterion is true at all time. Consider the percolation model and the probability $\mathbb{P}((0,2t)$ is wet). It is known [Gri99] that if p is strictly greater than a certain critical value p_c, this probability, which decreases with t, does not tend to 0. Thus, for $p > p_c$, there exists $\eta_p > 0$ such that $\mathbb{P}((0,2t)$ is wet$) > \eta_p$ for all $t \in \mathbb{N}$. By construction of the coupling, we obtain $\mathbb{P}(x_0^{2t} \neq x_1^{2t}$ or $x_0^{2t} \neq x_{-1}^{2t}) \geq \eta_p$ for all $t \in \mathbb{N}$. This proves that for $\alpha \geq \sqrt[3]{1-(1-p_c)^4}$, the PCA FINAE($\alpha$) has at least one invariant measure which is not in the convex hull of the Dirac masses at the configurations

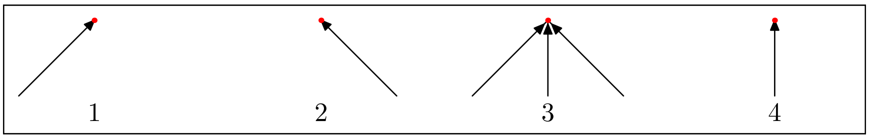


Figure 3.7: Construction of the graph G_1.

"all zeroes" and "all ones" (take any accumulation point of the Cesàro sums obtained from the sequence obtained from the iterated of δ_{x^0} by FINAE). This result can be translated to the majority-flip PCA. □

A duality result with the double branching annihilating random walk

The aim of this subsection is to prove a duality result between FINAE(α) and a double branching annihilating random walk (DBARW). The connection between these two models is interesting in itself and could provide a new way to study the PCA majority-flip(α) for small values of α. A similar duality result was already obtained for interacting particle systems [CD91], and the behaviour of the DBARW is very well understood in continuous time [Sud90], but its study appears to be more difficult in discrete time.

We now assume that $\alpha \leq 2/3$ (in particular, Prop. 3.8 does not apply). Let us define a process $(x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ in the following way. For each $(k,t) \in \mathbb{Z}\times\mathbb{N}$, we first choose independently to do one (and only one) of the following:

1. with probability $\alpha/2$, draw an arc from $(k-1,t)$ to $(k,t+1)$,
2. with probability $\alpha/2$, draw an arc from $(k+1,t)$ to $(k,t+1)$,
3. with probability $\alpha/2$, draw an arc from $(k-1,t)$ to $(k,t+1)$, an arc from (k,t) to $(k,t+1)$, and an arc from $(k+1,t)$ to $(k,t+1)$,
4. with probability $1-3\alpha/2$, draw an arc from (k,t) to $(k,t+1)$.

We thus obtain a directed graph G_1, that we will use to label each node of $\mathbb{Z}\times\mathbb{N}$ with a letter of $\{0,1\}$. The nodes of $\mathbb{Z}\times\{0\}$ are labeled according to the initial configuration x^0. A node labeled by a 1 will be interpreted as being occupied. A node $(k,t)\in\mathbb{Z}\times\mathbb{N}$ is then labeled by a 1 if and only if there is an odd number of paths leading to this node from an occupied node of $\mathbb{Z}\times\{0\}$. This define a random field $(x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ representing the labels of the nodes.

We claim that this field has the same distribution as a space-time diagram of FINAE(α) starting from x^0. Indeed, the value x_k^{t+1} is equal to x_{k-1}^t with probability $\alpha/2$, to x_{k+1}^t with probability $\alpha/2$, to $x_{k-1}^t + x_k^t + x_{k+1}^t \mod 2$ with probability $\alpha/2$ and to x_k^t with probability $1-3\alpha/2$. And one can check for each value $(x,y,z)\in\{0,1\}^3$ that these probabilities coincide with the ones obtained with the local function flip-if-not-all-equal. For example, if $(x_{k-1}^t, x_k^t, x_{k+1}^t) = (0,0,1)$, the value of x_k^{t+1} will be 1 if and only if case 2 or case 3 occurs, and they have together a probability $\alpha/2+\alpha/2=\alpha$. If $(x_{k-1}^t, x_k^t, x_{k+1}^t) = (0,1,0)$, we will have $x_k^{t+1}=1$ if and only if case 3 or case 4 occurs, which has a probability $\alpha/2+(1-3\alpha/2) = 1-\alpha$. And if $(x_{k-1}^t, x_k^t, x_{k+1}^t) = (0,0,0)$ (resp. $(1,1,1)$), we will get a 0 (resp. a 1) in all cases.

We now consider the process $(y_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ obtained from $(x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ by reversing time. Formally, for each $(k,t)\in\mathbb{Z}\times\mathbb{N}$, we first choose independently to do one (and only one) of the following things:

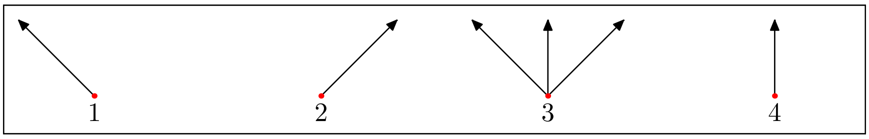


Figure 3.8: Construction of the graph G_2.

1. with probability $\alpha/2$, draw an arc from (k,t) to $(k-1,t+1)$;
2. with probability $\alpha/2$, draw an arc from (k,t) to $(k+1,t+1)$;
3. with probability $\alpha/2$, draw an arc from (k,t) to $(k-1,t+1)$, an arc from (k,t) to $(k,t+1)$, and an arc from (k,t) to $(k+1,t+1)$;
4. with probability $1-3\alpha/2$, draw an arc from (k,t) to $(k,t+1)$.

We thus obtain again a directed graph G_2, that we will use to label each node of $\mathbb{Z}\times\mathbb{N}$ with a letter of $\{0,1\}$. The nodes of $\mathbb{Z}\times\{0\}$ are labeled according to the initial configuration y^0. A node labeled by a 1 will be interpreted as being occupied. A node $(k,t)\in\mathbb{Z}\times\mathbb{N}$ is then labeled by a 1 if and only if there is an odd number of paths leading to this node from an occupied node of $\mathbb{Z}\times\{0\}$. This define a random field $(y_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ representing the labels of the nodes. We claim that this field has the same distribution as the double branching annihilating random walk that we now define. At time 0, a particle is placed on each cell k of $\mathbb{Z}$ such that $y_k^0=1$, and at each step of time, every particle chooses independently of the others do one (and only one) of the following things:

1. with probability $\alpha/2$, move from node k to $k-1$;
2. with probability $\alpha/2$, move from node k to $k+1$;
3. with probability $\alpha/2$, stay at node k and create two new particles at nodes $k-1$ and $k+1$;
4. with probability $1-3\alpha/2$, stay at node k.

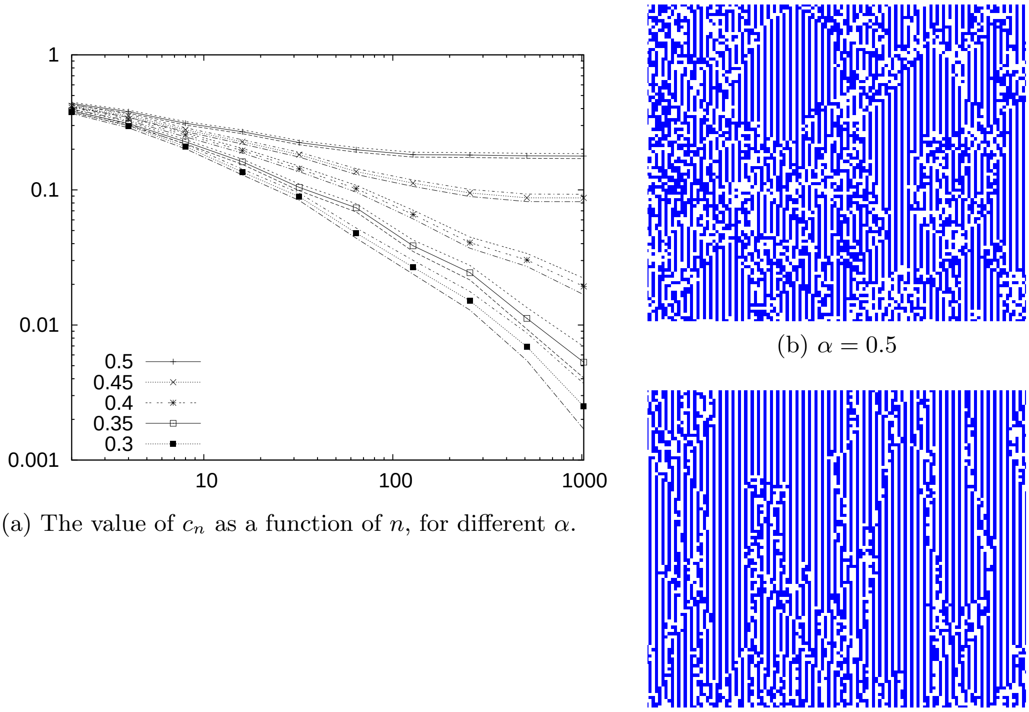
If after these choices, there is an even number of particles at a node, then all these particles annihilate. If there is an odd number of them, only one particle survives. We set $w_k^t=1$ if and only if at time t, there is a particle at node k.

To summarise, we have the following relations:

$$\text{FINAE} \sim (x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}} \overset{\text{time-reversal}}{\longleftrightarrow} (y_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}} \sim \text{DBARW}$$

The processes $(x_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ and $(y_k^t)_{k\in\mathbb{Z},t\in\mathbb{N}}$ are obtained one from another by reversing time. This can be used to get nontrivial information for FINAE. For instance, if A represents the set of occupied nodes at time 0 for x, that is to say $x^0=1_A$, we have the following duality relation:

$$\mathbb{P}^{\,x^0=1_A}(x_k^t\neq x_l^t)=\mathbb{P}^{\,x^0=1_A}(x_k^t+x_l^t=1)$$
$$=\mathbb{P}(\text{the total number of paths in } G_1 \text{ leading from } A\times\{0\} \text{ to } (k,t) \text{ or } (l,t) \text{ is odd})$$
$$=\mathbb{P}(\text{the total number of paths in } G_2 \text{ leading from } (k,0) \text{ or } (l,0) \text{ to } A\times\{t\} \text{ is odd})$$
$$=\mathbb{P}^{\,y^0=1_{\{k,l\}}}\Big(\sum_{i\in A} y_i^t \text{ is odd}\Big)$$



(a) The value of c_n as a function of n, for different α.

(b) $\alpha = 0.5$

(c) $\alpha = 0.3$

Figure 3.9: Experimental study of majority-flip(α) (the configurations at odd times only are represented on the space-time diagrams).

$$\leq \mathbb{P}^{\,y^0 = 1_{\{k,l\}}}(\exists i \in A, y_i^t = 1).$$

Thus, to prove that the probability for the PCA FINAE that two cells k and l will be in different states at time t tends to 0 as t tends to $+\infty$, it is sufficient to prove that in the DBARW, starting from two particles, the probability of extinction of the population of particles tends to 1.

3.3.3 Experimental study

We tried to get some numerical evidence for the conjecture of Sec. 3.3.1 using the perfect sampling tools developed in the previous section. To study the majority-flip PCA experimentally, a first idea would be to consider the same PCA on the set of cells $\mathbb{Z}_n$, n odd. This does not work well. First, due to the state space explosion, computing exactly the invariant measure is possible only for small values (we did it up to $n = 9$). Second, the algorithms of Sec. 3.2 cannot be applied since the EPCA is not ergodic.

Instead, we use approximations of the PCA by NH-PCA on a finite subset of cells, the methodology sketched in Sec. 3.2.6. Again, computing exactly the invariant measure is impossible except for very small windows. But now the sampling algorithms become effective.

Let P be the majority-flip PCA. Set $D_n = \{-n, \dots, n\}$, and let ν be the uniform measure on $\{0, 1\}$. Consider the NH-PCA $P(\nu, D_n)$. Let μ_n be the unique invariant measure of $P(\nu, D_n)$. We are interested in the quantity

$$c_n = \mu_n\{x \in \{0,1\}^{D_n} \mid x_0 = x_1 = 0\} + \mu_n\{x \in \{0,1\}^{D_n} \mid x_0 = x_1 = 1\}\,.$$

Indeed, by application of Lemma 3.1, if $\limsup_n c_n > 0$, then there exists a non-trivial invariant measure for the majority-flip PCA on $\mathbb{Z}$.

Now the NH-EPCA is ergodic, so the sampling algorithms of Sec. 3.2 can be used. We were able to run the algorithms up to a window size of $n = 1024$ before running into a timeout. The experimental results appear in Fig. 3.9, with a logarithmic scale. We ran the sampling algorithms 10^4 times. We show on the figure the confidence intervals calculated with Wilson score test at 95%.

It is reasonable to believe that the top two curves in Fig. 3.9 do not converge to 0 while the bottom three converge to 0. This would imply that at least for $\alpha \geq 0.45$, the PCA has several invariant measures, which is consistent with the visual impression of space-time diagrams.

Part II

Randomisation, conservation, classification

Chapter 4

Probabilistic cellular automata having Bernoulli or Markov invariant measures and random fields with i.i.d. directions

Beauty, the world seemed to say. And as if to prove it (scientifically) wherever he looked at the houses, at the railings, at the antelopes stretching over the palings, beauty sprang instantly. To watch a leaf quivering in the rush of air was an exquisite joy.

– Virginia Woolf, *Mrs. Dalloway*

Contents

Let us consider the simplest model of one-dimensional probabilistic cellular automata (PCA). The set of cells is $\mathbb{Z}$, the alphabet is $\mathcal{A} = \{0, 1\}$, and the neighbourhood is $\mathcal{N} = \{0, 1\}$, meaning that the new content of each cell is randomly chosen, independently of the others, according to a distribution depending only on the content of the cell itself and of its right neighbour.

There are necessary and sufficient conditions on the four parameters of such a PCA to have a Bernoulli product invariant measure. We study the properties of the random field given by the space-time diagram obtained when iterating the PCA starting from its Bernoulli product invariant measure.

It is a non-trivial random field with very weak dependences and nice combinatorial properties. In particular, not only the horizontal lines but also the lines in any other direction

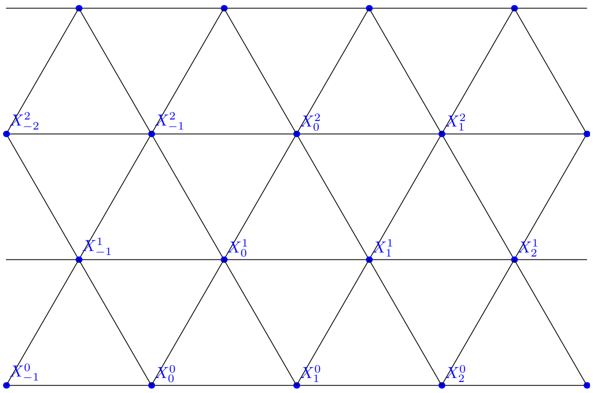


Figure 4.1: Space-time diagram.

consist of i.i.d. random variables. We study extensions of the results to Markov invariant measures, and to PCA with larger alphabets and neighbourhoods.

Let us consider a PCA of neighbourhood $\mathcal{N}$. For a time $n \in \mathbb{N}$ and a cell $i \in \mathbb{Z}^d$, the *dependence cone* $\mathcal{D}(i,n)$ of (i,n) is the set of coordinates in the space-time diagram that are likely to be influenced by the value of X_i^n. Precisely, we introduce the next definition.

Definition 4.1 (Dependence cone)**.** The *dependence cone* of $(i,n) \in \mathbb{Z}^d \times \mathbb{N}$ is the set

$$\mathcal{D}(i,n) = \{(k,t) \in \mathbb{Z}^d \times \mathbb{N}; t \geq n \text{ and } i = k + v_1 + \ldots + v_{t-n} \text{ for some } v_1, \ldots, v_{t-n} \in \mathcal{N}^{t-n}\}.$$

The next lemma follows from the definition of a PCA.

Lemma 4.1. *Let (i,n) belong to $\mathbb{Z}^d \times (\mathbb{N} \setminus \{0\})$ and let S be a subset of $\mathbb{Z}^d \times \mathbb{N}$ such that $\mathcal{D}(i,n) \cap S = \emptyset$. Then, X_i^n is independent of $(X_j^m)_{(j,m)\in S}$ conditionally to $(X_{i+v}^{n-1})_{v\in\mathcal{N}}$.*

4.1 Elementary PCA having Bernoulli invariant measures

For the time being, we assume that the neighbourhood is $\mathcal{N} = \{0,1\}$ and that the alphabet is $\mathcal{A} = \{0,1\}$.

When the neighbourhood is $\mathcal{N} = \{0,1\}$, for symmetry reasons, a natural choice can be to represent the space-time diagram on a regular triangular lattice, as in Fig. 4.1.

For convenience, we introduce the notations: for $x, y \in \mathcal{A}$,

$$\theta_{xy} = \theta_{xy}^1 = f(x,y)(1), \qquad \theta_{xy}^0 = f(x,y)(0) = 1 - \theta_{xy}\,.$$

Observe that a PCA is completely characterised by the four parameters: $\theta_{00}, \theta_{01}, \theta_{10}$, and θ_{11}.

4.1.1 Computation of the image of a product measure by a PCA

The goal of this section is to give an explicit description of the measure $\mu_p F$, where μ_p is the Bernoulli product measure of parameter p, as a function of the parameters $\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11}$.

Let us start with an observation. Consider $(Y_n)_{n\in\mathbb{Z}} \sim \mu_p F$. Let $q \in [0,1]$ be such that $Y_0 \sim \mathcal{B}_q$ (that is, $q = (1-p)^2\theta_{00} + (1-p)p(\theta_{01}+\theta_{10}) + p^2\theta_{11}$). Clearly, we have: $(Y_{2n})_{n\in\mathbb{Z}} \sim \mu_q$ and $(Y_{2n+1})_{n\in\mathbb{Z}} \sim \mu_q$. But the two i.i.d. sequences have a complex joint correlation structure. It makes it non-elementary to describe the finite-dimensional marginals of $\mu_p F$.

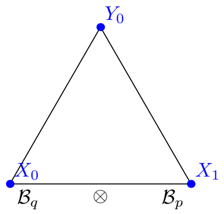
Assume that the parameters satisfy:

$$(\theta_{00},\theta_{01}),(\theta_{10},\theta_{11}) \notin \{(0,0),(1,1)\}\ . \tag{4.1}$$

For $p \in (0,1)$, $\alpha \in \{0,1\}$, define the function

$$\begin{array}{rccl} g_\alpha: & [0,1] & \longrightarrow & (0,1) \\ & q & \longmapsto & (1-q)\ (1-p)\ \theta_{00}^\alpha + (1-q)\ p\ \theta_{01}^\alpha + q\ (1-p)\ \theta_{10}^\alpha + q\ p\ \theta_{11}^\alpha\ . \end{array} \tag{4.2}$$

Consider three random variables X_0, X_1, Y_0 with $(X_0,X_1) \sim \mathcal{B}_q \otimes \mathcal{B}_p$ and $Y_0 \sim (\mathcal{B}_q \otimes \mathcal{B}_p)f$. In words, $g_\alpha(q)$ is the probability to have $Y_0 = \alpha$. With the condition (4.1), we have $g_\alpha(q) \in (0,1)$ for all $q \in [0,1]$. Observe also that $g_0(q) + g_1(q) = 1$.



For $p \in (0,1)$, $\alpha \in \{0,1\}$, we also define the function

$$\begin{array}{rccl} h_\alpha: & [0,1] & \longrightarrow & [0,1] \\ & q & \longmapsto & \big[(1-q)\ p\ \theta_{01}^\alpha + q\ p\ \theta_{11}^\alpha\big] g_\alpha(q)^{-1}\ . \end{array} \tag{4.3}$$

Consider X_0, X_1, Y_0 with $(X_0,X_1) \sim \mathcal{B}_q \otimes \mathcal{B}_p$ and $Y_0 \sim (\mathcal{B}_q \otimes \mathcal{B}_p)f$. In words, $h_\alpha(q)$ is the probability to have $X_1 = 1$ conditionally to $Y_0 = \alpha$.

Proposition 4.1. *Consider a PCA satisfying (4.1). Consider $p \in (0,1)$. For $\alpha_0 \cdots \alpha_{n-1} \in \mathcal{A}^n$, the probability of the cylinder $[\alpha_0 \cdots \alpha_{n-1}]$ under $\mu_p F$ is given by:*

$$\mu_p F[\alpha_0 \cdots \alpha_{n-1}] = g_{\alpha_0}(p) \prod_{i=1}^{n-1} g_{\alpha_i}(h_{\alpha_{i-1}}(h_{\alpha_{i-2}}(\ldots h_{\alpha_0}(p)\ldots)))\ .$$

By reversing the space-direction, we get an analogous proposition for a PCA satisfying the symmetric condition: $(\theta_{00},\theta_{10}),(\theta_{01},\theta_{11}) \notin \{(0,0),(1,1)\}$.

Proof. Let us compute recursively the value $\mu_p F[\alpha_0 \cdots \alpha_{n-1}]$. We set $X = X^0$ and $Y = X^1$. Assuming that $X \sim \mu_p$, by definition,

$$\mu_p F[\alpha_0] = \mathbb{P}(Y_0 = \alpha_0) = g_{\alpha_0}(p).$$

We can decompose the probability $\mu_p F[\alpha_0 \alpha_1]$ into

$$\mu_p F[\alpha_0\alpha_1] = \mathbb{P}(Y_0 = \alpha_0, Y_1 = \alpha_1) = \mathbb{P}(Y_1 = \alpha_1 \mid Y_0 = \alpha_0)\ \mathbb{P}(Y_0 = \alpha_0).$$

By definition, the conditional law of X_1 assuming that $Y_0 = \alpha_0$ is given by $\mathcal{B}_{h_{\alpha_0}(p)}$. So the law of (X_1, X_2) is $\mathcal{B}_{h_{\alpha_0}(p)} \otimes \mathcal{B}_p$ and we obtain

$$\mu_p F[\alpha_0\alpha_1] = g_{\alpha_1}(h_{\alpha_0}(p))\ g_{\alpha_0}(p).$$

More generally, we have:

$$\mathbb{P}(Y_0 = \alpha_0 \ldots Y_k = \alpha_k) = \mathbb{P}(Y_k = \alpha_k \mid Y_0 = \alpha_0 \ldots Y_{k-1} = \alpha_{k-1})\ \mathbb{P}(Y_0 = \alpha_0 \ldots Y_{k-1} = \alpha_{k-1}).$$

By induction, the law of X_k knowing that $Y_0 = \alpha_0 \ldots Y_{k-1} = \alpha_{k-1}$ is $\mathcal{B}_{h_{\alpha_{k-1}}(h_{\alpha_{k-2}}(\ldots h_{\alpha_0}(p)\ldots))}$. The result follows. $\square$

4.1.2 Conditions for a Bernoulli measure to be invariant

For $x \in \mathcal{X}$, denote by δ_x the Dirac probability measure concentrated on the configuration x. The probability measure $\mu_1 = \delta_{1^{\mathbb{Z}}}$ is invariant for the PCA F if and only if $\theta_{11} = 1$. Similarly, $\mu_0 = \delta_{0^{\mathbb{Z}}}$ is invariant for F if and only if $\theta_{00} = 0$.

Using Prop. 4.1, we get a necessary and sufficient condition for μ_p, $p \in (0,1)$, to be an invariant measure of F. The result is stated in Th. 4.1. The conditions we obtain were already known [BGM69, DKT90], but our proof is new and simpler.

Theorem 4.1. *The measure μ_p, $p \in (0,1)$, is an invariant measure of the PCA F of parameters $\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11}$ if and only if one of the two following conditions is satisfied:*

$$\begin{array}{ll} (i) & (1-p)\,\theta_{00} + p\,\theta_{01} = (1-p)\,\theta_{10} + p\,\theta_{11} = p \\ (ii) & (1-p)\,\theta_{00} + p\,\theta_{10} = (1-p)\,\theta_{01} + p\,\theta_{11} = p\,. \end{array}$$

In particular, a PCA has a (non-trivial) Bernoulli product invariant measure if and only if its parameters satisfy:

$$\theta_{00}\,(1-\theta_{11}) = \theta_{10}\,(1-\theta_{01}) \qquad or \qquad \theta_{00}\,(1-\theta_{11}) = \theta_{01}\,(1-\theta_{10})\,. \tag{4.4}$$

Proof. Let us assume that F satisfies condition *(i)* for some $p \in (0,1)$. Then, the function g_1 is given by $g_1(q) = (1-q)\,p + q\,p = p$, and $g_0(q) = 1 - g_1(q) = 1-p$. By Prop. 4.1, we have,

$$\forall \alpha = \alpha_0 \cdots \alpha_{n-1} \in \mathcal{A}^n, \quad \mu_p F[\alpha] = (1-p)^{|\alpha|_0} p^{|\alpha|_1} = \mu_p[\alpha]\,.$$

So μ_p is an invariant measure.

Now, assume that the PCA F satisfies condition *(ii)*. Let us reverse the space direction, that is, let us read the configurations from right to left. The same dynamic is now described by a new PCA $\widetilde{F}$ defined by the parameters $\tilde{\theta}_{00} = \theta_{00}, \tilde{\theta}_{01} = \theta_{10}, \tilde{\theta}_{10} = \theta_{01}, \tilde{\theta}_{11} = \theta_{11}$. So, the new PCA satisfies condition (i). According to the above, we have $\mu_p \widetilde{F} = \mu_p$. Let us reverse the space direction, once again. Since the Bernoulli product measure is unchanged, we obtain $\mu_p F = \mu_p$.

Conversely, assume that $\mu_p F = \mu_p$. It follows from Prop. 4.1 that for any value of the α_i, we must have $g_1(h_{\alpha_{n-1}}(h_{\alpha_{n-2}}(\ldots h_{\alpha_0}(p)\ldots))) = p$. Since g_1 is an affine function, there are only two possibilities: either g_1 is the constant function equal to p; or $h_{\alpha_{n-1}}(h_{\alpha_{n-2}}(\ldots h_{\alpha_0}(p)\ldots))) = p$ for all values of $\alpha_0, \ldots, \alpha_{n-1} \in \mathcal{A}$.

In the first case, observe that

$$g_1(q) = q\,[-(1-p)\,\theta_{00} - p\,\theta_{01} + (1-p)\,\theta_{10} + p\,\theta_{11}] + (1-p)\,\theta_{00} + p\,\theta_{01}\,.$$

To get: $\forall q \in [0,1],\ g_1(q) = p$, we must have condition *(i)*.

In the second case, we must have $h_0(p) = h_1(p) = p$ and $g_1(p) = p$. Using $g_0(p) = 1-p$ and $g_1(p) = p$, we get:

$$\begin{array}{rcl} h_0(p) & = & \big[(1-p)\,p\,(1-\theta_{01}) + p\,p\,(1-\theta_{11})\big](1-p)^{-1} \\ h_1(p) & = & \big[(1-p)\,p\,\theta_{01} + p\,p\,\theta_{11}\big]p^{-1} \;=\; (1-p)\,\theta_{01} + p\,\theta_{11}\,. \end{array}$$

The equality $h_1(p) = p$ provides the condition $(1-p)\,\theta_{01} + p\,\theta_{11} = p$. Let us switch to the equality $h_0(p) = p$. We have:

$$\begin{array}{rcl} h_0(p) = p & \iff & (1-p)\,(1-\theta_{01}) + p\,(1-\theta_{11}) = 1-p \\ & \iff & (1-p)\,\theta_{01} + p\,\theta_{11} = p\,. \end{array}$$

So, we obtain condition *(ii)*. $\square$

To complete Th. 4.1, let us quote another result [Vas78]. We recall that a PCA has positive rates if: $\forall u \in \mathcal{A}^{\mathcal{N}}, \forall a \in \mathcal{A},\ f(u)(a) > 0$.

Proposition 4.2. *Consider a positive-rate PCA F satisfying condition (i) or (ii), for some $p \in (0,1)$. Then F is ergodic, that is, μ_p is the unique invariant measure of F and for all initial measure μ, the sequence $(\mu F^n)_{n\geq 0}$ converges weakly to μ_p.*

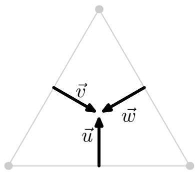
Assessing the ergodicity of a PCA is a difficult problem, which is algorithmically undecidable in general, see Chap. 3. In this complicated landscape, Prop. 4.2 gives a restricted setting in which ergodicity can be proved.

Observe that Prop. 4.2 is not true without the positive rates assumption. Consider for instance the PCA defined by: $\theta_{00} = p/(1-p), \theta_{01} = 0, \theta_{10} = 0, \theta_{11} = 1$ for some $p \in (0, 1/2]$. It satisfies (i) and (ii), but it is not ergodic since $\delta_{1^{\mathbb{Z}}}$ and μ_p are both invariant.

4.1.3 Transversal PCA

We assume that μ_p is invariant under the action of the PCA, and we focus on the correlation structure of the space-time diagram obtained when the initial measure is μ_p. Observe that this space-time diagram is both space-stationary and time-stationary. By time-stationarity, the space-time diagram can be extended from $\mathbb{Z} \times \mathbb{N}$ to $\mathbb{Z}^2$. From now on, we work with this extension.

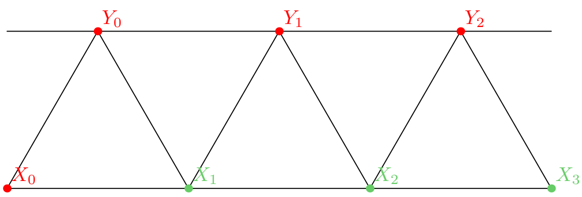
Let $(X_{k,n})_{k,n\in\mathbb{Z}\times\mathbb{Z}}$ be a realisation of the stationary space-time diagram.



It is convenient to define the three vectors $\vec{u}, \vec{v}$, and $\vec{w}$ as in the figure above. The PCA generating the space-time diagram is the PCA of direction $\vec{u}$. In some cases, the space-time diagram when rotated by an angle of $2\pi/3$ (resp. $-2\pi/3$) still has the correlation structure of a space-time diagram generated by a PCA of neighbourhood $\{0,1\}$. In this case, we say that, in the original space-time diagram, there is a *transversal PCA* of direction $\vec{v}$ (resp. $\vec{w}$).

Proposition 4.3. *Under condition (i), each line of angle $\pi/3$ of the space-time diagram is distributed according to μ_p. Moreover, their correlations are the ones of a* transversal *PCA of direction $\vec{v}$ and rates given by: $\vartheta_{00} = \theta_{00}$, $\vartheta_{01} = \theta_{10}$, $\vartheta_{10} = \theta_{01}$, $\vartheta_{11} = \theta_{11}$.*

To prove Prop. 4.3, we need two preliminary lemmas. Set $X = X^0$ and $Y = X^1$, so that we have in particular $(X, Y) \sim (\mu_p, \mu_p F)$.



Lemma 4.2. *Under condition (i), the variables $(Y_k)_{k\geq 0}$ are independent of X_0, that is, for any $n \geq 0$,*

$$\mathbb{P}\big(X_0 = x_0, (Y_i)_{0\leq i\leq n} = (y_i)_{0\leq i\leq n}\big) = \mu_p[x_0] \prod_{i=0}^{n} \mu_p[y_i]\,.$$

Proof. The left-hand side can be decomposed into:

$$\sum_{x_1\cdots x_{n+1}\in\{0,1\}^{n+1}} \mathbb{P}\big((X_i)_{0\le i\le n+1}=(x_i)_{0\le i\le n+1},(Y_i)_{0\le i\le n}=(y_i)_{0\le i\le n}\big),$$

which can be expressed with the transition rates of the PCA as follows:

$$\sum_{x_1\cdots x_{n+1}\in\{0,1\}^{n+1}} \mu_p[x_0]\prod_{i=0}^{n}\mu_p[x_{i+1}]\theta^{y_i}_{x_ix_{i+1}}$$

$$=\mu_p[x_0]\sum_{x_1\in\{0,1\}}\mu_p[x_1]\theta^{y_0}_{x_0x_1}\sum_{x_2\in\{0,1\}}\mu_p[x_2]\theta^{y_1}_{x_1x_2}\cdots\sum_{x_{n+1}\in\{0,1\}}\mu_p[x_{n+1}]\theta^{y_n}_{x_nx_{n+1}}.$$

Condition (i) can be rewritten as:

$$\forall a,c\in\{0,1\},\qquad \sum_{b\in\{0,1\}}\mu_p[b]\theta^c_{ab}=\mu_p[c]\,.$$

Using this, and simplifying from the right to the left, we obtain: $\mu_p[x_0]\prod_{i=0}^n\mu_p[y_i]$. □

Lemma 4.3. *Under condition (i), for any $n\ge 0$,*

$$\mathbb{P}(X_0=x_0,X_1=x_1,(Y_i)_{0\le i\le n}=(y_i)_{0\le i\le n})=\mu_p[x_0]\mu_p[x_1]\theta^{y_0}_{x_0x_1}\prod_{i=1}^n\mu_p[y_i]\,.$$

Proof. The proof is analogous. We decompose the left-hand side into:

$$\sum_{x_2\cdots x_{n+1}\in\{0,1\}^{n}} \mathbb{P}((X_i)_{0\le i\le n+1}=(x_i)_{0\le i\le n+1},(Y_i)_{0\le i\le n}=(y_i)_{0\le i\le n}),$$

which can be expressed with the transition rates of the PCA as follows:

$$\sum_{x_2\cdots x_{n+1}\in\{0,1\}^{n}} \mu_p[x_0]\prod_{i=0}^{n}\mu_p[x_{i+1}]\theta^{y_i}_{x_ix_{i+1}}$$

$$=\mu_p[x_0]\mu_p[x_1]\theta^{y_0}_{x_0x_1}\sum_{x_2\in\{0,1\}}\mu_p[x_2]\theta^{y_1}_{x_1x_2}\cdots\sum_{x_{n+1}\in\{0,1\}}\mu_p[x_{n+1}]\theta^{y_n}_{x_nx_{n+1}}.$$

Using (i) and simplifying from the right to the left, we get the result. □

Proof of Prop. 4.3. To prove the first part of the proposition, it is sufficient to prove that the sequence $(X_0^k)_{k\in\mathbb{Z}}$ is i.i.d. For a given $n\in\mathbb{N}$ and a sequence $(\alpha_k)_{0\le k\le n}$, let us prove recursively that $\mathbb{P}((X_0^n)_{0\le k\le n}=(\alpha_k)_{0\le k\le n})=\mu_p[\alpha_0\cdots\alpha_n]$. For $n=0$, the result is straightforward; and for $n=1$, it is a direct consequence of Lemma 4.2. For larger values of n, set $A=\mathbb{P}((X_0^k)_{0\le k\le n}=(\alpha_k)_{0\le k\le n})$, we have:

$$A\;=\sum_{y_1\cdots y_{n-1}\in\{0,1\}^{n-1}}\mathbb{P}\big((X_0^k)_{0\le k\le n}=(\alpha_k)_{0\le k\le n},(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}\big)\,.$$

Since $X_0^0=X_0, X_0^1=Y_0$, it can be rewritten as:

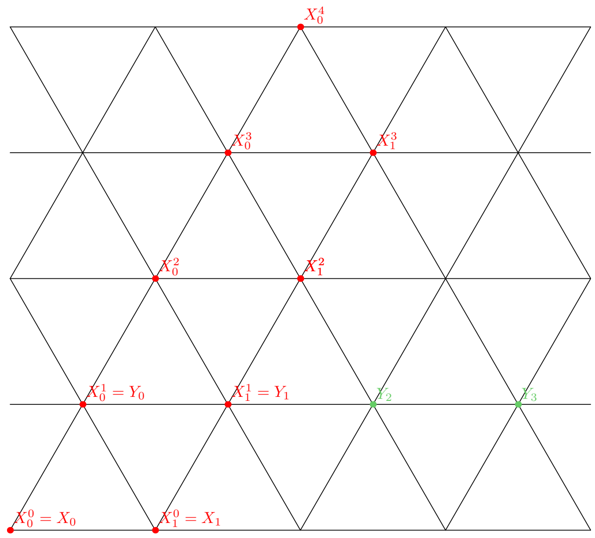
$$A\;=\sum_{y_1\cdots y_{n-1}\in\{0,1\}^{n-1}}\mathbb{P}\big((X_0^k)_{2\le k\le n}=(\alpha_k)_{2\le k\le n}\mid X_0=\alpha_0,Y_0=\alpha_1,(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}\big)$$

$$\times\,\mathbb{P}\big(X_0=\alpha_0,Y_0=\alpha_1,(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}\big)\,.$$

The law of $(X_0^k)_{2\le k\le n}$ conditionally to $(X_0,(Y_i)_{0\le i\le n-1})$ is equal to the law of $(X_0^k)_{2\le k\le n}$ conditionally to $(Y_i)_{0\le i\le n-1}$. Also, using Lemma 4.2, we have: $\mathbb{P}\big(X_0=\alpha_0, Y_0=\alpha_1, (Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}\big)=\mu_p[\alpha_0]\ \mathbb{P}(Y_0=\alpha_1,(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1})$. Coupling these two points, we get:

$$\begin{aligned} A &= \sum_{y_1\cdots y_{n-1}\in\{0,1\}^{n-1}} \mathbb{P}\big((X_0^k)_{2\le k\le n}=(\alpha_k)_{2\le k\le n}\mid Y_0=\alpha_1,(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}\big) \\ &\qquad\qquad \times\ \mu_p[\alpha_0]\ \mathbb{P}(Y_0=\alpha_1,(Y_i)_{1\le i\le n-1}=(y_i)_{1\le i\le n-1}) \\ &= \mu_p[\alpha_0]\ \mathbb{P}\big((X_0^k)_{1\le k\le n}=(\alpha_k)_{1\le k\le n}\big)\,. \end{aligned}$$

By induction, we obtain the result.



The second part of the proposition consists of proving that

$$\mathbb{P}\big((X_1^k)_{0\le k\le n}=(\beta_k)_{0\le k\le n}\mid (X_0^k)_{0\le k\le n+1}=(\alpha_k)_{0\le k\le n+1}\big)=\prod_{k=0}^{n}\vartheta_{\alpha_{k+1}\alpha_k}^{\beta_k}\,. \tag{4.5}$$

We prove the result recursively. For $n=0$, set $A=\mathbb{P}(X_1=\beta_0\mid Y_0=\alpha_1,X_0=\alpha_0)$. We want to prove that $A=\vartheta_{\alpha_1\alpha_0}^{\beta_0}$. Using the first part of the proposition, we have:

$$\begin{aligned} A &= \mathbb{P}(Y_0=\alpha_1\mid X_0=\alpha_0,X_1=\beta_0)\ \mathbb{P}(X_0=\alpha_0,X_1=\beta_0)\ \mathbb{P}(X_0=\alpha_0,Y_0=\alpha_1)^{-1} \\ &= \theta_{\alpha_0\beta_0}^{\alpha_1}\ \mu_p[\alpha_0]\ \mu_p[\beta_0]\ \mu_p[\alpha_0]^{-1}\ \mu_p[\alpha_1]^{-1} \quad = \quad \theta_{\alpha_0\beta_0}^{\alpha_1}\ \mu_p[\beta_0]\ \mu_p[\alpha_1]^{-1}\,. \end{aligned}$$

If $\alpha_1=\beta_0=u$, we get $A=\theta_{\alpha_0 u}^{u}=\vartheta_{u\alpha_0}^{u}$. Assume that $\alpha_1\neq\beta_0$. Condition (i) can be rewritten as:

$$\mu_p[\beta_0]\ \theta_{\alpha_0\beta_0}^{\alpha_1}+\mu_p[\alpha_1]\ \theta_{\alpha_0\alpha_1}^{\alpha_1}=\mu_p[\alpha_1]\,. \tag{4.6}$$

Dividing by $\mu_p[\alpha_1]$, we get:

$$A=\theta_{\alpha_0\beta_0}^{\alpha_1}\ \mu_p[\beta_0]\ \mu_p[\alpha_1]^{-1}=1-\theta_{\alpha_0\alpha_1}^{\alpha_1}=\theta_{\alpha_0\alpha_1}^{\beta_0}=\vartheta_{\alpha_1\alpha_0}^{\beta_0}\,.$$

For larger n, it is convenient to prove the next equality, which is equivalent to (4.5):

$$\mathbb{P}\big((X_0^k)_{0\le k\le n+1}=(\alpha_k)_{0\le k\le n+1},(X_1^k)_{0\le k\le n}=(\beta_k)_{0\le k\le n}\big)\ =\ \mu_p[\alpha_{n+1}]\prod_{k=0}^{n}\mu_p[\alpha_k]\vartheta_{\alpha_{k+1}\alpha_k}^{\beta_k}.$$

The left-hand side can be decomposed into:

$$\sum_{y_2\cdots y_n\in\{0,1\}^{n-1}} \mathbb{P}\big((X_0^k)_{0\le k\le n+1}=(\alpha_k)_{0\le k\le n+1},(X_1^k)_{0\le k\le n}=(\beta_k)_{0\le k\le n},(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big).$$

Let us decompose each term of the sum, conditioning by the values of X_0, X_1, Y_0, and Y_1.
We have:

$$\mathbb{P}\big((X_0^k)_{2\le k\le n+1}=(\alpha_k)_{2\le k\le n+1},(X_1^k)_{2\le k\le n}=(\beta_k)_{2\le k\le n}\mid (X_0,X_1,Y_0,Y_1)=(\alpha_0,\beta_0,\alpha_1,\beta_1),(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)$$
$$=\mathbb{P}\big((X_0^k)_{2\le k\le n+1}=(\alpha_k)_{2\le k\le n+1},(X_1^k)_{2\le k\le n}=(\beta_k)_{2\le k\le n}\mid (Y_0,Y_1)=(\alpha_1,\beta_1),(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)\,.$$

and using Lemma 4.3, and the equality $\mu_p[\beta_0]\theta^{\alpha_1}_{\alpha_0\beta_0}=\mu_p[\alpha_1]\vartheta^{\beta_0}_{\alpha_1\alpha_0}$ (see (4.6)):

$$\begin{aligned}
&\mathbb{P}\big((X_0,X_1,Y_0,Y_1)=(\alpha_0,\beta_0,\alpha_1,\beta_1),(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)\\
&\quad= \mu_p[\alpha_0]\mu_p[\beta_0]\theta^{\alpha_1}_{\alpha_0\beta_0}\ \mathbb{P}\big(Y_1=\beta_1,(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)\\
&\quad= \mu_p[\alpha_0]\mu_p[\alpha_1]\vartheta^{\beta_0}_{\alpha_1\alpha_0}\ \mathbb{P}\big(Y_1=\beta_1,(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)\\
&\quad= \mu_p[\alpha_0]\vartheta^{\beta_0}_{\alpha_1\alpha_0}\ \mathbb{P}\big((Y_0,Y_1)=(\alpha_1,\beta_1),(Y_i)_{2\le i\le n}=(y_i)_{2\le i\le n}\big)\,.
\end{aligned}$$

Assembling the pieces together, we obtain:

$$\mathbb{P}\big((X_0^k)_{0\le k\le n+1}=(\alpha_k)_{0\le k\le n+1},(X_1^k)_{0\le k\le n}=(\beta_k)_{0\le k\le n}\big)$$
$$=\mu_p[\alpha_0]\vartheta^{\beta_0}_{\alpha_1\alpha_0}\ \mathbb{P}\big((X_0^k)_{1\le k\le n+1}=(\alpha_k)_{1\le k\le n+1},(X_1^k)_{1\le k\le n}=(\beta_k)_{1\le k\le n}\big)\,.$$

We conclude the proof by induction. □

Lemma 4.4. *Let F be a PCA of neighbourhood $\{0,\dots,\ell\}$. Assume that $\mu_pF=\mu_p$ and consider the stationary space-time diagram obtained for that invariant measure. Then for any $\alpha>-1/\ell$, the line $L_\alpha=\{(k,n)\in\mathbb{Z}\times\mathbb{N}\mid n=\alpha k\}$ is such that the random variables $(X_k^n)_{(k,n)\in L_\alpha}$ are i.i.d.*

The lines described above are those which are outside the dependence cone of the PCA.

Proof. Let us show that any finite sequence of consecutive random variables on such a line is i.i.d. We can assume without loss of generality that the first of these points is X_0^0. Then, using the hypothesis on the slope, we obtain that the other random variables on that line are all outside the dependence cone of X_0^0. Thus, the $(n-1)$-tuple they constitute is independent of X_0^0. By induction, we get the result. □

Corollary 4.1. *Under condition (i), all the lines of the space-time diagram except possibly those of angle $2\pi/3$ consist of i.i.d. random variables.*

Proof. The previous proposition claims that the lines of angle $\pi/3$ are i.i.d. Lemma 4.4 provides the result for the lines of angles in $[0,\pi/3)\cup(2\pi/3,\pi]$. The angles in $(\pi/3,2\pi/3)$ correspond to lines that are outside the dependence cones of the transversal PCA, so we obtain the result by applying again Lemma 4.4 for the transversal PCA. □

In the same way, one can prove the following.

Proposition 4.4. *Under condition (ii), the lines of angle $2\pi/3$ of the space-time diagram are distributed according to μ_p and their correlations are those of a* transversal PCA of direction $\vec{w}$ and rates given by $\vartheta_{00}=\theta_{00}$, $\vartheta_{11}=\theta_{11}$ and $\vartheta_{01}=\theta_{10}$, $\vartheta_{10}=\theta_{01}$.

Corollary 4.2. *Under condition (ii), all the lines of the space-time diagram except possibly the ones of angle $\pi/3$ consist of i.i.d. random variables.*

For a PCA satisfying *(i)* (resp. *(ii)*), the lines of angle $2\pi/3$ (resp. $\pi/3$) are not i.i.d., except if the PCA also satisfies condition *(ii)* (resp. *(i)*). The distribution of the lines of angle $2\pi/3$ (resp. $\pi/3$) does not necessary have a Markov form either. For example, if $\theta_{00} = \theta_{01} = 1/2$ and $\theta_{10} = 0, \theta_{11} = 1$ (condition *(i)* is satisfied with $p = 1/2$), one can check that $\mathbb{P}(X_0^0 = 0, X_{-1}^1 = 0, X_{-2}^2 = 0) = 19/64$ which is different $\mathbb{P}(X_0^0 = 0)\mathbb{P}(X_{-1}^1 = 0 \mid X_0^0 = 0)\mathbb{P}(X_{-2}^2 = 0 \mid X_{-1}^1 = 0) = (1/2)(3/4)^2$.

It is an open problem to know if under condition *(i)* (resp. *(ii)*), it is possible to give an explicit description of the distribution of the lines of angle $2\pi/3$ (resp. $\pi/3$).

4.2 Spatial properties of the space-time diagram

We now concentrate on PCA satisfying **both** conditions (i) and (ii) for some $p \in (0,1)$. We consider the stationary space-time diagram associated with μ_p, and we still denote it by $(X_k^n)_{k,n\in\mathbb{Z}}$.

4.2.1 A random field with i.i.d. directions

For a given $p \in (0,1)$, conditions *(i)* and *(ii)* are both satisfied if and only if:

$$\exists s \in \Big[\frac{2p-1}{p}, \frac{p}{1-p}\Big], \qquad \theta_{00} = \frac{p(1-s)}{1-p},\ \theta_{01} = \theta_{10} = s,\ \theta_{11} = 1 - \frac{(1-p)s}{p}\ . \tag{4.7}$$

Example 4.1. For any value of $p \in (0,1)$, the choice $s = p$ is allowed. In that case, the transition rates θ_{ij} are all equal to p and the stationary random field is i.i.d., there is no dependence in the space-time diagram.

Example 4.2. If $p = 1/2$, every choice of $s \in [0,1]$ is valid and the corresponding PCA has the transition function $f(x,y) = s\ \delta_{x+y} + (1-s)\ \delta_{x+y+1}$, where the sums $x+y$ and $x+y+1$ are taken modulo 2. We recover the PCA of Ex. 2.2.

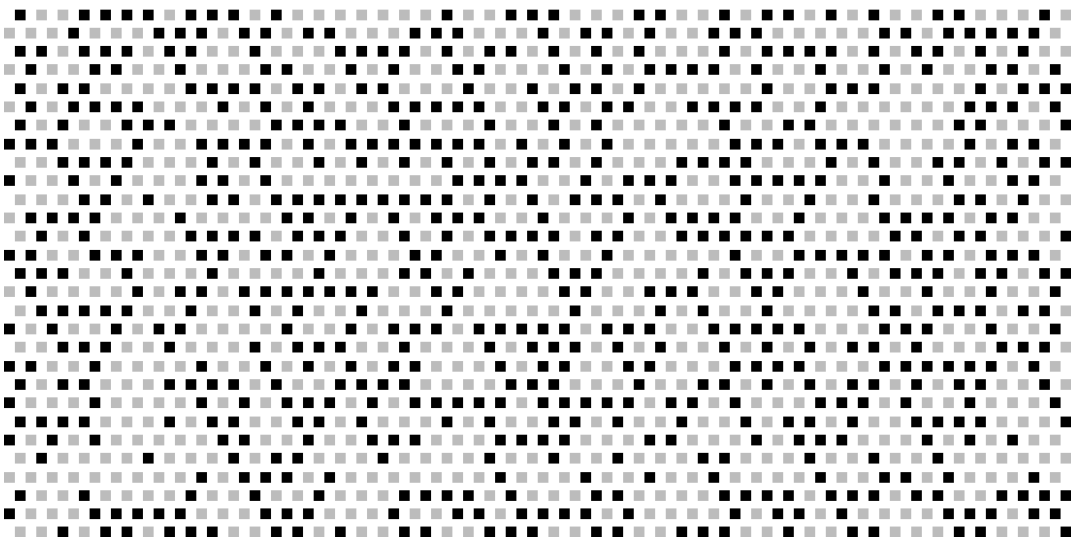


Figure 4.2: An example of space-time diagram for $p = 1/2$ and $s = 3/4$ (Ex. 4.2).

Example 4.3. For any value of $p \in (0, 1/2]$, it is possible to set $s = 0$ and then, $\theta_{01} = \theta_{10} = 0, \theta_{11} = 1$, and $\theta_{00} = p/(1-p)$. This PCA forbids the elementary triangles pointing up that have exactly one vertex labeled by a 0.

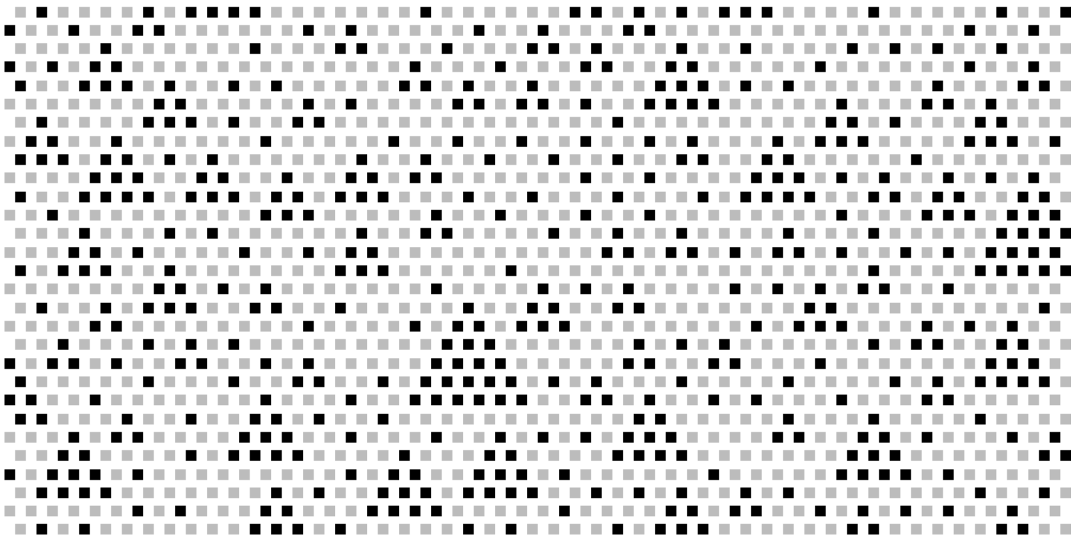


Figure 4.3: An example of space-time diagram for $p = 1/3$ and $s = 0$ (Ex. 4.3).

The next proposition is a direct consequence of Corollaries 4.1 and 4.2.

Proposition 4.5. *Consider a PCA satisfying (4.7). Every line of the stationary space-time diagram consists of i.i.d. random variables. In particular, any two different variables are independent.*

4.2.2 Correlations in triangles

We have seen that all the lines of the space-time diagram are i.i.d. But the whole space-time diagram is i.i.d. if and only if $s = p$. Indeed, if $s \neq p$, the random variable X_k^{n+1} is not independent of (X_k^n, X_{k+1}^n); in words, the three variables of an elementary triangle pointing up are correlated. Precisely, the triple $(X_k^n, X_{k+1}^n, X_k^{n+1})$ consists of random variables which are: (1) identically distributed; (2) pairwise independent; (3) globally dependent if $s \neq p$. The "converse" holds.

Proposition 4.6. *Let ν be a law on $\{0,1\}^3$ such that the three marginals on $\{0,1\}^2$ are i.i.d. Assume that ν is non-degenerate ($\nu \neq \delta_{000}, \nu \neq \delta_{111}$). Then ν can be realised as the law of an "elementary triangle pointing up" in the stationary space-time diagram of exactly one PCA satisfying (4.7).*

Proof. Consider $(X_0, X_1, Y_0) \sim \nu$. Assume that the common law of X_0, X_1, and Y_0 is $\mathcal{B}_p$. By the pairwise independence, we have:

$$\begin{aligned}\mathbb{P}(X_0 = 1, X_1 = 0, Y_0 = 0) &= \mathbb{P}(X_1 = 0, Y_0 = 0) - \mathbb{P}(X_0 = 0, X_1 = 0, Y_0 = 0) \\ &= (1-p)^2 - \mathbb{P}(X_0 = 0, X_1 = 0, Y_0 = 0)\,.\end{aligned}$$

We obtain:

$$\begin{aligned}&\mathbb{P}(X_0 = 1, X_1 = 0, Y_0 = 0) = \mathbb{P}(X_0 = 0, X_1 = 1, Y_0 = 0) = \mathbb{P}(X_0 = 0, X_1 = 0, Y_0 = 1) \\ &\mathbb{P}(X_0 = 0, X_1 = 1, Y_0 = 1) = \mathbb{P}(X_0 = 1, X_1 = 0, Y_0 = 1) = \mathbb{P}(X_0 = 1, X_1 = 1, Y_0 = 0)\,.\end{aligned}$$

Set $q_0 = \mathbb{P}(X_0 = 1, X_1 = 0, Y_0 = 0)$ and $q_1 = \mathbb{P}(X_0 = 0, X_1 = 1, Y_0 = 1)$. We have:

$$\mathbb{P}(X_0 = 0, X_1 = 0, Y_0 = 0) = (1-p)^2 - q_0, \qquad \mathbb{P}(X_0 = 1, X_1 = 1, Y_0 = 1) = p^2 - q_1 .$$

Furthermore:

$$q_0 + q_1 = \mathbb{P}(X_0 = 0, X_1 = 0, Y_0 = 1) + \mathbb{P}(X_0 = 1, X_1 = 0, Y_0 = 1) = \mathbb{P}(X_1 = 0, Y_0 = 1) = p(1-p).$$

Using the above, and expressing everything as a function of p and q_1, we get:

$$\begin{array}{rcl} \mathbb{P}(Y_0 = 1 \mid X_0 = 0, X_1 = 0) & = & (p(1-p) - q_1)/(1-p)^2 \\ \mathbb{P}(Y_0 = 1 \mid X_0 = 0, X_1 = 1) & = & q_1/(p(1-p)) \\ \mathbb{P}(Y_0 = 1 \mid X_0 = 1, X_1 = 0) & = & q_1/(p(1-p)) \\ \mathbb{P}(Y_0 = 1 \mid X_0 = 1, X_1 = 1) & = & 1 - q_1/p^2 . \end{array}$$

By setting $\theta_{ij} = \mathbb{P}(Y_0 = 1 \mid X_0 = i, X_1 = j)$ and $s = q_1/(p(1-p))$, we recover exactly (4.7). □

Lemma 4.5. *Consider a PCA satisfying (4.7). The random field $(X_{2k}^{2n})_{k,n\in\mathbb{Z}}$ corresponds to the space-time diagram of a new PCA, having a neighbourhood of size 2, and satisfying (4.7) for the same value of p.*

Proof. Let us consider the random field $(X_{2k}^{2n})_{k,n\in\mathbb{Z}}$. Observe that all its random variables are distributed according to $\mathcal{B}_p$, and that each line consists of i.i.d. random variables.

We complete the proof of Lemma 4.5 by considering a realisation of the space-time diagram. Let us assume that $(X_i^{2n})_{i\in\mathbb{Z}}$ is a sequence of i.i.d. random variables, of distribution $\mathcal{B}_p$. Let $(r_i^m)_{i\in\mathbb{Z},m>2n}$ be i.i.d. random variables, independent from the X_i^{2n}, and uniformly distributed on $[0,1]$, such that X_i^m is a deterministic function of X_i^{m-1}, X_{i+1}^{m-1} and r_i^m. Precisely, we define:

$$X_i^m = \begin{cases} 0 \text{ if } r_i^m < 1 - \theta_{X_i^{m-1} X_{i+1}^{m-1}} = f(X_i^{m-1}, X_{i+1}^{m-1})(0) \\ 1 \text{ otherwise.} \end{cases}$$

This provides a realisation of the space-time diagram (at least from time $2n$, but by stationarity, we can in fact also assume that the whole space-time diagram is built this way).

The random variable X_{2k}^{2n+2} can be written as a deterministic function of the following variables (see Fig. 4.4):

$$X_{2k}^{2n},\ X_{2k+1}^{2n},\ X_{2k+2}^{2n},\ r_{2k}^{2n+1},\ r_{2k+1}^{2n+1},\ r_{2k}^{2n+2}.$$

Conditionally on the variables $(X_{2k}^{2n})_{a\le k\le b+1}$, the variables:

$$(X_{2k+1}^{2n})_{a\le k\le b},\ (r_i^{2n+1})_{2a\le i\le 2b+1},\ (r_{2i}^{2n+2})_{a\le i\le b}$$

are still independent, and for different values of k, the variables X_{2k}^{2n+2} are deterministic functions of different variables among the above ones. Thus, for any $a < b$, the variables $(X_{2k}^{2n+2})_{a\le k\le b}$ are independent conditionally to the variables $(X_{2k}^{2n})_{a\le k\le b+1}$. □

Proposition 4.7. *Consider a PCA satisfying (4.7) with $s \neq p$. The correlations between three random variables that form an equilateral triangle pointing up decrease exponentially as a function of the size of the triangle.*

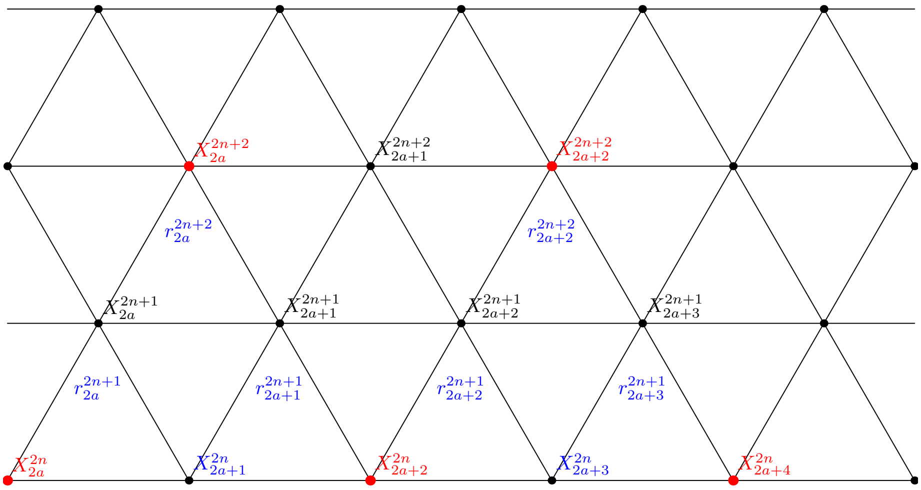


Figure 4.4: Illustration of the proof of Lemma 4.5.

Proof. Let us consider the random field $(X_{2k}^{2n})_{k,n\in\mathbb{Z}}$. By Lemma 4.5, this "extracted" random field corresponds to the space-time diagram of a new PCA, having a neighbourhood of size 2 and satisfying (4.7) for the same value of p. To know its transition rates $\theta_{ij}^{(2)} = \mathbb{P}(X_0^2 = 1 \mid X_0^0 = i, X_2^0 = j)$, it is enough to compute $\theta_{10}^{(2)} = \theta_{01}^{(2)}$. We denote this value by $\phi(s)$, since it is a function of $s = \theta_{01} = \theta_{10}$.

Summing over all possible values of X_1^0, X_0^1, X_1^1 (we first consider the case $X_1^0 = 1$ and then the one $X_1^0 = 0$), we get:

$$\phi(s) = p\,[\theta_{01}\theta_{11}\,\theta_{11} + (1-\theta_{01})\theta_{11}\,\theta_{01} + \theta_{01}(1-\theta_{11})\,\theta_{10} + (1-\theta_{01})(1-\theta_{11})\,\theta_{00}]$$

$$+(1-p)\,[\theta_{00}\theta_{01}\,\theta_{11} + (1-\theta_{00})\theta_{01}\,\theta_{01} + \theta_{00}(1-\theta_{01})\,\theta_{10} + (1-\theta_{00})(1-\theta_{01})\,\theta_{00}].$$

Replacing the coefficients θ_{ij} by their expression as a function of p and s and simplifying the result, we obtain:

$$\phi(s) = p + \frac{(s-p)^3}{p(1-p)}.$$

We proceed similarly for the random field $(X_{2^ik}^{2^in})_{k,n\in\mathbb{Z}}$. The coefficient $\theta_{01}^{(2^i)} = \mathbb{P}(X_0^{2^i} = 1 \mid X_0^0 = 0, X_{2^i}^0 = 1)$ is equal to $\phi^i(s)$, which satisfies:

$$\phi^i(s) - p = \frac{(s-p)^{3^i}}{(p(1-p))^{\frac{3^i-1}{2}}} = \sqrt{p(1-p)}\left(\frac{s-p}{\sqrt{p(1-p)}}\right)^{3^i}.$$

Similar computations can be performed for equilateral triangles pointing up of other sizes. The decay of correlation for equilateral triangles pointing up is exponential in function of their size. □

The next lemma will allow us to characterise completely the triples of random variables that are not independent.

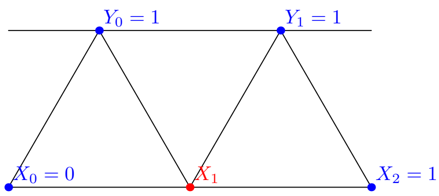
Lemma 4.6. *Consider a PCA satisfying (4.7). The variable X_0^0 is independent of the sequence $(X_k^n)_{k\in\mathbb{Z},n\in\mathbb{N}\setminus\{0\}}$.*

Proof. Set $X = X^0$ and $Y = X^1$. It is sufficient to prove that X_0 is independent of $(Y_k)_{k\in\mathbb{Z}}$. But $(Y_k)_{k\geq 0}$ and $(Y_k)_{k<0}$ are independent conditionally to X_0, so that we can conclude with Lemma 4.2 and its analogue for condition *(ii)*. □

Proposition 4.8. *Consider a PCA satisfying (4.7) with $s \neq p$. Three random variables of the stationary space-time diagram are correlated if and only if they form an equilateral triangle pointing up.*

Proof. Three variables that form an equilateral triangle pointing up are correlated, see the proof of Prop. 4.7. Let us now consider three variables (Z_1, Z_2, Z_3) that do not constitute such a triangle. Then, if we consider the smallest equilateral triangle pointing up that contains them, there is an edge of that triangle that contains exactly one of these variables. By rotation of angle $2\pi/3$ or translation of the diagram, one can assume that this edge is the horizontal one and that it contains the variable Z_1, and not the variables Z_2, Z_3. Now, using Lemma 4.6, we obtain that Z_1 is independent of (Z_2, Z_3). But since Z_2 and Z_3 are independent, the three variables (Z_1, Z_2, Z_3) are independent. □

There are subsets of four variables that do not contain equilateral triangles pointing up and that are correlated. It is the case in general of (X_0, X_2, Y_0, Y_1). Let us consider for instance the PCA of Ex. 4.3. The event $(X_0, X_2, Y_0, Y_1) = (0, 1, 1, 1)$ has probability zero, since whatever the value of X_1, the space-time diagram would have an elementary triangle pointing up with exactly one zero.



4.2.3 Incremental construction of the random field

Let us show how to construct incrementally the stationary space-time diagram of a PCA satisfying conditions *(i)* and *(ii)*, using two elementary operations, based respectively on Lemmas 4.1 and 4.6.

Consider a PCA satisfying *(i)* and *(ii)* for some $p \in (0, 1)$. Let $S \subset \mathbb{Z}^2$ be the finite set of points of the space-time diagram that has been constructed at some step. Initially $S = \{(0, 0)\}$ and $X_0^0 \sim \mathcal{B}_p$.

- If $(i, n), (i + 1, n) \in S, (i, n + 1) \notin S$, and $\mathcal{D}(i, n + 1) \cap S = \emptyset$. Choose X_i^{n+1} knowing (X_i^n, X_{i+1}^n) according to the law of the PCA.

  If $(i, n), (i, n+1) \in S, (i+1, n) \notin S$, and if no point of the dependence cone of $(i+1, n)$ with respect to the transversal PCA of direction $\vec{v}$ belongs to S: choose X_{i+1}^n knowing (X_i^{n+1}, X_i^n) according to the law of the transversal PCA of direction $\vec{v}$.

  If $(i, n + 1), (i + 1, n) \in S, (i, n) \notin S$, and if no point of the dependence cone of (i, n) with respect to the transversal PCA of direction $\vec{w}$ belongs to S: choose X_i^n knowing (X_{i+1}^n, X_i^{n+1}) according to the law of the transversal PCA of direction $\vec{w}$.

- If $(i, n) \notin S$, and if $(j, m) \in S$ implies $m > n$: choose X_i^n according to $\mathcal{B}_p$ and independently of the variables $X_j^m, (j, m) \in S$.

  If $(i, n) \notin S$, and if $(j, m) \in S$ implies $j > i$: choose X_i^n according to $\mathcal{B}_p$ and independently of the variables $X_j^m, (j, m) \in S$.

  If $(i, n) \notin S$, and if $(j, m) \in S$ implies $j + m < i + n$: choose X_i^n according to $\mathcal{B}_p$ and independently of the variables $X_j^m, (j, m) \in S$.

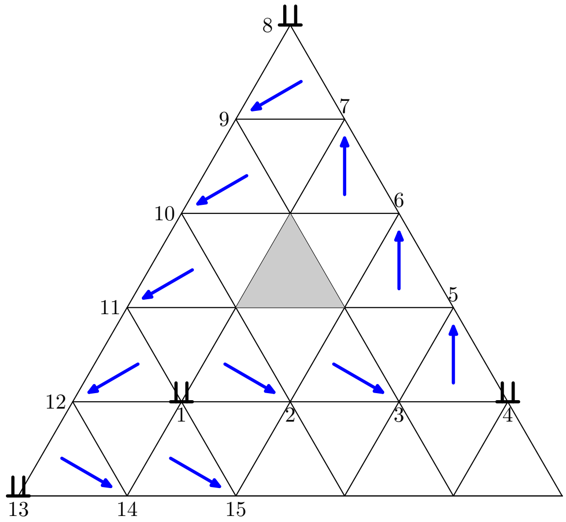


Figure 4.5: Illustration of the incremental construction of the random field.

By applying the above rules in the order illustrated by the figure below, one can progressively build the stationary space-time diagram of the PCA. Indeed the rules enlarge S in such a way that, at each step, the variables of S have the same distribution as the corresponding finite-dimensional marginal of the stationary space-time diagram. This is proved by Lemmas 4.1 and 4.6.

On Fig. 4.5, the labelling of the nodes corresponds to the step at which the corresponding variable is computed (after the three variables of the grey triangle). An arrow pointing to a variable means that it has been constructed according to the PCA of the direction of the arrow (first rule). The nodes labelled by $\amalg$ are the ones which have been constructed by independence (second rule).

In the next sections, we consider two types of extensions. First, PCA with an alphabet and neighbourhood of size 2 but having a Markov invariant measure. Second, PCA having a Bernoulli product invariant measure but with a general alphabet and neighbourhood.

4.3 Elementary PCA having Markov invariant measures

Markov measures are a natural extension of Benoulli product measures. In a nutshell, the tools of Sec. 4.1 can be extended to find conditions for having a Markov invariant measure, but the spatial properties presented in Sec. 4.2 do not remain.

Consider $a, b \in (0,1)$, and let us consider the Markov measure on $\{0,1\}^{\mathbb{Z}}$ of transition matrix

$$Q = \begin{pmatrix} 1-a & a \\ 1-b & b \end{pmatrix},$$

it is the measure ν_Q defined on cylinders by:

$$\forall x = x_m \cdots x_n, \qquad \nu_Q[x] = \pi_{x_m} \prod_{i=m}^{n-1} Q_{x_i, x_{i+1}},$$

where $\pi = (\pi_0, \pi_1)$ is such that $\pi Q = \pi$, $\pi_0 + \pi_1 = 1$, that is, $\pi_0 = (1-b)/(1-b+a)$ and $\pi_1 = a/(1-b+a)$.

The Markov measure ν_Q is space-stationary. If $a = b$, then $\nu_Q = \mu_a$, the Bernoulli product measure of parameter a.

Let us fix the PCA, that is, the parameters $(\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11})$ and assume that (4.1) holds. Let us fix the parameters a and b in $(0,1)$ (defining Q and π as above). We introduce the analogues of the functions defined in (4.2) and (4.3).

For $\alpha \in \{0,1\}$, define the function:

$$\begin{array}{rccl} g_\alpha : & [0,1] & \longrightarrow & (0,1) \\ & r & \longmapsto & (1-r)\,(1-a)\,\theta_{00}^\alpha + (1-r)\,a\,\theta_{01}^\alpha + r\,(1-b)\,\theta_{10}^\alpha + r\,b\,\theta_{11}^\alpha\,. \end{array} \tag{4.8}$$

In words, $g_\alpha(r)$ is the probability that $Y_0 = \alpha$ if the law of (X_0, X_1) is given by $\mathbb{P}(X_0 = x_0, X_1 = x_1) = r_{x_0}\; Q_{x_0,x_1}$ with $r_0 = 1-r$ and $r_1 = r$. With condition (4.1) on the parameters, we have $g_\alpha(r) \in (0,1)$ for all r. Observe also that: $g_0(r) + g_1(r) = 1$.

For $\alpha \in \{0,1\}$, we also define the function:

$$\begin{array}{rccl} h_\alpha : & [0,1] & \longrightarrow & [0,1] \\ & r & \longmapsto & \big[(1-r)\,a\,\theta_{01}^\alpha + r\,b\,\theta_{11}^\alpha\big] g_\alpha(r)^{-1}\,. \end{array} \tag{4.9}$$

In words, $h_\alpha(r)$ is the probability to have $X_1 = 1$ conditionally to $Y_0 = \alpha$ if (X_0, X_1) is distributed according to the above law.

Proposition 4.9. *Consider the Markov measure ν_Q and the PCA F as above. For any $\alpha_0 \cdots \alpha_{n-1} \in \mathcal{A}^n$, the probability of the cylinder $[\alpha_0 \cdots \alpha_n]$ under $\nu_Q F$ is given by:*

$$\nu_Q F[\alpha_0 \cdots \alpha_{n-1}] = g_{\alpha_0}(\pi_1) \prod_{i=1}^{n-1} g_{\alpha_i}(h_{\alpha_{i-1}}(h_{\alpha_{i-2}}(\ldots h_{\alpha_0}(\pi_1)\ldots)))\,.$$

Using Prop. 4.9, we obtain sufficient conditions for having a Markov invariant measure. This provides a new proof of a result already mentioned in different works [BGM69, DKT90, Ver76].

Theorem 4.2. *Consider a PCA F such that: $\exists i,j,\ \theta_{ij} \in (0,1)$, that is, a PCA which is not a deterministic CA. The PCA F has an invariant Markov measure associated to $a, b \in (0,1)$ if we are in one of the three following cases:*

1. *The parameters satisfy:*

$$(\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11}) \in (0,1)^4,\ \ \theta_{00}\theta_{11}(1-\theta_{01})(1-\theta_{10}) = \theta_{01}\theta_{10}(1-\theta_{00})(1-\theta_{11}). \tag{4.10}$$

In which case, a and b are the unique solutions in $(0,1)$ of the equations:

$$b(1-\theta_{11}) = (1-a)\theta_{00}, \qquad a(1-b)\theta_{01}\theta_{10} = b(1-a)\theta_{00}\theta_{11}.$$

2. *The parameters satisfy:*

$$\begin{array}{rllll} & \theta_{00} = 1, & \theta_{01} \in (0,1], & \theta_{10} = 1, & \theta_{11} \in (0,1) \\ \text{or} & \theta_{00} = 1, & \theta_{01} = 1, & \theta_{10} \in (0,1], & \theta_{11} \in (0,1). \end{array}$$

In which case, a and b are the unique solutions in $(0,1)$ of the equations:

$$b(1-\theta_{11}) = (1-a), \qquad a(1-b)\theta_{01}\theta_{10} = b(1-a)\theta_{11}.$$

3. *The parameters satisfy:*

$$\begin{array}{rllll} & \theta_{00} \in (0,1), & \theta_{01} = 0, & \theta_{10} \in [0,1), & \theta_{11} = 0 \\ \text{or} & \theta_{00} \in (0,1), & \theta_{01} \in [0,1), & \theta_{10} = 0, & \theta_{11} = 0. \end{array}$$

In which case, a and b are the unique solutions in $(0,1)$ of the equations:

$$b = (1-a)\theta_{00}, \qquad a(1-b)(1-\theta_{01})(1-\theta_{10}) = b(1-a)(1-\theta_{00}). \tag{4.11}$$

Let us point out that if $a \notin \{b, 1-b\}$, the condition (4.10) is also necessary.

Proof. We treat the case $\big[(\theta_{00},\theta_{01}) \neq (1,1), (\theta_{10},\theta_{11}) \neq (0,0)\big]$ (observe that Prop. 4.9 holds). The case $\big[(\theta_{00},\theta_{10}) \neq (1,1), (\theta_{01},\theta_{11}) \neq (0,0)\big]$ can be treated by reversing the space-direction.

Let us assume that the following conditions are satisfied:

1. for $\alpha \in \{0,1\}$, $g_\alpha(\pi_1) = \pi_\alpha$;
2. for $\alpha \in \{0,1\}$, there exists $c_\alpha \in [0,1]$ such that: $\forall r,\ h_\alpha(r) = c_\alpha$;
3. for $\alpha, \beta \in \{0,1\}$, $g_\beta(c_\alpha) = Q_{\alpha,\beta}$.

Then, by a direct application of Prop. 4.9, the measure ν_Q is invariant. When are these conditions fulfilled?

For $\alpha = 1$, condition 2 tells us that there exists $c_1 \in [0,1]$ such that for any $r \in [0,1]$,

$$(1-r)\ a\ \theta_{01} + r\ b\ \theta_{11} = c_1\big((1-r)\ (1-a)\ \theta_{00} + (1-r)\ a\ \theta_{01} + r\ (1-b)\ \theta_{10} + r\ b\ \theta_{11}\big).$$

This is the case if and only if:

$$a\ \theta_{01} = c_1((1-a)\ \theta_{00} + a\ \theta_{01}), \qquad b\ \theta_{11} = c_1((1-b)\ \theta_{10} + b\ \theta_{11})\ .$$

Thus, condition 2 for $\alpha = 1$ is equivalent to:

$$a\ (1-b)\ \theta_{01}\ \theta_{10} = (1-a)\ b\ \theta_{00}\ \theta_{11}\ . \tag{4.12}$$

In the same way, condition 2 for $\alpha = 0$ is equivalent to:

$$a\ (1-b)\ (1-\theta_{01})\ (1-\theta_{10}) = (1-a)\ b\ (1-\theta_{00})\ (1-\theta_{11})\ . \tag{4.13}$$

Eliminating a and b in (4.12) and (4.13), we obtain the relation (4.10) for the parameters of the PCA.

Conversely, let us assume that relation (4.10) holds. We will prove that there exist $a, b \in (0,1)$ such that the three above conditions are satisfied.

First observe that (4.12) holds if and only if (4.13) holds. So, we have a first relation to be satisfied by the parameters $a, b \in (0,1)$ which is (4.12). Under this relation, condition 2 is satisfied with:

$$c_0 = \frac{a\ (1-\theta_{01})}{(1-a)\ (1-\theta_{00}) + a\ (1-\theta_{01})} = \frac{b\ (1-\theta_{11})}{(1-b)\ (1-\theta_{10}) + b\ (1-\theta_{11})}\ , \tag{4.14}$$

and

$$c_1 = \frac{a\ \theta_{01}}{(1-a)\ \theta_{00} + a\ \theta_{01}} = \frac{b\ \theta_{11}}{(1-b)\ \theta_{10} + b\ \theta_{11}}\ . \tag{4.15}$$

Now consider condition 3 for $\alpha = \beta = 1$. Simplifying using (4.15), we obtain:

$$g_1(c_1) = Q_{11} = b \iff (1-a)\ \theta_{00} = b\ (1-\theta_{11})\ . \tag{4.16}$$

Condition 3 for other values of α and β provides the same relation after simplification.

Let us show that if equations (4.12) and (4.16) are satisfied, then the PCA also fulfills condition 1. Is is sufficient to prove that $g_1(\pi_1) = \pi_1$. Expanding both sides of (4.13) and simplifying using (4.12), we obtain:

$$a(1-b)\ (1-\theta_{01}-\theta_{10}) = (1-a)b\ (1-\theta_{00}-\theta_{11})\ . \tag{4.17}$$

Applying the definition (4.8), we have:

$$g_1(\pi_1) = \frac{1}{1-b+a}\Big((1-b)(1-a)\,\theta_{00} + (1-b)a\,\theta_{01} + a(1-b)\,\theta_{10} + ab\,\theta_{11}\Big)\,.$$

Using (4.17), we can replace $a(1-b)(\theta_{01}+\theta_{10})$ by $a(1-b)-(1-a)b\,(1-\theta_{00}-\theta_{11})$. With (4.16), we finally obtain $g_1(\pi_1) = a/(1-b+a) = \pi_1$.

Now, observe that the system:

$$\begin{cases} (1-b)\,a\,\theta_{01}\,\theta_{10} = b\,(1-a)\,\theta_{00}\,\theta_{11} \\ (1-a)\,\theta_{00} = b\,(1-\theta_{11}) \end{cases} \tag{4.18}$$

has a unique solution $(a,b)\in(0,1)^2$. Let Q be the matrix associated with (a,b). Since the three above conditions are satisfied, the Markov measure ν_Q is invariant by the PCA. □

In the Markov case, unlike the Bernoulli case, there is no simple description of the law of other lines in the stationary space-time diagram. Nevertheless, the stationary space-time diagram has a different but still remarkable property: if $\theta_{01}=\theta_{10}$, it is *time-reversible*, meaning it has the same distribution if we reverse the direction of time [Vas78]. This is closely related to the results of Sec. 1.4, and can be proved by considering the class of PCA obtained in Ex. 1.1 when $c_3 = 1$.

Bernoulli product measures are special cases of Markov measures. Therefore it is natural to ask whether all the cases covered by Th. 4.1 are retrieved in (4.10). The answer is no. Indeed, the measure ν_Q is a Bernoulli product measure if and only if $a=b$. Simplifying in (4.18) and (4.10), we obtain:

$$\big[\theta_{00}=\theta_{01},\theta_{11}=\theta_{10}\big] \quad \text{or} \quad \big[\theta_{00}=\theta_{10},\theta_{11}=\theta_{01}\big]\,.$$

The corresponding PCA have a neighbourhood of size 1. This is far from exhausting the PCA with a Bernoulli product measure.

Finite set of cells. It is also interesting to draw a parallel between the result of Th. 4.2 and Prop. 4.6 of Bousquet-Mélou [BM98]. In this last article, the author studies PCA of alphabet $\mathcal{A}=\{0,1\}$ and neighbourhood $\mathcal{N}=\{0,1\}$, but defined on a finite ring of size N (periodic boundary conditions: $X_N = X_0$), and proves that the invariant measure has a Markov form if the parameters satisfy the same relation (4.10) as in the infinite case. The expression of the measure is then given by:

$$\mathbb{P}(X_0=x_0, X_1=x_1,\dots,X_{N-1}=x_{N-1}) = \frac{1}{Z}\prod_{i=0}^{N-1} Q_{x_i,x_{i+1}},$$

where Z is a normalising constant, and where the coefficients a and b defining the matrix Q are the solution of the same system (4.18) as in the infinite case.

For a PCA satisfying condition (4.10), we have a Markov invariant measure both on a finite ring and on $\mathbb{Z}$. This is not the case for Bernoulli product measures: except when the actual neighbourhood is of size 1, PCA satisfying the conditions of Th. 4.1 do not have a product form invariant measure on finite rings.

Example 4.4. Consider for instance the PCA of transition function $f(x,y) = (3/4)\,\delta_{x+y \mod 2} + (1/4)\,\delta_{x+y+1 \mod 2}$ (Ex. 4.2), on the ring of size 4. Its invariant measure μ is different from the uniform measure:

$$\mu(0000)=573/8192,\ \mu(0001)=963/16384,\ \mu(0011)=33/512,$$
$$\mu(0101)=69/1024,\ \mu(0111)=957/16384,\ \mu(1111)=563/8192\,.$$

4.4 General alphabet and neighbourhood

In this section, the neighbourhood is $\mathcal{N} = \{0, \ldots, \ell\}$ and the alphabet is $\mathcal{A} = \{0, \ldots, n\}$. For $p = (p_0, \ldots, p_n)$ such that $p_0 + \ldots + p_n = 1$, we still denote by μ_p the corresponding Bernoulli product measure on $\mathcal{A}^{\mathbb{Z}}$.

For convenience, we introduce the following notations: $\forall x_0, \ldots, x_\ell \in \mathcal{A},\ \forall k \in \mathcal{A}$,

$$\theta^k_{x_0 \cdots x_\ell} = f(x_0, \ldots, x_\ell)(k)\ .$$

We define new functions g_k and h_k, which generalise the ones in (4.2) and (4.3). These new functions g_k and h_k are not functions of a single variable, but of probability measures on $\mathcal{A}^\ell$. Assume that:

$$\forall k \in \mathcal{A},\ \forall x_0 \cdots x_{\ell-1} \in \mathcal{A}^\ell,\ \exists i \in \mathcal{A}, \qquad \theta^k_{x_0 \cdots x_{\ell-1} i} > 0\ . \tag{4.19}$$

Let us define:

$$\begin{aligned} g_k : \mathcal{M}(\mathcal{A}^\ell) &\longrightarrow (0,1), \\ \mathcal{D} &\longmapsto \text{the probability that } Y_0 = k \text{ if } (X_0, \ldots, X_\ell) \sim \mathcal{D} \otimes \mathcal{B}_p\ , \end{aligned}$$

$$\begin{aligned} h_k : \mathcal{M}(\mathcal{A}^\ell) &\longrightarrow \mathcal{M}(\mathcal{A}^\ell), \\ \mathcal{D} &\longmapsto \text{the distribution of } (X_1, \ldots, X_\ell) \text{ conditionally to } Y_0 = k \\ &\qquad \text{if } (X_0, \ldots, X_\ell) \sim \mathcal{D} \otimes \mathcal{B}_p\ . \end{aligned}$$

We have the following analogue of Prop. 4.1.

Proposition 4.10. *Consider a PCA satisfying (4.19). Consider $p = (p_i)_{i \in \mathcal{A}}$ with $p_i > 0$ for all $i \in \mathcal{A}$. For $\alpha_0 \cdots \alpha_{n-1} \in \mathcal{A}^n$, the probability of the cylinder $[\alpha_0 \cdots \alpha_{n-1}]$ under $\mu_p F$ is given by:*

$$\mu_p F[\alpha_0 \cdots \alpha_{n-1}] = g_{\alpha_0}(\mathcal{B}_p^{\otimes \ell}) \prod_{i=1}^{n-1} g_{\alpha_i}(h_{\alpha_{i-1}}(h_{\alpha_{i-2}}(\ldots h_{\alpha_0}(\mathcal{B}_p^{\otimes \ell}) \ldots)))\ .$$

By reversing the space-direction, we get an analogue of Prop. 4.10 under the symmetric condition: $\forall k \in \mathcal{A}, \forall x_0 \cdots x_{\ell-1} \in \mathcal{A}^\ell, \exists i \in \mathcal{A},\ \theta^k_{i x_0 \cdots x_{\ell-1}} > 0$.

Applying Prop. 4.10, we obtain the following result, that had already appeared in a more complicated setting [Vas78].

Theorem 4.3. *Consider $p = (p_i)_{i \in \mathcal{A}}$ with $p_i > 0$ for all $i \in \mathcal{A}$. The measure μ_p is an invariant measure of the PCA F if one of the two following conditions is satisfied:*

$$\forall x_0, \ldots, x_{\ell-1} \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{i \in \mathcal{A}} p_i\, \theta^k_{x_0 \cdots x_{\ell-1} i} = p_k, \tag{4.20}$$

$$\forall x_0, \ldots, x_{\ell-1} \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{i \in \mathcal{A}} p_i\, \theta^k_{i x_0 \cdots x_{\ell-1}} = p_k. \tag{4.21}$$

Proof. Let us assume that F satisfies condition (4.20). Then, the function g_k is constant. Indeed,

$$g_k(\mathcal{D}) = \sum_{i \in \mathcal{A}, x_0 \cdots x_{\ell-1} \in \mathcal{A}^\ell} \mathcal{D}(x_0, \ldots, x_{\ell-1})\ p_i\ \theta^k_{x_0 \cdots x_{\ell-1} i} = p_k\ .$$

By Prop. 4.10, we obtain that $\mu_p F = \mu_p$.

Now, like in the proof of Th. 4.1, we can reverse the space direction and define a new PCA $\widetilde{F}$. The PCA F satisfies condition (4.21) if and only if the PCA $\widetilde{F}$ satisfies condition (4.20). Therefore, if F satisfies condition (4.21), then we have $\mu_p \widetilde{F} = \mu_p$, which implies in turn that $\mu_p F = \mu_p$. $\square$

As opposed to Th. 4.1, the conditions in Th. 4.3 are sufficient but not necessary. To illustrate this fact, the simplest examples are provided by PCA that do not depend on all the elements of their neighbourhood. Consider for instance the PCA of alphabet $\mathcal{A} = \{0,1\}$ and neighbourhood $\mathcal{N} = \{0,1,2\}$, defined, for some $a,b \in (0,1)$, by: $\forall u,v \in \mathcal{A}, \theta^1_{u0v} = a$, $\theta^1_{u1v} = b$. This PCA has a Bernoulli invariant measure, but if $a \neq b$, it satisfies neither condition (4.20), nor condition (4.21).

Let us state a related result, which extends Prop. 4.2, and completes Th. 4.3. (For the relevance of this result, see the discussion following Prop. 4.2.)

Proposition 4.11 ([Vas78]). *Consider a positive rates PCA F satisfying condition (4.20) or (4.21), for some $p = (p_i)_{i\in\mathcal{A}}$, $p_i > 0$ for all $i \in \mathcal{A}$. Then F is ergodic, that is, μ_p is the unique invariant measure of F and for all initial measure μ, the sequence $(\mu F^n)_{n\geq 0}$ converges weakly to μ_p.*

Condition (4.20) implies that the variables $X_0, \dots, X_{\ell-1}, Y_0$ are mutually independent, since for any $v \in \{0,1\}^\ell$ and $\alpha \in \{0,1\}$, we have $\mathbb{P}((X_0,\dots,X_{\ell-1}) = v, Y_0 = \alpha) = \mu_p[v] \sum_{i\in\mathcal{A}} p_i\, \theta^\alpha_{vi} = \mu_p[v]\mu_p[\alpha]$. Similarly, condition (4.21) implies that the variables $X_1, \dots,$ X_ℓ, Y_0 are mutually independent.

The next lemma is a generalisation of Lemma 4.6.

Lemma 4.7. *Under conditions (4.20) and (4.21), the variable X^0_0 is independent of $(X^n_k)_{k\in\mathbb{Z},n\in\mathbb{N}\setminus\{0\}}$.*

Proof. Set $X = X^0$ and $Y = X^1$. Like in Lemma 4.6, it is sufficient to prove that X_0 is independent of $Y = (Y_k)_{k\in\mathbb{Z}}$. Let us fix some $a,b \in \mathbb{Z}, (a < 0 < b)$, and prove that X_0 is independent of $(Y_a, Y_{a+1}, \dots, Y_b)$. We have:

$$S = \mathbb{P}\big(X_0 = x_0, (Y_i)_{a\leq i\leq b} = (y_i)_{a\leq i\leq b}\big)$$

$$= \sum_{\substack{x_i\in\mathcal{A}\\ i\in\{a,a+1,\dots,b+\ell\}\setminus\{0\}}} \mathbb{P}\big((X_i)_{a\leq i\leq b+\ell} = (x_i)_{a\leq i\leq b+\ell}, (Y_i)_{a\leq i\leq b} = (y_i)_{a\leq i\leq b}\big)\,.$$

Furthermore

$$\mathbb{P}\big((X_i)_{a\leq i\leq b+\ell} = (x_i)_{a\leq i\leq b+\ell}, (Y_i)_{a\leq i\leq b} = (y_i)_{a\leq i\leq b}\big)$$
$$= \mu_p[x_0] \prod_{i=a}^{-1} \mu_p[x_i]\, \theta^{y_i}_{x_i\cdots x_{i+\ell}} \prod_{j=\ell}^{b+\ell} \mu_p[x_j]\, \theta^{y_{j-\ell}}_{x_{j-\ell}\cdots x_j} \prod_{k=1}^{\ell-1} \mu_p[x_k]\,.$$

If we compute the sum S in the order: $x_a, \dots, x_{-1}$ first (simplifications using condition (4.20)) then $x_{b+\ell}, x_{b+\ell-1}, \dots, x_\ell$ (simplifications using condition (4.21)), and finally $x_1, \dots, x_{\ell-1}$, we obtain eventually: $S = \mu_p[x_0] \prod_{i=a}^b \mu_p[y_i]$. □

Corollary 4.3. *If both conditions (4.20) and (4.21) are satisfied, then every line of the stationary space-time diagram consists of i.i.d. random variables. In particular, any two different random variables are independent.*

If the neighbourhood is $\mathcal{N} = \{0,1\}$, under conditions (4.20) or (4.21), the spatial properties of Sec. 4.2 remain for a general alphabet (existence of transversal PCA, properties of triangles,...). These two conditions can then be rewritten:

$$\forall i \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{j\in\mathcal{A}} p_j\, \theta^k_{ij} = p_k,$$
$$\forall j \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{i\in\mathcal{A}} p_i\, \theta^k_{ij} = p_k,$$

and the transversal PCA are defined respectively by the parameters

$$\tilde\theta^k_{ij} = \frac{p_k}{p_i}\theta^i_{jk} \qquad \text{and} \qquad \tilde\theta^k_{ij} = \frac{p_k}{p_j}\theta^j_{ki}.$$

For other neighbourhoods, there is no natural transversal PCA.

Chapter 5

Randomisation versus conservation in one-dimensional CA

Tout cela tourbillonnant, se chevauchant en désordre... Mais il connaît pour les avoir mille fois observées ces infimes particules en mouvement. Il les a isolées d'autres particules avec lesquelles elles avaient formé d'autres systèmes très différents, il les connaît bien. Maintenant elles montent, affleurent, elles forment sur le visage de son père un fin dépôt, une mince couche lisse qui lui donne un aspect figé, glacé.

– Nathalie Sarraute, *Le planétarium*

Contents

In this chapter, we focus on deterministic one-dimensional CA, and on the role played by their Bernoulli invariant measures.

We study the necessary and sufficient conditions for CA to have Bernoulli product invariant measures. These conditions can be described by a conservation law [KT12]. In particular, it appears that the fact, for a CA, to admit all Bernoulli measures as invariant measures is very restrictive: the CA fulfilling this property are exactly the ones that are both surjective and state-conserving. It remains true if we consider PCA, since the only PCA that admits every Bernoulli measures as invariant measures are deterministic ones.

We compare the known criterion on deterministic CA for having a Bernoulli invariant measure with the ones obtained in the previous chapter for PCA. When specialising to deterministic CA the sufficient conditions for having a Bernoulli product measure developed in Chap. 4, Sec. 4.4, a particular class of CA appears, namely permutative CA. The combinatorical structure of these CA gives them rich properties. We study their *rigidity* and their *randomisation* capacities. Informally, a CA is rigid if its only invariant measure satisfying some non-degeneracy condition (e.g. positive entropy) is the uniform measure. The randomisation is the property, for a CA, to converge (either simply, or in Cesàro mean) to the uniform measure from a large class of initial measures.

5.1 Bernoulli invariant measures and conservation laws

Recall that a deterministic cellular automaton is a PCA having a transition function f such that, for all $x \in \mathcal{A}^{\mathcal{N}}$, the probability measure $f(x)$ is concentrated on a single letter of the alphabet. Thus, the transition function of a one-dimensional CA can be described by a mapping $f : \mathcal{A}^{\mathcal{N}} \longrightarrow \mathcal{A}$, and the CA can be viewed as a deterministic mapping $F : \mathcal{A}^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$.

5.1.1 CA having Bernoulli invariant measures

Let us fix the alphabet $\mathcal{A}$. We recall that we denote by λ the uniform measure on $\mathcal{A}^{\mathbb{Z}}$, that is, the product measure of uniform measures on $\mathcal{A}$. Next proposition gives a complete characterisation of CA for which λ is an invariant measure.

Proposition 5.1 ([Hed69]). *Let F be a cellular automaton. We have:*

$$F \text{ is surjective} \quad \Longleftrightarrow \quad \lambda F = \lambda\,.$$

Let us present a recent result which refines Prop. 5.1. Given a finite and non-empty word $u \in \mathcal{A}^+$, let $u^{\mathbb{Z}} = \cdots uuu \cdots \in \mathcal{A}^{\mathbb{Z}}$ be a periodic bi-infinite word of period u (the starting position is indifferent). If $F : \mathcal{A}^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ is a CA, then $F(u^{\mathbb{Z}}) = v^{\mathbb{Z}}$ for some word v with $|v| = |u|$. For simplicity, we write $v = F(u)$.

Theorem 5.1 ([KT12]). *Consider a CA F on the alphabet $\mathcal{A}$. The Bernoulli product measure μ_p, $p = (p_i)_{i\in\mathcal{A}}$, $p_i > 0$ for all $i \in \mathcal{A}$, is invariant for F if and only if:*

$$(i)\ F \text{ is surjective} \qquad \text{and} \qquad (ii)\ \forall u \in \mathcal{A}^+,\ \sum_{i\in\mathcal{A}} |u|_i \log(p_i) = \sum_{i\in\mathcal{A}} |F(u)|_i \log(p_i)\,.$$

This theorem is a *conservation law*. It claims that a CA preserves a Bernoulli measure μ_p if and only if, when attributing a weight $\log p_i$ to the letter $i \in \mathcal{A}$, the total weight

$$\sum_{i\in\mathcal{A}} |u|_i \log(p_i)$$

of the periodic configuration $u^{\mathbb{Z}}$ is preserved by the CA.

If a CA has an invariant Bernoulli product measure μ_p (with $p_i > 0$ for all $i \in \mathcal{A}$), then it is surjective, so that by Th. 5.1, the uniform measure is also invariant.

5.1.2 PCA having all Bernoulli measures as invariant measures

Definition 5.1. A cellular automaton F is *state-conserving* if:

$$\forall u \in \mathcal{A}^+, \forall i \in \mathcal{A},\ |u|_i = |F(u)|_i.$$

As a consequence of Th. 5.1, a surjective and state-conserving CA admits all Bernoulli product measures μ_p as invariant measures. We also have a converse proposition.

Proposition 5.2. *The two following properties are equivalent.*

(i) F is a PCA such that for every Bernoulli measure μ_p, we have $\mu_p F = \mu_p$.

(ii) F is a surjective and state-conserving CA.

Proof. The implication: $(ii) \implies (i)$ follows from Th. 5.1. Let us consider a PCA satisfying (i), and consider two words $u, v \in \mathcal{A}^n$ (with n larger than the size of the neighbourhood) such that on a finite ring of size n, after one iteration of the PCA, there is a positive probability $\theta_{u\to v}$ to reach v from u. On the line $\mathbb{Z}$, for any $k \in \mathbb{N}$, starting from the word u^{k+2}, there is a probability $\theta_{u\to v}^k$ to obtain the word v^k after one iteration.

$$\begin{array}{ccccccccc} \dots & \dots & v & v & v & v & \dots & \dots \\ \dots & u & u & u & u & u & u & \dots \end{array}$$

Since μ_p is an invariant measure, we have $\mu_p[v^k] \geq \mu_p[u^{k+2}] \cdot \theta_{u\to v}^k$, that is $\mu_p[v]^k \geq \mu_p[u]^{k+2} \cdot \theta_{u\to v}^k$. This is true for any $k \in \mathbb{N}$. By raising to the power $1/k$ on each side, when k tends toward infinity, we obtain $\mu_p(v) \geq \mu_p(u) \cdot \theta_{u\to v}$, that is:

$$\frac{\mu_p[v]}{\mu_p[u]} \geq \theta_{u\to v}.$$

We want this inequality to be true for any value of p. But:

$$\frac{\mu_p[v]}{\mu_p[u]} = \frac{\prod_{i\in\mathcal{A}} p_i^{|v|_i}}{\prod_{i\in\mathcal{A}} p_i^{|u|_i}} = \prod_{i\in\mathcal{A}} p_i^{|v|_i-|u|_i},$$

so that if $|u|_i < |v|_i$ for some $i \in \mathcal{A}$, we obtain a contradiction for $p_i \to 0$, and if $|u|_i > |v|_i$, for $p_i \to 1$.

So, for any $u, v \in \mathcal{A}^n$, if there is a positive probability to go from u to v on the ring of size n, then $|u|_i = |v|_i$ for all $i \in \mathcal{A}$. Let us assume that the PCA is not deterministic. Then, there exists a value of the neighbourhood making possible a transition from some $i \in \mathcal{A}$ to some $j \in \mathcal{A}$ or some $k \neq j$, both with positive probability. It means that if this value of the neighbourhood appears on a finite ring, then there are at least two words with different numbers of j that can be reached with a positive probability. They cannot have both the same number of j as the initial configuration, so that we get a contradiction. Consequently, the PCA is in fact a deterministic CA. □

There are non-trivial examples of surjective and state-conserving CA. The following example is suggested by García-Ramos [GR12].

Example 5.1. Let $A = 100010000$ and $B = 100100000$. These two blocks are non-overlapping. We define a CA in the following way: if there are two consecutive blocks of A or B, then the one at the right is changed into A if they are the same, and into B if there are different; in all other cases the state of the cells are unchanged. This CA is surjective and state-conserving.

$$\begin{array}{cc} & A \\ & \uparrow \\ A & A \\ B & B \end{array} \qquad \begin{array}{cc} & B \\ & \uparrow \\ A & B \\ B & A \end{array}$$

We would like to know if there are surjective and state-conserving CA whose rules are "less constrained" than the one of the CA described above. In order to give a precise meaning of this, we introduce the following definition.

Definition 5.2. A configuration $x \in \mathcal{X}$ is an equicontinuity point of F if:

$$\forall \varepsilon > 0, \exists \delta > 0, d(x, y) < \delta \implies [\forall n \in \mathbb{N}, d(F^n(x), F^n(y)) < \varepsilon].$$

It is an equicontinuity point of F in the direction $(p, q) \in \mathbb{Z} \times \mathbb{N} \setminus \{0\}$ if it is an equicontinuity point of $\sigma^p F^q$.

The CA of Ex. 5.1 has equicontinuity points, for example all the points that do not contain a block A or a block B, which is the case in particular of the point $0^{\mathbb{Z}}$. From any configuration close to such a point (for example, a configuration having about ten 0's around the origin), the cells located around the origin behave endlessly in a trivial way, and there is no interaction between the left part and the right part of the initial configuration.

It is an open problem to know if any surjective and state-conserving CA has at least one direction with equicontinuity points.

In our attempt to explore surjective and state-conserving CA, we have proved the following proposition.

Proposition 5.3. *Let F be a CA of alphabet $\mathcal{A} = \{0, \ldots, n\}$ and of neighbourhood $\mathcal{N} = \{\ell, \ldots, r-1, r\}$ for some $\ell < r$, and set $L = r - \ell$. We have the following equivalence.*

(i) F is a surjective and state-conserving CA.

(ii) For any word $u \in \mathcal{A}^s$, and any $k_0, \ldots, k_n$ such that $k_0 + \ldots + k_n = L$ there are exactly $\binom{L}{k_0,\ldots,k_n}$ words of length $s + L$ with a number $|u|_i + k_i$ of i, for $0 \leq i \leq n$, that are mapped to u.

Proof. Let F be a surjective and state-conserving CA, and let $u \in \mathcal{A}^s$. We have: $\mu_p[u] = \prod_{i\in\mathcal{A}} p_i^{|u|_i}$, and since all the Bernoulli measures are invariant by Prop. 5.2, it is also equal to:

$$\mu_p F[u] = \sum_{t_0+\ldots+t_n=s+L} \alpha_t \prod_{i\in\mathcal{A}} p_i^{t_i},$$

where α_t is the number of words v of length $s + L$ such that $|v|_i = t_i$ for $0 \leq i \leq n$, that are mapped to u. Thus, for any probability vector p, we have:

$$\prod_{i\in\mathcal{A}} p_i^{|u|_i} = \sum_{t_0+\ldots+t_n=s+L} \alpha_t \prod_{i\in\mathcal{A}} p_i^{t_i}.$$

If we specialise this equality to the identity CA, we recover the following combinatorial formula:

$$\prod_{i\in\mathcal{A}} p_i^{|u|_i} = \sum_{k_0+\ldots+k_n=L} \binom{L}{k_0,\ldots,k_n} \prod_{i\in\mathcal{A}} p_i^{|u|_i+k_i}.$$

Since this is true for any probability vector p, we can identify the coefficients. Thus, $\alpha_t = \binom{L}{k_0-|u|_0,\ldots,k_n-|u|_n}$ if k is such that $k_i \geq |u|_i$ for all $0 \leq i \leq n$, and $\alpha_t = 0$ otherwise. Equivalently, for any word $u \in \mathcal{A}^n$, and any $k_0, \ldots, k_n$ such that $k_0 + \ldots + k_n = L$ there are exactly $\binom{L}{k_0,\ldots,k_n}$ words of length $s + L$ with a number $|u|_i + k_i$ of i, for $0 \leq i \leq n$, that are mapped to u.

Conversely, if (ii) is satisfied, then any Bernoulli measure μ_p is invariant, which implies (i). □

5.1.3 Permutative CA

The results in Sec. 4.1, and 4.4 give conditions for a PCA to admit invariant Bernoulli product measures. The above results, Sec. 5.1.1, give conditions for a CA to admit invariant Bernoulli product measures. The natural question is whether we obtain the latter conditions by specialising the former ones.

Definition 5.3. A CA of transition function $f : \mathcal{A}^{\mathcal{N}} \longrightarrow \mathcal{A}$, where the neighbourhood is of the form $\mathcal{N} = \{\ell, \ldots, r-1, r\}$ for some $\ell < r$, is *left-permutative* (resp. *right-permutative*) if, for all $w = w_\ell \cdots w_{r-1} \in \mathcal{A}^{r-\ell}$, the mapping from $\mathcal{A}$ to $\mathcal{A}$ defined by: $a \longmapsto f(aw)$ (resp. $a \longmapsto f(wa)$), is bijective. A CA is *permutative* if it is either left or right-permutative. It is *bipermutative* if it is both left and right-permutative.

Example 5.2. If the alphabet is $\mathcal{A} = \mathbb{Z}_n$, and if $a, b, c \in \mathbb{Z}_n$, then the CA of local rule given by $f(x, y) = ax + by + c$ is called an *affine CA*. For a (resp. b) invertible in $\mathbb{Z}_n$, the CA is left-permutative (resp. right-permutative).

Let $F : \mathcal{A}^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ be a permutative CA. The existence of the bijections, see Def. 5.3, has two direct consequences: (i) F is surjective; (ii) the uniform measure is invariant: $\lambda F = \lambda$. In fact, these last two properties are equivalent, by Prop. 5.1.

Recall that the conditions (4.20) or (4.21) of Th. 4.3, see below, are sufficient for the Bernoulli product measure μ_p (with $\forall k \in \mathcal{A}, p_k > 0$) to be invariant for the PCA F.

$$\forall x_0, \ldots, x_{\ell-1} \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{i \in \mathcal{A}} p_i\, \theta^k_{x_0 \cdots x_{\ell-1} i} = p_k \quad (4.20)$$
$$\forall x_0, \ldots, x_{\ell-1} \in \mathcal{A}, \forall k \in \mathcal{A}, \quad \textstyle\sum_{i \in \mathcal{A}} p_i\, \theta^k_{i x_0 \cdots x_{\ell-1}} = p_k \quad (4.21)$$

Let us specialise these conditions to CA, that is, let us assume that all the coefficients $\theta^k_{x_0 \ldots x_{\ell-1} i}$ are equal to 0 or 1.

Lemma 5.1. *A cellular automaton satisfies condition (4.20) for some $p = (p_k)_{k \in \mathcal{A}}$ with $\forall k \in \mathcal{A}, p_k > 0$, resp. condition (4.21), if and only if it is right-permutative, resp. left-permutative.*

Proof. Consider a CA of transition function f satisfying condition (4.20) for some $p = (p_k)_{k \in \mathcal{A}}$ with $\forall k \in \mathcal{A}, p_k > 0$. The coefficients $\theta^k_{x_0 \cdots x_{\ell-1} i}$ can only be equal to 0 or to 1, the CA being deterministic. Let us fix $x_0, \ldots, x_{\ell-1} \in \mathcal{A}$. For any $k \in \mathcal{A}$, we have $p_k > 0$, so that there exists at least one letter $i \in \mathcal{A}$ for which $\theta^k_{x_0 \cdots x_{\ell-1} i} = 1$. The mapping from $\mathcal{A}$ to $\mathcal{A}$ defined by: $a \mapsto f(x_0 \cdots x_{\ell-1} a)$ is surjective and therefore bijective. □

To summarise, we recover the permutative CA. On the other hand, the surjective but non-permutative CA (like Ex. 5.1) are not captured by the sufficient conditions of Th. 4.3.

Let the neighbourhood be $\mathcal{N} = \{0, 1\}$ and consider a general alphabet $\mathcal{A}$. For a left-permutative CA (resp. right-permutative), the transversal CA, see Sec. 4.1.3 and also the last paragraph of Sec. 4.4, is right-permutative (resp. left-permutative), and explicitly computable. Moreover, it is well-defined even if the space-time diagram is not assumed to be stationary. We recover here a folk result.

In the special case $\mathcal{N} = \{0, 1\}$ and $\mathcal{A} = \{0, 1\}$, one can check by hand that all the surjective CA are permutative. So in this case, we recover all the surjective CA. This is consistent with the fact that in this case, the conditions of Th. 4.3 are necessary and sufficient (see Th. 4.1). In fact, Kari, Salo, and Törmä have proved in 2013 that when the neighbourhood is $\mathcal{N} = \{0, 1\}$ and $|\mathcal{A}|$ is a prime number, all surjective CA are permutative [KST13].

Remark. Condition (4.20) can be interpreted as "*being right-permutative in expectation*" for a PCA. And similarly, condition (4.21) amounts to "*being left-permutative in expectation*".

There are permutative CA that have other Bernoulli invariant measures μ_p than the uniform measure, but in that case, they admit a power of the shift as a non-trivial factor. Precisely, we have the following result.

Proposition 5.4. *Let F be a left-permutative CA of alphabet $\mathcal{A} = \{0, \ldots, n\}$ and neighbourhood $\mathcal{N} = \{\ell, \ldots, r\}$, with $L = r - \ell$. For any word $w \in \mathcal{A}^L$, we denote by τ_w the permutation of $\mathcal{A}$ such that for any $a \in \mathcal{A}$, $f(aw) = \tau_w(a)$. Let p be a probability measure on $\mathcal{A}$. We set $S_i = \{j \in \mathcal{A}; p_j = p_i\}$.*

If $\mu_p F = \mu_p$, then for any $w \in \mathcal{A}^L$, and any $i \in \mathcal{A}$, $\tau_w(S_i) = S_i$. In particular,

- *if p is such that $0 < p_0 < p_1 < \ldots < p_n$, it implies that $F = \sigma^r$.*

- *if F is bipermutative, its only Bernoulli invariant measure of full-support is the uniform measure.*

Proof. If a Bernoulli measure μ_p of full support is invariant under F, then for any $k \in \mathcal{A}$, we have:

$$p_k = \sum_{w \in \mathcal{A}^L} p_{\tau_w^{-1}(k)} \, p_{w_1} \dots p_{w_L}.$$

Let us assume without loss of generality that $0 < p_0 \leq p_1 \leq \dots \leq p_n$. Then,

$$p_0 = \sum_{w \in \mathcal{A}^L} p_{\tau_w^{-1}(0)} \, p_{w_1} \dots p_{w_L}$$

$$\geq \sum_{w \in \mathcal{A}^L} p_0 \, p_{w_1} \dots p_{w_L} = p_0.$$

It follows that for any $w \in \mathcal{A}^L$, $p_{\tau_w^{-1}(0)} = p_0$. With the setting $S_0 = \{k \in \mathcal{A}\,; p_k = p_0\}$, we obtain that for any $w \in \mathcal{A}^L$, $\tau_w(S_0) = S_0$. Consequently, if $k \notin S_0, \tau_w(k) \notin S_0$. Iterating the same argument, one can prove that for any $k \in \mathcal{A}$, and any $w \in \mathcal{A}^L$, $p_{\tau_w^{-1}(k)} = p_k$. Thus, all the permutations τ_w preserve the subsets of letters of $\mathcal{A}$ having the same weight under the Bernoulli measure.

In particular, we obtain that if a Bernoulli measure with parameters $0 < p_0 < p_1 < \dots < p_n$ is an invariant measure of a left-permutative CA, then this CA is equal to σ^r.

Let us now assume that the CA is bipermutative. There cannot be a (non-empty) subset $T \subsetneq \mathcal{A}$ such that for any $w \in \mathcal{A}^L$, $\tau_w(T) = T$, since in that case, for any word $v \in \mathcal{A}^L$ beginning with a letter in T, and any $a \in \mathcal{A}$, we would have $f(va) \in T$, which would be in contradiction with the right-permutativity. $\square$

Example 5.3. We define below a permutative CA of alphabet $\mathcal{A} = \{0, 1, 2, 3\}$ and neighbourhood $\mathcal{N} = \{0, 1\}$. The table gives the value of $f(x, y)$ for $x, y \in \mathcal{A}$.

	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	0	1	2	3
3	1	0	3	2

For any parameter $s \in [0, 1/2]$, the Bernoulli measure μ_p of parameter $p = (s, t, s, t)$ is an invariant measure, where t is defined by $t = 1/2 - s$.

5.2 Rigidity and randomisation

In Sec. 4.1.2, we have characterised PCA of alphabet $\mathcal{A} = \{0, 1\}$ and neighbourhood $\mathcal{N} = \{0, 1\}$ having Bernoulli invariant measures. As mentioned in Prop. 4.2, if they have positive rates, these PCA are ergodic. Surjective CA have the uniform Bernoulli measure as invariant measure, but they are non-ergodic, since they also have other invariant measures. Nevertheless, under some conditions, one can prove a *rigidity* result, which consists in proving that the only invariant measure satisfying some properties (so that the measures that are too "degenerated" are excluded) is the uniform measure [Sab10]. One can also look at *randomisation* results, that is, proving that from a large class of initial measures, the iterates of the CA converge to the uniform measure.

5.2.1 A first rigidity result: mixing criterion

Let us consider the sum CA, defined on the alphabet $\mathcal{A} = \mathbb{Z}_2$ by $F(x)_k = x_k + x_{k+1}$. The uniform Bernoulli measure $\lambda = \mu_{1/2}$ is an invariant measure of F, but F also has many other invariant measures, such as the measure concentrated on the configuration $\ldots 000 \ldots$, or measures stemming from different periodic orbits. We prove below a first *rigidity* result for that CA. Let us mention that Miyamoto has also obtained close results with a different approach [Miy79, Miy94].

We recall that a measure μ is a *shift-mixing measure* if for any cylinders $[u]$ and $[v]$,

$$\mu([u] \cap \sigma^{-n}[v]) \longrightarrow_{n\to\infty} \mu[u]\mu[v].$$

Proposition 5.5. *Let us consider the CA F of alphabet $\mathcal{A} = \mathbb{Z}_2$ defined by $F(x)_k = x_k + x_{k+1}$, and let μ be a shift-mixing measure with full support on $\mathcal{A}^{\mathbb{Z}}$. If μ is an invariant measure of F, then μ is equal to the uniform measure λ.*

Proof. For $n \in \mathbb{N}$, let us define the CA

$$G_n = F^{2^n}.$$

It is known that for any $n \in \mathbb{N}$, G_n satisfies the following scaling property:

$$\forall x \in \mathcal{A}^{\mathbb{Z}}, \forall k \in \mathbb{Z}, G_n(x)_k = x_k + x_{k+2^n}. \tag{5.1}$$

Let μ be a shift-mixing measure with full support on $\mathcal{A}^{\mathbb{Z}}$, that is invariant under the action of F. Then, for any $n \in \mathbb{N}$, μ is an invariant measure of G_n. Let us fix some $\ell \geq 1$ and consider the cylinders of length ℓ. Let $w \in \mathcal{A}^\ell$ be such that $\mu[w] = \min_{u\in\mathcal{A}^\ell} \mu[u]$, and assume that there exists some $w' \in \mathcal{A}^\ell$ such that $\mu[w'] > \mu[w]$. We set $\varepsilon = \mu[w'] - \mu[w] > 0$.

Since μ is mixing, for any $\alpha > 0$, there exists some $n \in \mathbb{N}$ such that for any $u, v \in \mathcal{A}^\ell$, we have :

$$\mu[u]\mu[v] - \mu([u] \cap \sigma^{-2^n}[v]) < \frac{\alpha}{2^\ell}.$$

Using the scaling property (5.1), we obtain that for any $u \in \mathcal{A}^\ell$,

$$\mu[u] = \mu G_n[u] = \sum_{v\in\mathcal{A}^\ell} \mu([v] \cap \sigma^{-2^n}([u - v]))$$

where $u - v$ is the word of $\mathcal{A}^\ell$ defined by $(u - v)_k = u_k - v_k$.

In particular, we obtain :

$$\begin{aligned}
\mu[w] &= \sum_{v\in\mathcal{A}^\ell} \mu([v] \cap \sigma^{-2^n}[w - v]) \\
&\geq \Big(\sum_{v\in\mathcal{A}^\ell} \mu[v]\mu[w - v] - \frac{\alpha}{2^\ell} \Big) \\
&= \Big(\sum_{v\in\mathcal{A}^\ell, v\neq w-w'} \mu[v]\mu[w - v] \Big) + \mu[w - w']\mu[w'] - \alpha \\
&\geq \Big(\sum_{v\in\mathcal{A}^\ell, v\neq w-w'} \mu[v]\mu[w] \Big) + \mu[w - w'](\mu[w] + \varepsilon) - \alpha \\
&= \mu[w] + \varepsilon\mu[w - w'] - \alpha.
\end{aligned}$$

Any choice of $\alpha < \varepsilon\mu[w - w']$ gives a contradiction. Thus, the only invariant measure of full support that is shift-mixing is the uniform measure. $\square$

5.2.2 Entropy criteria

In this section, we will give criterion based on ergodicity and entropy that ensure that an invariant measure μ of a CA F on $\mathcal{A}^{\mathbb{Z}}$ is the uniform measure λ. Since F commutes with the shift σ, the ordered pair (F, σ) defines a $\mathbb{N} \times \mathbb{Z}$ action on $\mathcal{A}^{\mathbb{Z}}$. We recall that a (F, σ)-invariant probability measure μ is ergodic for the action (F, σ) if every Borel set invariant under F and σ has μ-measure 0 or 1.

We denote by $\mathcal{P}_\ell$ the partition corresponding to the $|\mathcal{A}|^{2\ell+1}$ cylinders of base $\{-\ell, -\ell+1, \dots, \ell\}$. The refinement of two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ is the partition defined by:

$$\mathcal{P}_1 \vee \mathcal{P}_2 = \{A \cap B\,; A \in \mathcal{P}_1 \text{ and } B \in \mathcal{P}_2\}.$$

Definition 5.4 (Entropy)**.** Let $\mathcal{P}$ be a finite partition of $\mathcal{A}^{\mathbb{Z}}$. The entropy, with respect to μ, of the partition $\mathcal{P}$ is defined by:

$$H_\mu(\mathcal{P}) = -\sum_{A \in \mathcal{P}} \mu(A) \log(\mu(A)).$$

The entropy, with respect to μ, of $F : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ can be defined by:

$$h_\mu(F) = -\lim_{\ell \to \infty} \lim_{N \to \infty} \frac{1}{N} H_\mu\Big(\bigvee_{n=0}^{N-1} F^{-n}(\mathcal{P}_\ell) \Big).$$

This limit exists by subadditivity. We refer for example to the work of Walters [Wal82] for a complete introduction to entropy.

Note that the entropy of the shift σ can also be written:

$$h_\mu(\sigma) = -\lim_{k \to \infty} \frac{1}{k} \sum_{a_0, \dots, a_k \in \mathcal{A}} \mu[a_0 \dots a_k] \log \mu[a_0 \dots a_k].$$

Theorem 5.2 ([HMM03])**.** *Let $\mathcal{A} = \mathbb{Z}_p$ with p prime, and let F be an affine CA on $\mathcal{A}^{\mathbb{Z}}$, of neighbourhood $\mathcal{N} = \{0, 1\}$ and local rule $f(x, y) = ax + by + c$ for some $a, b \in \mathbb{Z}_p^*, c \in \mathbb{Z}_p$. If μ is a (F, σ)-invariant measure such that:*

(i) μ is ergodic for σ,

(ii) $h_\mu(F) > 0$,

then $\mu = \lambda$.

Ergodicity with respect to σ is an extremly strong assumption, but the assumption of ergodicity for the action (F, σ) is not sufficient to guarantee the result.

We will generalise the tools used to prove this theorem [HMM03] and prove a rigidity result for CA of local function of the form $F(x, y) = \rho(ax+by+c)$, where ρ is any permutation of the alphabet $\mathcal{A}$. Let $\mathfrak{S}(\mathcal{A})$ be the group of permutations of $\mathcal{A}$.

Theorem 5.3. *Let $\mathcal{A} = \mathbb{Z}_n$, and let F be a bipermutative CA on $\mathcal{A}^{\mathbb{Z}}$ of neighbourhood $\mathcal{N} = \{0, 1\}$ and local rule $f(x, y) = \rho(ax + by + c)$ for some $a, b \in \mathbb{Z}_n^*, c \in \mathbb{Z}_n$, and $\rho \in \mathfrak{S}(\mathcal{A})$. If μ is a (F, σ)-invariant measure such that:*

(i) μ is ergodic for σ,

(ii) $h_\mu(F) > 0$,

then $h_\mu(F) = \log k$, where k divides n. In particular, if n is a prime number, then $h_\mu(F) = \log n$ and $\mu = \lambda$.

To begin with, let us introduce some preliminary results.

We denote by $\mathfrak{B}$ the Borel $\sigma-$algebra of $\mathcal{A}^{\mathbb{Z}}$. We set $\mathfrak{B}_1 = F^{-1}(\mathfrak{B})$, and given a measure $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$, we define μ_x as its conditional measure with respect to $\mathfrak{B}_1$ at point x. That is, for $A \in \mathfrak{B}$, one has $\mu_x(A) = \mathbb{E}(\mathbf{1}_A|\mathfrak{B}_1)(x)$, and $\mu(\cdot) = \int_{\mathcal{A}^{\mathbb{Z}}} \mu_x(\cdot)\mathrm{d}\mu(x)$.

For $x \in \mathcal{A}^{\mathbb{Z}}$, we set $\mathcal{F}(x) = \{y \in \mathcal{A}^{\mathbb{Z}}\,; F(y) = F(x)\}$. The set $\mathcal{F}(x)$ is the *fiber* of x.

Lemma 5.2. *Let $(\mathcal{A}^{\mathbb{Z}}, F)$ be a cellular automaton and let $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$, one has:*

1. *if μ is σ-invariant then $\mu_x\sigma = \mu_{\sigma(x)}$,*
2. *the support of the measure μ_x is $\mathrm{supp}(\mu_x) = \mathcal{F}(x)$.*

Let us now assume that F is a bipermutative CA of neighbourhood $\mathcal{N} = \{0, 1\}$. Since F is bipermutative, for any $x \in \mathcal{A}^{\mathbb{Z}}$ and for any $\alpha \in \mathcal{A}$, there exists a unique element $y \in \mathcal{F}(x)$ such that $y_0 = \alpha$. For $\omega \in \mathfrak{S}(\mathcal{A})$, we can thus define the operator:

$$\begin{array}{rccll} T_\omega : & \mathcal{A}^{\mathbb{Z}} & \longrightarrow & \mathcal{A}^{\mathbb{Z}} & \\ & x & \longmapsto & y & \text{such that } y \in \mathcal{F}(x) \text{ and } y_0 = \omega(x_0). \end{array}$$

For $\omega \in \mathfrak{S}(\mathcal{A})$, we also define:

$$\phi_\omega(x) = \mu_x(T_\omega^{-1}(x)) = \mu_x(\{T_{\omega^{-1}}(x)\}).$$

In particular, we have $\phi_{\mathrm{Id}}(x) = \mu_x(\{x\})$ and $\phi_\omega(x) = \phi_{\mathrm{Id}}(T_{\omega^{-1}}(x))$. Finally, we set:

$$E_\omega = \{x \in \mathcal{A}^{\mathbb{Z}}\,; \phi_\omega(x) > 0\}.$$

Proposition 5.6. *Let μ be a (σ, F)-invariant measure, ergodic for σ and of positive entropy for F. The following properties are satisfied:*

1. *$\phi_{Id} \circ \sigma = \phi_{Id}$ μ-a.e. ,*
2. *$\phi_{Id} \circ F = \phi_{Id}$ μ-a.e. ,*
3. *$\mu(E_{Id}) = 1$,*
4. *for $\omega \in \mathfrak{S}(\mathcal{A})$, $T_{\omega^{-1}}(\mathbf{1}_{E_\omega}\mu)$ is absolutely continuous with respect to μ, that is: if $\mu(A) = 0$, then $\mu(T_\omega(A) \cap E_\omega) = 0$,*
5. *$\phi_\omega = \phi_{Id}$ μ-a.e. in E_ω.*

Proof. 1. Since μ is σ-invariant, we have $\mu_x\sigma = \mu_{\sigma(x)}$. Consequently, ϕ_{Id} is σ-invariant.

2. By Property 1 and the σ-ergodicity of μ, the function ϕ_{Id} is equal μ-a.e. to some constant c. Since μ is F-invariant, it follows that $\phi_{\mathrm{Id}}(F(x)) = c$ for μ-a.e. x. We thus obtain $\phi_{\mathrm{Id}}(F(x)) = \phi_{\mathrm{Id}}(x) = c$ for μ-a.e. x.

3. By Property 1 and the σ-ergodicity of μ, we have $\mu(E_{\mathrm{Id}}) = 0$ or $\mu(E_{\mathrm{Id}}) = 1$. Moreover, we know that $h_\mu(F) = \int_{\mathcal{A}^{\mathbb{Z}}} -\ln \phi_{\mathrm{Id}}(x)d\mu(x)$ (entropy formula for bipermutative CA [HMM03, Sec. 4.3]). Since $h_\mu(F) > 0$, it follows that $\mu(E_{\mathrm{Id}}) = 1$.

4. We have $\mu(A) = \int_{\mathcal{A}^{\mathbb{Z}}} \mu_x(A)d\mu(x)$, so that if $\mu(A) = 0$, then $\mu_x(A) = 0$ μ-a.e. In particular, for μ-a.e. $x \in T_\omega(A)$, $0 = \mu_x(A) \geq \mu_x(T_\omega^{-1}(x)) = \phi_\omega(x)$, thus $x \notin E_\omega$.

5. By Property 2, $\phi_{\mathrm{Id}}(F(x)) = \phi_{\mathrm{Id}}(x)$ for μ-a.e. x. Using Property 4, we obtain that for μ-almost every $x \in E_\omega$, $\phi_{\mathrm{Id}}(F(T_{\omega^{-1}}(x))) = \phi_{\mathrm{Id}}(T_{\omega^{-1}}(x))$. And since $F(T_{\omega^{-1}}(x)) = F(x)$, it comes $\phi_{\mathrm{Id}}(x) = \phi_{\mathrm{Id}}(T_{\omega^{-1}}(x))$, that is, $\phi_\omega(x) = \phi_{\mathrm{Id}}(x)$ for μ-a.e. $x \in E_\omega$.

$\square$

Proposition 5.7. *Let $\omega \in \mathfrak{S}(\mathcal{A})$. If there exists $d \in \mathbb{N}$ such that $T_\omega \circ \sigma^d = \sigma^d \circ T_\omega$, then for any $\omega' \in \mathfrak{S}(\mathcal{A})$, we have: $\phi_\omega = \phi_{\omega' \circ \omega}$ μ-a.e. in $E_{\omega'}$.*

Proof. Let ν be some ergodic component of μ for σ^d. The measure νF is σ^d-invariant and ergodic for σ^d, and it is absolutely continuous with respect to $\mu F = \mu$. Thus, νF is an ergodic component of μ for σ^d, and it is equal to $\nu\sigma^j$ for some $j \in \{0, \ldots, d-1\}$, so that $\nu F^d = \nu\sigma^{jd} = \nu$.

The function ϕ_ω is σ^d-invariant, since $\phi_\omega(\sigma^d(x)) = \phi_{\mathrm{Id}}(T_{\omega^{-1}}(\sigma^d(x))) = \phi_{\mathrm{Id}}(\sigma^d(T_{\omega^{-1}}(x)))$, and by Property 1 of Prop. 5.6, $\phi_{\mathrm{Id}}(\sigma^d(T_{\omega^{-1}}(x))) = \phi_{\mathrm{Id}}(T_{\omega^{-1}}(x)) = \phi_\omega(x)$ μ-a.e. Thus, for each ergodic component ν of μ (for σ^d), ϕ_ω is equal ν-a.e. to some constant $c_{\nu,\omega}$. And since $\nu F^d = \nu$, we obtain that $\phi_\omega(F^d(x)) = \phi_\omega(x) = c_{\nu,\omega}$ ν-a.e. This is true for each ergodic component of μ. Consequently, $\phi_\omega(F^d(x)) = \phi_\omega(x)$ μ-a.e.

Using Property 4 of Prop. 5.6, we obtain that for μ-a.e. $x \in E_{\omega'}$, $\phi_\omega(F^d(T_{\omega'^{-1}}(x))) = \phi_\omega(T_{\omega'^{-1}}(x))$. Since $F^d(T_{\omega'^{-1}}(x)) = F^d(x)$, it follows that $\phi_\omega(F^d(T_{\omega'^{-1}}(x))) = \phi_\omega(F^d(x)) = \phi_\omega(x)$ μ-a.e. Finally, $\phi_\omega(T_{\omega'^{-1}}(x)) = \phi_\omega(x)$ for μ-a.e. $x \in E_{\omega'}$, that is: $\phi_\omega = \phi_{\omega'\circ\omega}$ μ-a.e. in $E_{\omega'}$. □

Proof of Th. 5.3. For $k \in \mathbb{Z}_n$, let $\omega_k \in \mathfrak{S}(\mathcal{A})$ be the permutation defined by $\omega_k(x) = x + k$. For simplicity, we replace the notations $T_{\omega_k}, \phi_{\omega_k}, E_{\omega_k}$ by T_k, ϕ_k, E_k respectively.

Set $v = b^{-1}a$ (by hypothesis, F is bipermutative and a, b are invertible in $\mathbb{Z}_n$) and let d be such that $v^{2d} = 1$. Observe that $T_k \circ \sigma^{2d} = \sigma^{2d} \circ T_k$, since two elements of the same fiber can be represented as follows.

$$\begin{array}{ccccccc} \ldots & x_0 & x_1 & x_2 & x_3 & \ldots & x_{2d} & \ldots \\ \ldots & x_0 + k & x_1 - kv & x_2 + kv^2 & x_3 - kv^3 & \ldots & x_{2d} + k & \ldots \end{array}$$

Let μ be a (σ, F)-invariant measure, ergodic for σ and of positive entropy for F. We know by Prop. 5.6 that $\mu(E_0) = 1$, and as we have seen in the proof of that proposition, there exists a constant c such that $\phi_k(x) = c$ μ-a.e. in E_k.

By Prop. 5.7, for any $i, k \in \mathcal{A}$, $\phi_k = \phi_{i+k}$ μ-a.e. in E_i.

Let us notice that by definition, $\sum_{j=0}^{n-1} \phi_j(x) = 1$. Let j be the smallest element of $\{1, \ldots, n\}$ such that $\mu(E_j) > 0$ (there exists such a j, since otherwise, we would have $c = 1$ and $h_\mu(F) = 0$). Then in E_j, we have μ-a.s. : $c = \phi_0 = \phi_j = \phi_{2j} = \phi_{3j} = \ldots$. Moreover, for values i that are not in the subgroup of $\mathbb{Z}_n$ generated by j, we have μ-a.s. $\phi_i = 0$, since otherwise, we would get a contradiction with the definition of j. Consequently, $c = \gcd(j, n)/n$, and by the entropy formula, $h_\mu(F) = -\log c$. If n is prime, then the only possibility is that $\gcd(j, n) = 1$ and $h_\mu(F) = \log n$, so that $\mu = \lambda$, meaning that μ is the uniform measure. □

5.2.3 Randomisation

We introduce the following definition.

Definition 5.5 (Randomisation)**.** Let F be a CA on $\mathcal{A}$ and let $M \subset \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$. We say that:

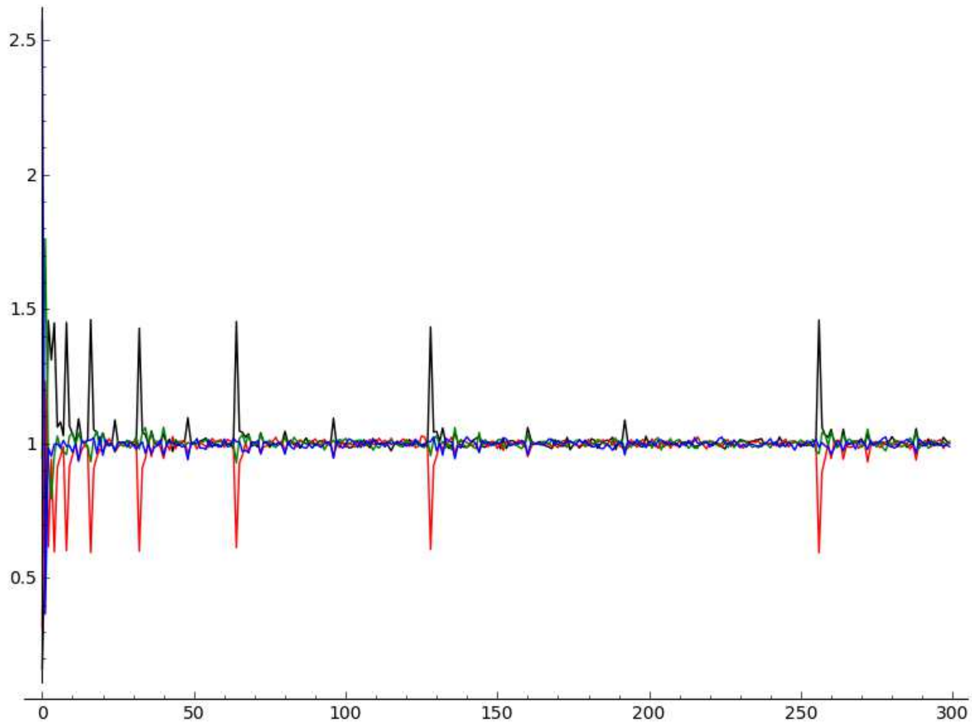
- F randomises M if, for any $\mu \in M$, we have $\mu F^n \to \lambda$;
- F randomises M in Cesàro mean if, for any $\mu \in M$, $\frac{1}{n}\sum_{k=0}^{n} \mu F^k \to \lambda$.

The Cesàro limits have been studied for the CA of algebraic origin [Lin84, MM98, FMMN00]. In particular, it is proved that affine permutative CA randomise in Cesàro mean the set of (non-degenerated) Bernoulli measures. Nevertheless, these CA do not randomise Bernoulli measure (without Cesàro mean), because of their scaling properties. Numerical evidence let us think that some CA of local function of the form $f(x, y) = \rho(ax + by + c)$ could randomise Bernoulli measure.

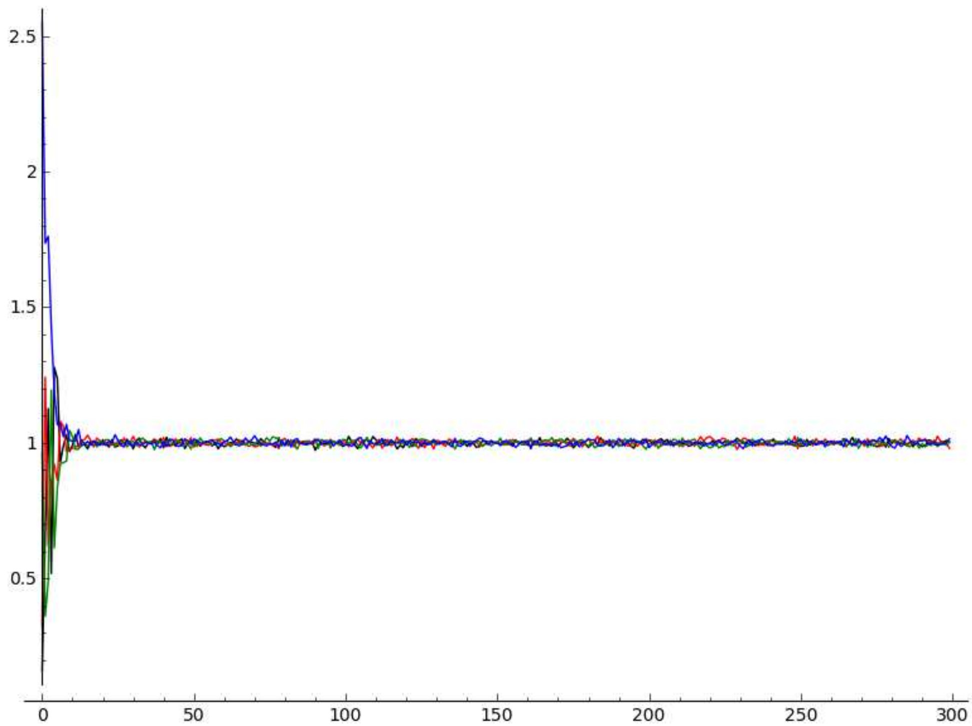
Example 5.4. Let us set $\mathcal{A} = \mathbb{Z}/4\mathbb{Z}$, and $\mathcal{N} = \{0, 1\}$ and compare the behaviour of the CA F and G defined respectively by the local functions $f(x, y) = x + y$ and $g(x, y) = \rho(x + y)$ where ρ is the transposition exchanging 2 and 3, that is, the permutation defined by $\rho(0) = 0$, $\rho(1) = 1, \rho(2) = 3, \rho(3) = 2$.

We start at time $t = 0$ from the Bernoulli product measure of parameters given by the vector $(1/25, 2/25, 6/25, 16/25)$. In a single graphic, we represent the evolution for times $t \in \{0...300\}$ of the number of occurrences of each word with a given length ($\ell = 1$ on Fig. 5.1, and then $\ell = 2, 3, 4$ on Fig. 5.2). The scaling is such that for the uniform measure, the value would be 1 for all the words.

For the first CA, we observe peaks at each power of 4. They are due to the fact that $F^{4^n}(x)_k = x_k + x_{k+4^n}$, which has for consequence that this CA cannot randomise. Such peaks do not appear for the CA G. If the behaviour is the same for larger lengths, it means that G randomises this Bernoulli measure.



$f(x, y) = x + y$



$g(x, y) = \rho(x + y)$

Figure 5.1: Comparison of the behaviours of the CA F and G of Ex. 5.4 for words of length $\ell = 1$. Here, each curve represents a different letter of the alphabet $\mathcal{A}$ (simulations of Hellouin de Menibus).

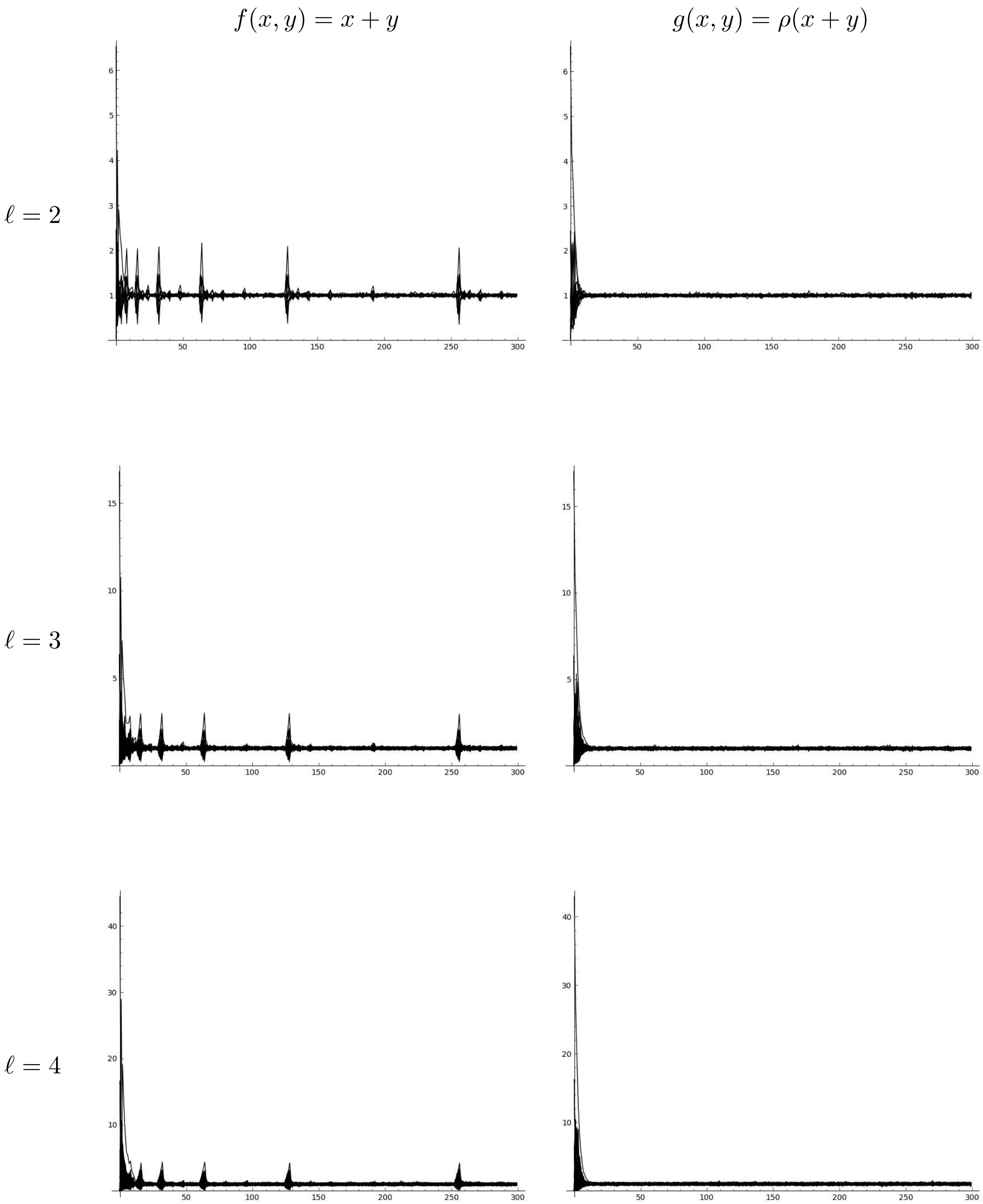


Figure 5.2: Comparison of the behaviours of the CA F and G of Ex. 5.4 for words of length $\ell = 2, 3, 4$ (simulations of Hellouin de Menibus).

Chapter 6

Density classification on infinite lattices and trees

Und weit, weit her Trommelwirbel. Nun gibt es keinen Zweifel mehr, der Aufstand sammelt sich. Ein paar Stunden noch, und die Entscheidung wird fallen. Erregt eilt die Königin immer wieder ans Fenster, um zu lauschen, ob die drohenden Anzeichen sich verstärken. Diese Nacht kennt keinen Schlaf. Endlich, um vier Uhr morgens erhebt sich blutrot die Sonne aus dem wolkenlosen Himmel. Es wird ein heißer Tag werden.

– Stefan Zweig, *Marie-Antoinette*

Contents

Consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter p. We address the density classification problem, that is, we want to design a (probabilistic or deterministic) cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether p is smaller or larger than 1/2. Precisely, the trajectories should converge to the uniform configuration with only 0's if $p < 1/2$, and only 1's if $p > 1/2$. We present solutions to the problem on the regular grids of dimension d, for any $d \geq 2$, and on the regular infinite trees. For the bi-infinite line, we propose some candidates that we back up with numerical simulations.

6.1 The density classification problem

Consider a finite or a countably infinite set of cells, which are spatially arranged according to a group structure G. The *density classification problem* consists in deciding, in a decentralised way, if an initial configuration on G contains more 0's or more 1's. More precisely, the goal is to design a deterministic or probabilistic dynamical system that evolves in the configuration space $\{0,1\}^G$ with a local and homogeneous updating rule and whose trajectories converge to 0^G or to 1^G if the initial configuration contains more 0's or more 1's, respectively. To attack the problem, two natural instantiations of dynamical systems are considered, one with synchronous updates of the cells, and one with asynchronous updates. In the first case, time is discrete, all cells are updated at each time step, and the model is known as a *Probabilistic Cellular Automaton (PCA)* [DKT90]. A *Cellular Automaton (CA)* is a PCA in which the updating rule is deterministic. In the second case, time is continuous, cells are updated at random instants, at most one cell is updated at any given time, and the model is known as a (finite range) *Interacting Particle System (IPS)* [Lig05].

The general spirit of the problem is that of distributed computing: gathering a global information by exchanging only local information. The challenge is two-fold: first, it is impossible to centralise the information (cells are indistinguishable); second, it is impossible to use classical counting techniques (cells contain only binary information).

The density classification problem was originally introduced for synchronous models and rings of finite size ($G = \mathbb{Z}/n\mathbb{Z}$) [Pac88]. After experimentally observing that finding good rules to perform this task was difficult, it was shown that perfect classification with CA is impossible, that is, there exists no given CA that solves the density classification problem for all values of n [LB95]. However, this result did not stop the quest for the best – although imperfect – models as nothing was known about how well CA could perform. The use of PCA opened a new path [Fas02, SOS09] and it was shown that there exist PCA that can classify with an arbitrary precision [Fat11, Fat13]. In the present paper, we complement in Prop. 6.1 the known results by showing that there exists no PCA that perfectly solves the density classification problem for all values of n.

The challenge is now to extend the research to infinite groups whose Cayley graphs are lattices or regular trees. First, we need to specify the meaning of "having more 0's or more 1's" in this context. Consider a random configuration on $\{0,1\}^G$ obtained by assigning independently to each cell a value 1 with probability p and a value 0 with probability $1-p$. We say that a model "classifies the density" if the trajectories converge weakly to 1^G for $p > 1/2$, and to 0^G for $p < 1/2$. A couple of conjectures and negative results exist in the literature. Density classification on $\mathbb{Z}^d$ is referred to by Cox and Durrett under the name of "bifurcation" [CD91]. These two authors study variants of the famous voter model IPS [Lig05, Ch. V] and they propose two instances that are conjectured to bifurcate.

The density classification question has also been addressed for the Glauber dynamics associated to the Ising model at temperature 0, both for lattices and for trees [FSS02, How00, KM11]. The Glauber dynamics defines an IPS or PCA having 0^G and 1^G as invariant measures. Depending on the cases, there is either a proof that the Glauber dynamics does not classify the density, or a conjecture that it does with a proof only for densities sufficiently close to 0 or 1.

The density classification problem has been approached with different perspectives on finite and infinite groups, as emphasized by the results collected above. For finite groups, the problem is studied *per se*, as a benchmark for understanding the power and limitations of cellular automata as a computational model. The community involved is rather on the computer science side. For infinite groups, the goal is to understand the dynamics of specific models that are relevant in statistical mechanics. The community involved is rather on the theoretical physics and probability theory side.

The aim of the present chapter is to investigate how to generalise the finite group approach to the infinite group case.

We want to build models of PCA and IPS, as simple as possible, that correct random noise in the initial configuration, even if the density of errors is close to 1/2. We consider the groups $\mathbb{Z}^d$, whose Cayley graphs are lattices (Sec. 6.2), and the free groups, whose Cayley graphs are infinite regular trees (Sec. 6.3). In all cases, except for $\mathbb{Z}$, we obtain both PCA and IPS models that classify the density. To the best of our knowledge, they constitute the first known such examples. The case of $\mathbb{Z}$ is more complicated and still open. We provide some potential candidates for density classification together with simulation experiments (Sec. 6.4).

6.1.1 The density classification problem on $\mathbb{Z}_n$

The density classification problem was originally stated as follows: find a finite neighbourhood $\mathcal{N} \subset \mathbb{Z}$ and a transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$ such that for any integer $n \geq 1$ and any configuration $x \in \mathcal{A}^{\mathbb{Z}_n}$, when applying the CA F of transition function f to x, the sequence of iterates $(F^k(x))_{k\geq 0}$ reaches the fixed point $\mathbf{0} = 0^n$ if $|x|_0 > |x|_1$ and the fixed point $\mathbf{1} = 1^n$ if $|x|_1 > |x|_0$. The problem can be extended to PCA by requiring the measure $(\delta_x F^t)_{t\geq 0}$ to converge to $\delta_{\mathbf{0}}$, resp. $\delta_{\mathbf{1}}$. (Or equivalently, by requiring the space-time diagram to converge almost surely to $\mathbf{0}$, resp. $\mathbf{1}$.)

Land and Belew have proved that there exists no CA that perfectly performs this density classification task for all values of n [LB95]. We now prove that this negative result can be extended to PCA. It provides at the same time a new proof for CA as a particular case.

Denote by δ_x the probability measure corresponding to a Dirac distribution centered on x.

Proposition 6.1. *There exist no PCA or IPS that solves perfectly the density classification problem on $\mathbb{Z}_n$, that is, for any integer $n \geq 1$, and for any configuration $x \in \mathcal{A}^{\mathbb{Z}_n}$, $(\delta_x F^t)_{t\geq 0}$ converges to $\delta_{\mathbf{0}}$ if $|x|_0 > n/2$ and to $\delta_{\mathbf{1}}$ if $|x|_1 > n/2$.*

Proof. We carry out the proof for PCA. For IPS, the argument is similar and even simpler. Let us assume that F is a PCA that solves perfectly the density classification problem on $\mathbb{Z}_n$. Let $\mathcal{N}$ be the neighbourhood of F, and let ℓ be such that $\mathcal{N} \subset [-\ell+1, \ell-1]$. We will prove that for any $x \in \mathcal{A}^{\mathbb{Z}_n}$ (with $n \geq 2\ell$), the number of occurrences of 1's after application of F to x is almost surely equal to $|x|_1$. Let us assume that it is not the case. Then, there exist two words x and y having different numbers of 1's, such that from the word x, there is a positive probability to reach y in one step. Formally, this can be written:

$$\exists x, y \in \mathcal{A}^{\mathbb{Z}_n},\ |x|_1 \neq |y|_1, \qquad \delta_x F(y) > 0\,. \tag{6.1}$$

Assume for instance that $|y|_1 > |x|_1$ (the case $|y|_1 < |x|_1$ is treated similarly). We first assume that $|x|_1 = a > n/2$. We will construct a configuration z of density smaller than 1/2, from which there is a positive probability to reach a configuration w of density larger than 1/2. For integers $k \geq 2, m \geq 2\ell$, let us consider the configuration $z = x^k 0^m \in \mathcal{A}^{\mathbb{Z}_{kn+m}}$. We have $|z|_1 = ka$. Let $y_s = y_{\ell+1} \dots y_n$ be the suffix of length $n-\ell$ of y, and let $y_p = y_1 \dots y_{n-l}$. be the prefix of length $n-\ell$ of y. By applying equation (6.1), it follows that:

$$\exists u, v, u', v' \in \mathcal{A}^{\ell}, \qquad \delta_z F(u y_s y^{k-2} y_p v u' 0^{m-2\ell} v') > 0\,.$$

Set $w = uy_s y^{k-2} y_p v u' 0^{m-2\ell} v'$.

$$z = \underbrace{\overbrace{x \;\; x \;\; \ldots \;\; x}^{k}}_{\leftarrow\ kn\ \rightarrow} \; \underbrace{\overbrace{0 \;\; 0 \;\; \ldots \;\; 0}^{m}}_{\leftarrow\ m\ \rightarrow}$$

$$w = \underset{\ell}{u} \;\; \underset{n-\ell}{y_s} \;\; \underbrace{\overbrace{y \;\; y \;\; \ldots \;\; y}^{k-2}}_{\leftarrow\ (k-2)n\ \rightarrow} \;\; \underset{n-\ell}{y_p} \;\; \underset{\ell}{v} \;\; \underset{\ell}{u'} \;\; \underbrace{\overbrace{0 \;\; 0 \;\; \ldots \;\; 0}^{m-2\ell}}_{\leftarrow\ (m-2\ell)\ \rightarrow} \;\; \underset{\ell}{v'}$$

We have $|w|_1 \geq k|y|_1 - 2\ell \geq k(a+1) - 2\ell$.

For large enough m, if we set k to be the largest integer such that $k(a - n/2) < m/2$ (implying that $(k+1)(a - n/2) \geq m/2$, so that $ka \geq (kn+m)/2 + n/2 - a$), we have:

$$|z|_1 = ka < \frac{kn+m}{2}, \qquad |w|_1 \geq k(a+1) - 2\ell \geq \frac{kn+m}{2} + \frac{n}{2} - a + k - 2\ell > \frac{kn+m}{2},$$

the last inequality coming from the fact that for large enough m, $k > a + 2\ell$. So, with a positive probability, we can reach a configuration with more ones than zeros starting from a configuration with more zeros than ones. Since F classifies the density with probability 1, the new configuration can be considered as an initial condition that needs to be classified and will thus almost surely evolve to the fixed point $\mathbf{1}$, that is, a bad classification will occur, which contradicts our hypothesis.

The case $|x|_1 < n/2$ can be handled by swapping the roles of 0 and 1.

We have proved that for any $x \in \mathcal{A}^{\mathbb{Z}_n}$ (with $n \geq \ell$), the number of occurrences of ones after application of F to x is almost surely equal to $|x|_1$. This is in contradiction with the fact that F classifies the density. □

The proof can be adapted to larger dimensions and we obtain the following.

Proposition 6.2. *For any $d \geq 1$, there is no d-dimensional PCA or IPS such that for any integers $n_1, \ldots, n_d \geq 1$, and for any configuration $x \in \mathcal{A}^{\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_d}}$, $(\delta_x F^t)_{t \geq 0}$ converges to $\delta_{\mathbf{0}}$ if $|x|_0 > (n_1 \ldots n_d)/2$ and to $\delta_{\mathbf{1}}$ if $|x|_1 > (n_1 \ldots n_d)/2$.*

6.1.2 The density classification problem on infinite groups

Let us define formally the density classification problem on infinite groups.

We denote by μ_p the Bernoulli measure of parameter p, that is, the product measure of density p on $X = \mathcal{A}^G$. A realisation of μ_p is obtained by assigning independently to each element of G a label 1 with probability p and a label 0 with probability $1-p$. Set $\mathbf{0} = 0^G$ and $\mathbf{1} = 1^G$.

The *density classification problem* consists in finding a PCA or an IPS F, such that:

$$\begin{cases} p < 1/2 \implies \mu_p F^t \xrightarrow[t\to\infty]{w} \delta_{\mathbf{0}}, \\ p > 1/2 \implies \mu_p F^t \xrightarrow[t\to\infty]{w} \delta_{\mathbf{1}}. \end{cases} \tag{6.2}$$

The notation $\xrightarrow{w}$ stands for the weak convergence of measures. In our case, the interpretation of this convergence is that for any *finite* subset $K \subset G$, the probability that all the cells of K are labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$). Or, equivalently, that for any single cell, the probability that it is labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$).

From subgroups to groups. Next result will be used several times.

Proposition 6.3. *Let H be a subgroup of G, and let F_H be a process (PCA or IPS) of neighbourhood $\mathcal{N}$ and transition function f that classifies the density on $\mathcal{A}^H$. We denote by F_G the process on $\mathcal{A}^G$ having the same neighbourhood $\mathcal{N}$ and the same transition function f. Then, F_G classifies the density on $\mathcal{A}^G$.*

Proof. Since H is a subgroup, the group G is partitioned into a union of classes $g_1H, g_2H, \ldots$ We have $\mathcal{N} \subset H$, so that if an element $g \in G$ is in some class g_iH, then for any $v \in \mathcal{N}$, the element $g \cdot v$ is also in g_iH. Since F_H classifies the density, on each class g_iH, the process F_G satisfies equation (6.2). Thus for any cell of G, the probability that it is labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$). □

6.2 Classifying the density on $\mathbb{Z}^d, d \geq 2$

According to Prop. 6.3, given a process that classifies the density on $\mathbb{Z}^2$, we can design a new process that classifies on $\mathbb{Z}^d$ for $d > 2$, by considering $\mathbb{Z}^d$ as a pile of $\mathbb{Z}^2$-layers, and by classifying the density independently on each of these layers. It doesn't mean that there are no other ways to classify the density, for which the different layers would interact together, but it gives at least one elementary way to classify the density of $\mathbb{Z}^d, d > 1$, if we know how to do on $\mathbb{Z}^2$.

Below, we concentrate on $\mathbb{Z}^2$.

To classify the density on $\mathbb{Z}^2$, a first natural idea is to apply the majority rule on a cell and its four direct neighbours. Unfortunately, this does not work, neither in the CA nor in the IPS version. Indeed, a 2×2 square of four cells in state 1 (resp. 0) remains in state 1 (resp. 0) forever. For $p \in (0, 1)$, monochromatic elementary squares of both colors appear almost surely in the initial configuration which makes the convergence to $\mathbf{0}$ or $\mathbf{1}$ impossible. We prove more generally that on $\mathbb{Z}^d$, the majority rule over a symmetric neighbourhood that contains the cell itself has a finite stable pattern (Fig. 6.1 represents two examples on $\mathbb{Z}^2$). Classification of the density is thus impossible. We recover the "forbidden symmetry" of Pippenger [Pip94].

Lemma 6.1. *Let us consider a set $\mathcal{N} = \{e_0, e_1, \ldots, e_n, -e_1, \ldots, -e_n\}$ of $(2n + 1)$ different elements of $\mathbb{Z}^d$, with $e_0 = (0, \ldots, 0)$. If the cells of the set $\mathcal{D} = \{\sum_{i \in S} e_i | S \subset \{0, \ldots, n\}\}$ are initially in the same state, then they remain in that same state when iterating the majority CA or IPS of neighbourhood $\mathcal{N}$.*

Proof. Let us fix any subset S of $\{0, \ldots, n\}$, and consider the cell $c = \sum_{i \in S} e_i$. We want to prove that c has at least $n + 1$ neighbours which belong to $\mathcal{D}$. First the cell c is in its own neighbourhood. For $j \in S$, the cell $c - e_j = \sum_{i \in S \setminus \{j\}} e_i$ belongs to $\mathcal{D}$, and for $j \in \{1, \ldots, n\} \setminus S$, the cell $c + e_j = \sum_{i \in S \cup \{j\}} e_i$ belongs to $\mathcal{D}$. Therefore c has at least $n + 1$ neighbours in $\mathcal{D}$. If all the cells of $\mathcal{D}$ are in the same state, when applying the majority rule, this state is preserved. □

On $\mathbb{Z}^2$, another idea is to apply the majority rule on the four nearest neighbours (excluding the cell itself) and to choose uniformly the new state of the cell in case of equality. In the IPS setting, this process is known as the Glauber dynamics associated to the Ising model. It has been conjectured to classify the density, but the result has been proved only for values of p that are sufficiently close to 0 or 1 [FSS02].

To overcome the difficulty, we consider the majority CA but on the asymmetric neighbourhood $\mathcal{N} = \{(0, 0), (0, 1), (1, 0)\}$. This CA, known as Toom's rule [DKT90, Too80], has

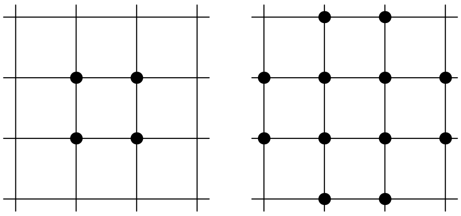


Figure 6.1: Stable patterns obtained respectively for the von Neumann neighbourhood (the cell and its four nearest neighbours) and the Moore neighbourhood (the cell and its eight surrounding neighbours).

been introduced in connection with the positive rates problem, see Sec. 6.2.3. Here we prove that Toom's CA classifies the density on $\mathbb{Z}^2$. Our proof relies on the properties of the percolation clusters on the triangular lattice [Gri99]. We then define an IPS inspired by this local rule and prove with the same techniques that it also classifies the density.

6.2.1 A cellular automaton that classifies the density

Let us denote by $\mathtt{maj} : \mathcal{A}^3 \to \mathcal{A}$, the majority function, so that

$$\mathtt{maj}(x,y,z) = \begin{cases} 0 & \text{if } x+y+z < 2 \\ 1 & \text{if } x+y+z \geq 2 \end{cases}.$$

Theorem 6.1. *The cellular automaton $\mathcal{T} : \mathcal{A}^{\mathbb{Z}^2} \to \mathcal{A}^{\mathbb{Z}^2}$ defined by:*

$$\mathcal{T}(x)_{i,j} = \mathtt{maj}(x_{i,j}, x_{i,j+1}, x_{i+1,j})$$

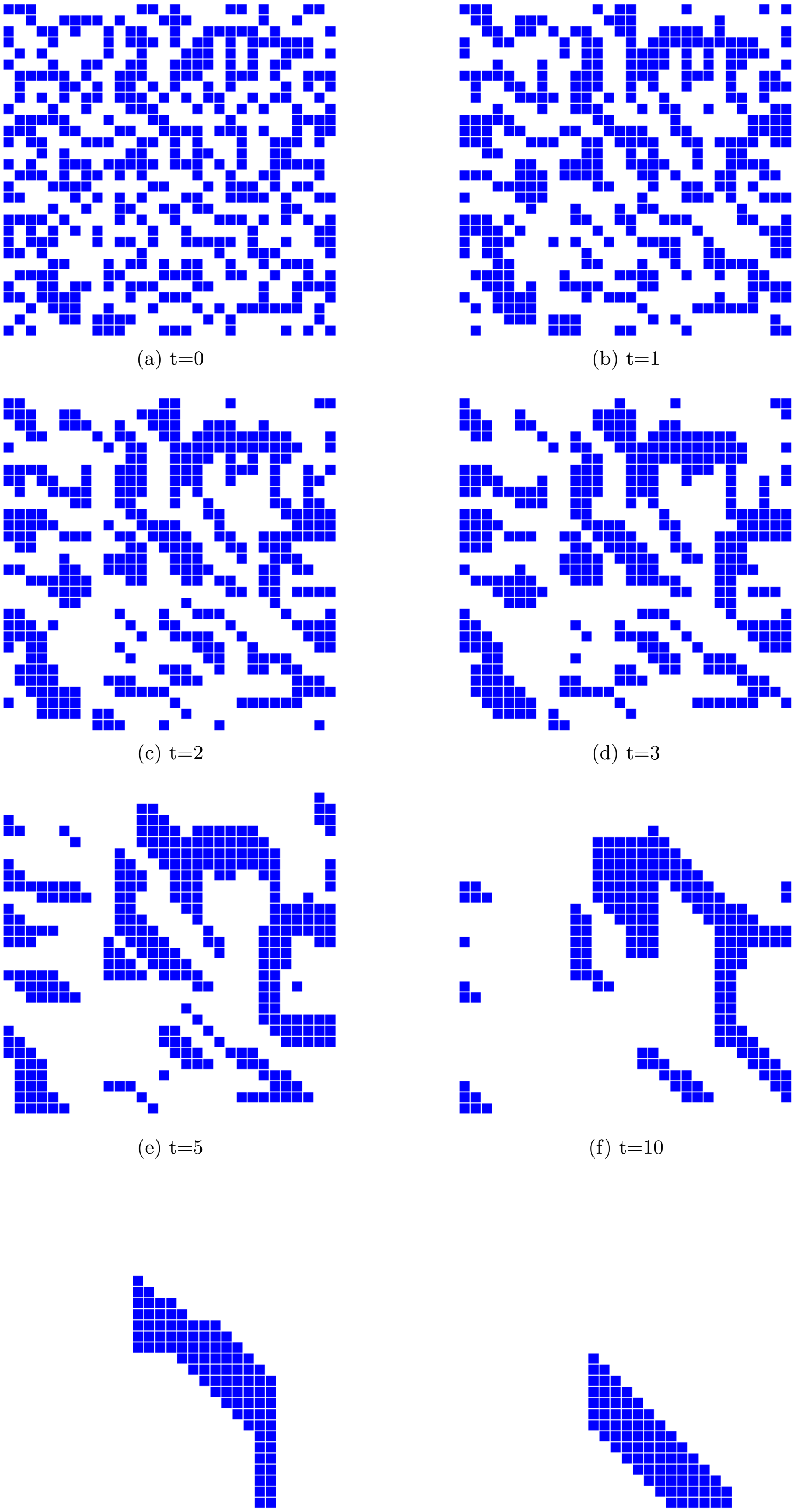
for any $x \in \mathcal{A}^{\mathbb{Z}^2}$, $(i,j) \in \mathbb{Z}^2$, classifies the density.

Proof. By symmetry, it is sufficient to prove that if $p > 1/2$, then $(\mu_p \mathcal{T}^n)_{n \geq 0}$ converges weakly to $\delta_{\mathbf{1}}$.

Let us consider the graph defined with $\mathbb{Z}^2$ as the set of sites (vertices) and $\{\{(i,j),(i,j+1)\},\{(i,j),(i+1,j)\},\{(i+1,j),(i,j+1)\},(i,j) \in \mathbb{Z}^2\}$ as the set of bonds (edges). This graph is equivalent to a triangular lattice, on which our notion of connectivity is defined. We recall that a 0-*cluster* is a subset of connected sites labelled by 0 which is maximal for inclusion. The site percolation threshold on the triangular lattice is equal to 1/2 so that, for $p > 1/2$, there exists almost surely no infinite 0-cluster [Gri99]. Thus, if $S_{\mathbf{0}}$ denotes the set of sites labelled by 0, the set $S_{\mathbf{0}}$ consists almost surely of a countable union $S_{\mathbf{0}} = \cup_{k \in \mathbb{N}} S_k$ of finite 0-clusters. Moreover, the size of the 0-clusters decays exponentially: there exist some constants κ and γ such that the probability for a given site to be part of a 0-cluster of size larger than n is smaller than $\kappa e^{-\gamma n}$ [Gri99].

Let us describe how the 0-clusters are transformed by the action of the CA. For $S \subset \mathbb{Z}^2$, let 1_S be the configuration defined by $(1_S)_x = 1$ if $x \in S$ and $(1_S)_x = 0$ otherwise. Let $\mathcal{T}(S)$ be the subset S' of $\mathbb{Z}^2$ such that $\mathcal{T}(1_S) = 1_{S'}$. By a simple symmetry argument, this last equality is equivalent to $\mathcal{T}(1_{\mathbb{Z}^2 \setminus S}) = 1_{\mathbb{Z}^2 \setminus S'}$. We observe the following.

- The rule does not break up or connect different 0-clusters (proved by Gács [Gác90, Fact 3.1]). More precisely, if S consists of the 0-clusters $(S_k)_k$, then the components of $\mathcal{T}(S)$ are the nonempty sets among $(\mathcal{T}(S_k))_k$.



(a) t=0 (b) t=1

(c) t=2 (d) t=3

(e) t=5 (f) t=10

(g) t=20 (h) t=30

Figure 6.2: Simulation of Toom CA from a Bernoulli measure of parameter 0.45.

- Any finite 0-cluster disappears in finite time: if S is a finite and connected subset of $\mathbb{Z}^2$, then there exists an integer $n \geq 1$ such that $\mathcal{T}^n(S) = \emptyset$. This is the *eroder* property [DKT90].

- Let us consider a 0-cluster and a rectangle in which it is contained. Then the 0-cluster always remains within this rectangle. More precisely, if R is a rectangle set, that is, a set of the form $\{(x, y) \in \mathbb{Z}^2 \mid a_1 \leq x \leq a_2,\ b_1 \leq y \leq b_2\}$, and if $S \subset R$, then for all $n \geq 1$, $\mathcal{T}^n(S) \subset R$ (the proof follows from $\mathcal{T}(S) \subset \mathcal{T}(R) \subset R$).

Let us now consider all the 0-clusters for which the minimal enveloping rectangle contains the origin $(0, 0)$. By the exponential decay of the size of the clusters, one can prove that the number of such 0-clusters is almost surely finite. Indeed, the probability that the point of coordinates (m, n) is a part of such a cluster is smaller than the probability for this point to belong to a 0-cluster of size larger than $\max(|m|, |n|)$. And since

$$\sum_{(m,n)\in\mathbb{Z}^2} \kappa e^{-\gamma \max(|m|,|n|)} < 4\kappa \sum_{m\in\mathbb{N}} (me^{-\gamma m} + \sum_{n\geq m} e^{-\gamma n}) < \infty,$$

we can apply the Borel-Cantelli lemma to obtain the result. Let T_0 be the maximum of the time needed to erase these 0-clusters. The random variable T_0 is almost surely finite, and after T_0 time steps, the site $(0, 0)$ will always be labelled by a 1. As the argument can be generalised to any site, it ends the proof. $\square$

We point out that Toom's CA classifies the density despite having many different invariant measures. For example:

- Any configuration x that can be decomposed into monochromatic North-East paths (that is, $x_{i,j} = x_{i,j+1}$ or $x_{i,j} = x_{i+1,j}$ for any i, j) is a fixed point and δ_x is an invariant measure.

- Let y be the checkerboard configuration defined by $y_{i,j} = 0$ if $i + j$ is even and $y_{i,j} = 1$ otherwise, and let z be defined by $z_{i,j} = 1 - y_{i,j}$. Since we have $\mathcal{T}(y) = z$ and $\mathcal{T}(z) = y$, the two configurations y and z form a periodic orbit and $(\delta_y + \delta_z)/2$ is an invariant measure.

6.2.2 An interacting particle system that classifies the density

We now define an IPS for which we use the same steps as above to prove that it classifies the density.

Note that the exact IPS analogue of Toom's rule might classify the density but the above proof does not carry over since, in some cases, different 0-clusters may merge. To overcome the difficulty, we introduce a different IPS with a new neighbourhood of size 7: the cell itself and the six cells that are connected to it in the triangular lattice defined in the previous section.

For $\alpha \in \mathcal{A}$, set $\bar{\alpha} = 1 - \alpha$.

Theorem 6.2. *Let us consider the following IPS: for a configuration $x \in \mathcal{A}^{\mathbb{Z}^2}$, we update the state of the cell (i, j) by applying the majority rule on the North-East-Centre neighbourhood, except in the following cases (for which we keep the state unchanged):*

1. $x_{i,j} = x_{i-1,j+1} = x_{i+1,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j}$ *and* $(x_{i,j-1} = \bar{x}_{i,j}$ *or* $x_{i-1,j} = \bar{x}_{i,j})$,

2. $x_{i,j} = x_{i-1,j+1} = x_{i,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j} = \bar{x}_{i+1,j-1}$ *and* $x_{i-1,j} = \bar{x}_{i,j}$,

3. $x_{i,j} = x_{i-1,j} = x_{i+1,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j} = \bar{x}_{i-1,j+1}$ *and* $x_{i,j-1} = \bar{x}_{i,j}$.

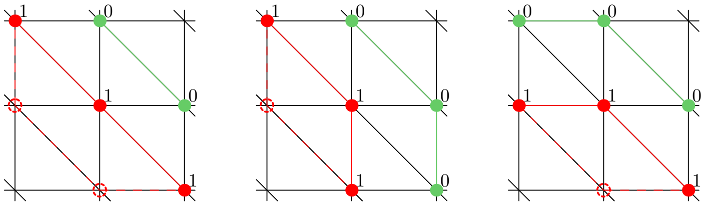


Figure 6.3: Illustration of the definition of the IPS.

This IPS classifies the density.

The three cases for which we always keep the state unchanged are illustrated below for $x_{i,j} = 1$ (central cell). In the first case, we allow to flip the central cell if and only if the two cells marked by a dashed circle are also labelled by 1. Otherwise, the updating could connect two different 0-clusters and break up the 1-cluster to which the cell (i, j) belongs to. The second and third cases are analogous.

The proof is similar to the one of Th. 6.1 but involves some additional technical points.

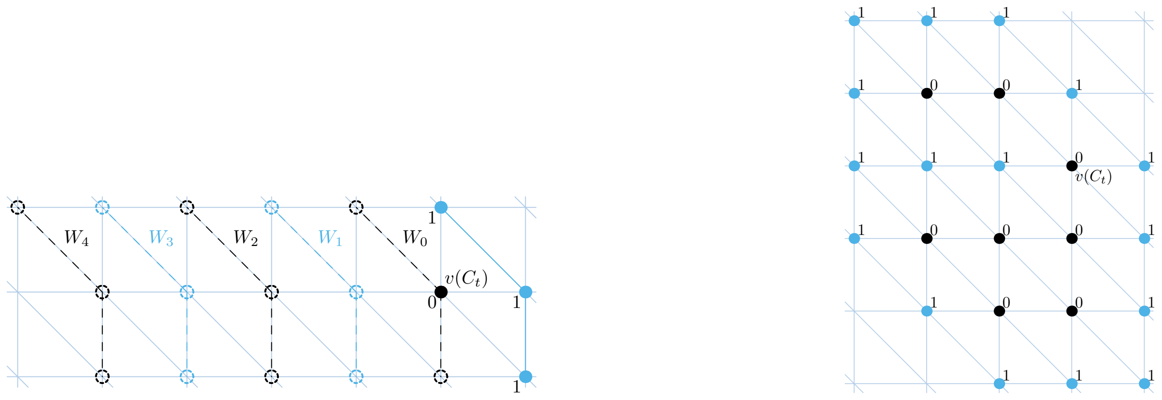
Proof. We assume as before that $p > 1/2$. Like the CA of the previous section, the new process that we have defined never breaks up a cluster or connects different ones. Furthermore, if we consider a 0-cluster and the smallest rectangle in which it is contained, we can check again that the 0-cluster will never go beyond this rectangle. As before, we only need to prove that any finite 0-cluster disappears almost surely in finite time to conclude the proof. We consider a realisation of the trajectory of the IPS with initial density μ_p. We associate to any finite 0-cluster $C \subset \mathbb{Z}^2$ the point $v(C) = \max\{(i, j) \in C\}$, using the lexicographic order on the coordinates (we set $v(\emptyset) = (-\infty, -\infty)$). In other words, the point $v(C)$ is the upmost point of C among its rightmost points. Let us consider at time 0 some finite 0-cluster C_0. We denote by C_t the state of this cluster at time t.

Claim. *The sequence $v(C_t)$ is nonincreasing. Moreover, if $t \geq 0$ is such that $C_t \neq \emptyset$, then there exists almost surely a time $t' > t$ such that $v(C_{t'}) < v(C_t)$.*

Let us prove the claim. Let us denote by $x \in \mathcal{A}^{\mathbb{Z}^2}$ a configuration attained at some time t, and let $(i, j) = v(C_t)$. By definition of $v(C_t)$, if a cell of coordinates $(i+1, j')$ is connected to a cell of C_t, then $x_{i+1,j'} = 1$. Either we have also $x_{i+1,j'+1} = 1$ and the cell $(i+1, j')$ will not flip, or $x_{i+1,j'+1} = 0$, but in this case, since $(i+1, j'+1)$ does not belong to C_t, $x_{i,j'+1} = 1$ and the cell of C_t to which is connected $(i+1, j')$ is necessarily (i, j'). So, $x_{i,j'} = 0$ and $x_{i+1,j'-1} = 1$, once again by definition of $v(C_t)$. Depending on the value of $x_{i+2,j'-1}$, either rule 1 or rule 2 forbids the cell $(i+1, j')$ to flip. In the same way, we can prove that if a cell of coordinates $(i, j'), j' > j$ is connected to C_t, then it is not allowed to flip. This proves that $v(C_t)$ is nonincreasing.

In order to prove the second part of the claim, we need to show that the cell (i, j) will almost surely be flipped in finite time. By definition of $(i, j) = v(C_t)$, we know that $x_{i,j+1} = x_{i+1,j} = x_{i+1,j-1} = 1$. The cell (i, j) will thus be allowed to flip, except if $x_{i-1,j+1} = x_{i,j-1} = 0$ and $x_{i-1,j} = 1$. But in that case, the cell $(i-1, j)$ will end up flipping, except if $x_{i-1,j-1} = x_{i-2,j+1} = 1, x_{i-2,j} = 0$, and so on. Let $W_n = \{(i-n, j), (i-1-n, j+1), (i-n, j-1)\}$. If for each n, the cells of W_n are in the state $(n \mod 2)$, then none of the cells $(i-n, j)$ is allowed to flip (see Fig. 6.4.a). But recall now that the initial measure is μ_p. There exists almost surely an integer $n \geq 0$ such that the initial state of the cell $(i-n, j)$ is *not* $(n \mod 2)$.

Let $m(t)$ be the smallest integer n whose value at time t is not $n \mod 2$. Then, one can easily check that $m(t)$ is non-increasing, and that it reaches 0 in finite time. Thus, the cell (i, j) ends up flipping and we have proved the claim.



(a) (b)

Figure 6.4: Illustration of the proof of Th. 6.2.

The example of Fig. 6.4.b illustrates how the proof works. Here, no cell of the cluster C_t is allowed to flip, but since the cells on the right and on the top of $v(C_t)$ cannot flip either, $v(C_t)$ does not increase. The cell at the left of $v(C_t)$ will end up flipping, and $v(C_t)$ will then be allowed to flip.

Since we know that a 0-cluster cannot go beyond its enveloping rectangle, a direct consequence of the claim is that any 0-cluster disappears in finite time. This allows us to conclude the proof in the same way as for the majority cellular automaton. $\square$

6.2.3 The positive rates problem in $\mathbb{Z}^2$

Let us mention a connected problem and result. By definition, a PCA or an IPS of local function $\varphi : \mathcal{A} \longrightarrow \mathcal{M}(\mathcal{A})$ has *positive rates* if:

$$\forall u \in \mathcal{A}^{\mathcal{N}}, \ \forall a \in \mathcal{A}, \qquad \varphi(u)(a) > 0 \,. \tag{6.3}$$

The *positive rates problem* consists in finding a positive-rate model which is non-ergodic (with several invariant measures). This is a natural question, also relevant in the context of fault-tolerant models of computation, and which has been extensively studied.

In $\mathbb{Z}^2$, the positive rates problem is solved by a "perturbation" of Toom's CA. In fact, this was the motivation that led Toom to introduce the CA that bears his name. Let φ_0 be the local function of Toom's CA, seen here as a function into $\mathcal{M}(\mathcal{A})$, and define the positive-rate PCA F with local function $\varphi : \mathcal{A}^{\mathcal{N}} \longrightarrow \mathcal{M}(\mathcal{A})$ given by:

$$\forall u \in \mathcal{A}^{\mathcal{N}}, \qquad \varphi(u) = (1-\varepsilon)\varphi_0(u) + \varepsilon \, \texttt{Unif} \,, \tag{6.4}$$

where $\varepsilon \in (0,1)$ and where `Unif` is the uniform probability distribution on $\mathcal{A}$. The interpretation is that the computations are done according to Toom's rule, but, at each time and in each cell, an error may occur with probability ε in which case the new cell value is chosen uniformly. It is proved by Toom [DKT90, Too80] that for ε small enough, the positive-rate PCA F has several invariant measures, with at least one close to "all 0", and one close to "all 1".

Intuitively and roughly, this non-ergodicity result and the one in Th. 6.1 can be viewed as being complementary, expressing the very strong "erasing" capacities of Toom's CA. Density classification amounts to erasing "errors" in the initial configuration (the symbols which are in minority), and non-ergodicity amounts to almost-erasing "errors" occurring in the whole space time diagram (the 1's if we are close to "all 0", or the 0's if we are close to "all 1").

6.3 Classifying the density on regular trees

Consider the finitely presented group $T_n = \langle a_1, \ldots, a_n \mid a_i^2 = 1 \rangle$. The Cayley graph of T_n is the infinite n-regular tree. For $n = 2k$, we also consider the free group with k generators, that is, $T'_{2k} = \langle a_1, \ldots, a_k \mid \cdot \rangle$. The groups T_{2k} and T'_{2k} are not isomorphic, but they have the same Cayley graph.

6.3.1 Shortcomings of the nearest neighbour majority rules

For odd values of n, a natural candidate for classifying the density is to apply the majority rule on the n neighbours of a cell. But it is proved that neither the CA (see the work of Kanoria [KM11] for $n = 3, 5$, and 7) nor the IPS (see the work of Howard [How00] for $n = 3$) classify the density.

For $n = 4$, a natural candidate would be to apply the majority on the four neighbours and the cell itself. We now prove that it does not work either.

Proposition 6.4. *Consider the group $T'_4 = \langle a, b \mid \cdot \rangle$. Consider the majority CA or IPS with neighbourhood $\mathcal{N} = \{1, a, b, a^{-1}, b^{-1}\}$. For $p \in (1/3, 2/3)$, the trajectories do not converge weakly to a uniform configuration.*

Proof. If $p \in (1/3, 2/3)$, then we claim that at time 0, there are almost surely infinite chains of zeros and infinite chains of ones that are fixed. Let us choose some cell labelled by 1. Consider the (finite or infinite) subtree of 1's originating from this cell viewed as the root. If we forget the root, the random tree is exactly a Galton-Watson process. The expected number of children of a node is $3p$ and since $3p > 1$, this Galton-Watson process survives with positive probability. Consequently, there exists almost surely an infinite chain of 1's at time 0 somewhere in the tree. In the same way, since $3(1-p) > 0$, there exists almost surely an infinite chain of 0's. □

As for $\mathbb{Z}^2$, we get round the difficulty by keeping the majority rule but choosing a non-symmetrical neighbourhood.

6.3.2 A rule that classifies the density on T'_4

In this section, we consider the free group $T'_4 = \langle a, b | \cdot \rangle$, see Fig. 6.5 (a).

Theorem 6.3. *The cellular automaton $F : \mathcal{A}^{T'_4} \to \mathcal{A}^{T'_4}$ defined by:*

$$F(x)_g = \mathtt{maj}(x_{ga}, x_{gab}, x_{gab^{-1}})$$

for any $x \in \mathcal{A}^{T'_4}, g \in T'_4$, classifies the density.

Proof. We consider a realisation of the trajectory of the CA with initial distribution μ_p. Let us denote by X_g^n the random variable describing the state of the cell g at time n. Since the process is homogeneous, it is sufficient to prove that X_1^n converges almost surely to 0 if $p < 1/2$ and to 1 if $p > 1/2$. Let us denote by $h : [0,1] \to [0,1]$ the function that maps a given $p \in [0,1]$ to the probability $h(p)$ that $\mathtt{maj}(X, Y, Z) = 1$ when X, Y, Z are three independent Bernoulli random variables of parameter p. An easy computation provides $h(p) = 3p^2 - 2p^3$, and one can check that the sequence $(h^n(p))_{n \geq 0}$ converges to 0 if $p < 1/2$ and to 1 if $p > 1/2$.

We prove by induction on $n \in \mathbb{N}$ that for any $k \in \mathbb{N}$, the family $\mathcal{E}_k(n) = \{X^n_{u_1 u_2 \ldots u_k} \mid u_1, u_2, \ldots, u_k \in \{a, ab, ab^{-1}\}\}$ consists of independent Bernoulli random variables of parameter $h^n(p)$. By definition of μ_p, the property is true at time $n = 0$. Let us assume that it is true at some time $n \geq 0$, and let us fix some $k \geq 0$.

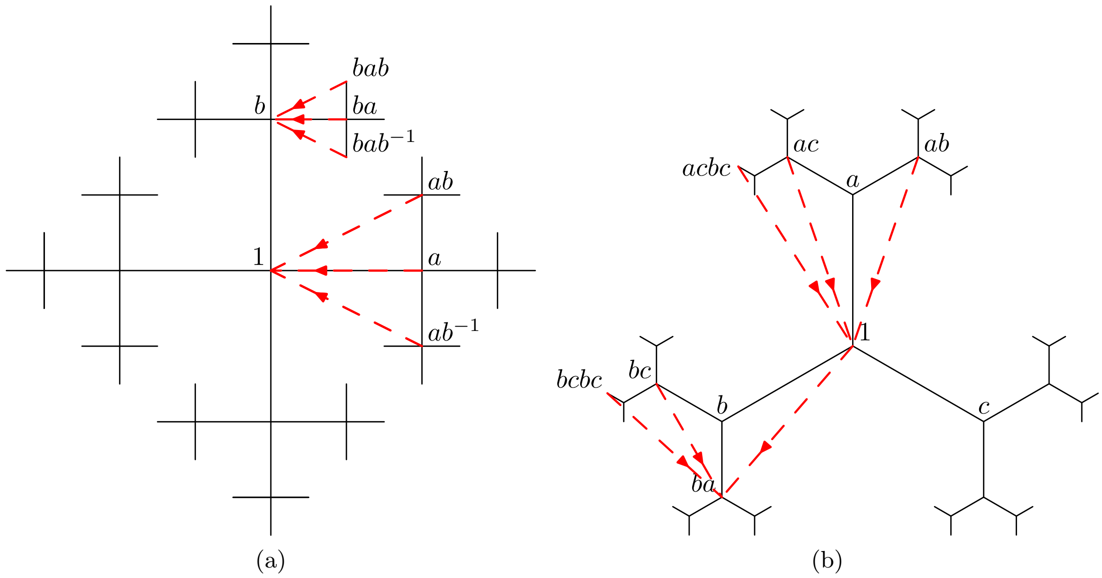


Figure 6.5: The cellular automata described by Th. 6.3 and Th. 6.4.

Let $u_1, u_2, \dots, u_k$ and $v_1, v_2, \dots, v_k$ be two different sequences of elements of $\{a, ab, ab^{-1}\}$. We have:

$$X^{n+1}_{u_1u_2\dots u_k} = \mathtt{maj}(X^n_{u_1u_2\dots u_k a}, X^n_{u_1u_2\dots u_k ab}, X^n_{u_1u_2\dots u_k ab^{-1}}),$$

$$X^{n+1}_{v_1v_2\dots v_k} = \mathtt{maj}(X^n_{v_1v_2\dots v_k a}, X^n_{v_1v_2\dots v_k ab}, X^n_{v_1v_2\dots v_k ab^{-1}}).$$

Thus, two different elements of $\mathcal{E}_k(n+1)$ can be written as the majority on two disjoint triples of $\mathcal{E}_{k+1}(n)$. The fact that the triples are disjoint is a consequence of the fact that $\{a, ab, ab^{-1}\}$ is a code: a given word $g \in G$ written with the elementary patterns a, ab, ab^{-1} can be decomposed in only one way as a product of such patterns. By hypothesis, the family $\mathcal{E}_{k+1}(n)$ is made of i.i.d. Bernoulli variables of parameter $h^n(p)$, so the variables of $\mathcal{E}_k(n+1)$ are independent Bernoulli random variables of parameter $h^{n+1}(p)$. Consequently, the process F classifies the density on T'_4. □

Let us mention that from time $n \geq 1$, the field $(X^n_g)_{g\in G}$ is not i.i.d. For example, X^1_1 and $X^1_{ab^{-1}a^{-1}}$ are not independent since both of them depend on X^0_a.

On $T'_{2k} = \langle a_1, \dots, a_k | \cdot \rangle$, one can either apply Prop. 6.3 to obtain a cellular automaton that classifies the density, or define a new CA by the following formula: $F(x)_g = \mathtt{maj}(x_{ga_1}, x_{ga_1a_2}, x_{ga_1a_2^{-1}}, \dots, x_{ga_1a_k}, x_{ga_1a_k^{-1}})$ and check that it also classifies the density.

It is also possible to adapt the above proof to show that the IPS with the same local rule also classifies the density.

6.3.3 A rule that classifies the density on T_3

We now consider the group $T_3 = \langle a, b, c \mid a^2 = b^2 = c^2 = 1 \rangle$.

Theorem 6.4. *The cellular automaton $F : \mathcal{A}^{T_3} \to \mathcal{A}^{T_3}$ defined by:*

$$F(x)_g = \mathtt{maj}(x_{gab}, x_{gac}, x_{gacbc})$$

for any $x \in \mathcal{A}^{T_3}, g \in T_3$, classifies the density.

Proof. The proof is analogous to the previous case. We prove by induction on $n \in \mathbb{N}$ that for any $k \in \mathbb{N}$, that the family $\mathcal{E}_k(n) = \{X^n_{u_1u_2\dots u_k} \mid u_1, u_2, \dots, u_k \in \{ab, ac, acbc\}\}$ consists of independent Bernoulli random variables of parameter $h^n(p)$, the key point being that $\{ab, ac, acbc\}$ is a code. $\square$

Once again, as explained in Prop. 6.3, since we have a solution on T_3, we obtain a CA that classifies the density for any $T_n, n \geq 3$, by applying exactly the same rule. The corresponding IPS on T_n also classifies the density.

The positive rates problem in regular trees. The positive rates problem is defined in Sec. 6.2.3. PCA solving the problem on regular trees appear in the literature [Daw77]. Here, we obtain new examples by considering the CA of Th. 6.3 or the one of Th. 6.4, and by defining its "perturbation" as in equation (6.4). It is not difficult to prove that for ε small enough, the resulting positive-rate PCA is non-ergodic.

Again, this non-ergodicity result complements the density classification result, both of them reflecting strong erasing capacities of the CA (see the discussion at the end of Sec. 6.2.2).

6.4 Classifying the density on $\mathbb{Z}$

The density classification problem on $\mathbb{Z}$ appears as much more difficult than the other cases. We are not aware of any previous result in the literature (even partial), neither for (P)CA nor for IPS.

Below we focus on the synchronous version of the classification problem. First, we show that simple solutions do exist if we slightly relax the formulation of the problem (Sec. 6.4.1). Then we go back to the original problem. We first present a couple of naive (P)CA and show that they do not classify the density (Sec. 6.4.2). We then describe three models, two CA and one PCA, that are conjectured to classify the density (Sec. 6.4.3). We provide some preliminary analytical results (Sec. 6.4.4), as well as experimental investigations of the conjecture by using numerical simulations (Sec. 6.4.5).

In the examples below, the *traffic* cellular automaton, rule 184 according to Wolfram's notation, plays a central role. It is the CA with neighbourhood $\mathcal{N} = \{-1, 0, 1\}$ and local function `traf` defined by:

x, y, z	111	110	101	100	011	010	001	000
$\mathtt{traf}(x, y, z)$	1	0	1	1	1	0	0	0

This CA can be seen as a simple model of traffic flow on a single lane: the cars are represented by 1's moving one step to the right if and only if there are no cars directly in front of them. It is a density-preserving rule.

6.4.1 An exact solution with weakened conditions

On finite rings, several models have been proposed that solve relaxed variants of the density classification problem. We concentrate on one of these models introduced by Kari and Le Gloannec [KLG12]. The original setting is modified since the model operates on an extended alphabet, and the criterium for convergence is also weakened. Modulo this relaxation, it solves the problem on finite rings $\mathbb{Z}_n$. We show the same result on $\mathbb{Z}$.

Proposition 6.5. *Consider the cellular automaton F on the alphabet $\mathcal{B} = \mathcal{A}^2$, with neighbourhood $\mathcal{N} = \{-1, 0, 1\}$, and local function $f = (f_1, f_2)$ defined by:*

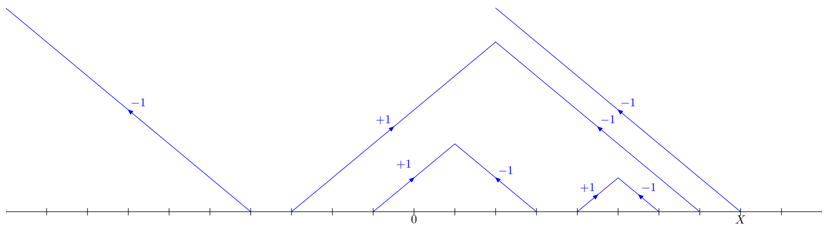
$$f_1(x, y, z) = \mathtt{traf}(x_1, y_1, z_1) \quad ; \quad f_2(x, y, z) = \begin{cases} 0 & \text{if } x_1 = y_1 = 0 \\ 1 & \text{if } x_1 = y_1 = 1 \\ y_2 & \text{otherwise} \end{cases} \tag{6.5}$$

The projections $\mu_p F^n(\mathcal{A}^{\mathbb{Z}} \times \cdot)$ converge to $\delta_{\mathbf{0}}$ if $p < 1/2$ and to $\delta_{\mathbf{1}}$ if $p > 1/2$.

Intuitively, the CA operates on two tapes: on the first tape, it simply performs the traffic rule; on the second tape, what is recorded is the last occurrence of two consecutive zeros or ones in the first tape. If $p < 1/2$, then, on the first tape, there is a convergence to configurations which alternate between patterns of types 0^k and $(10)^\ell$. Consequently, on the second tape, there is convergence to the configuration $\delta_{\mathbf{0}}$. We formalise the argument below.

Proof. Let $T : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ be the traffic CA, see above. Following an idea of Belitsky and Ferrari [BF05], we define the recoding $\psi : \mathcal{A}^{\mathbb{Z}} \to \{-1, 0, 1\}^{\mathbb{Z}}$ by $\psi(x)_i = 1 - x_i - x_{i-1}$. Consider $(\psi \circ T^n(x))_{n \geq 0}$, the recodings of the trajectory of the CA originating from $x \in \{0, 1\}^{\mathbb{Z}}$. There is a convenient alternative way to describe $(\psi \circ T^n(x))_{n \geq 0}$. It corresponds to the trajectories in the so-called *Ballistic Annihilation* model: 1 and -1 are interpreted as particles that we call respectively positive and negative particles. Negative particles move one cell to the left at each time step while positive particles move one cell to the right; and when two particles of different types meet, they annihilate.

Consider the Ballistic Annihilation model with initial condition $\mu_p \psi$ for $p > 1/2$. The density of negative particles is p^2, while the density of positive particles is $(1 - p)^2$. During the evolution, the density of positive particles decreases to 0, while the density of negative particles decreases to $2p - 1$. In particular, the negative particles that will never disappear have density $2p - 1$ [BF05]. We can track back the position at time 0 of the "eternal" negative particles. Let X be the (random) position at initial time of the first eternal particle on the right of cell 0. After time X, the column 0 in the space-time diagram contains only 0 or -1 values. This key point is illustrated in the figure below.



We now go back to the traffic CA with initial condition distributed according to μ_p for $p > 1/2$ and concentrate on two consecutive columns of the space-time diagram. The property tells us that after some almost surely finite time, the columns do not contain the pattern 00.

For the CA defined by equation (6.5) with an initial condition distributed according to a measure μ satisfying $\mu(\cdot \times \mathcal{A}^{\mathbb{Z}}) = \mu_p$ for $p > 1/2$, the above key point gets translated as follows: in any given column of the space-time diagram, after some a.s. finite time, the column contains only the letters $(0, 1)$ or $(1, 1)$. In particular, $\mu_p F^t(\mathcal{A}^{\mathbb{Z}} \times \cdot)$ converges weakly to $\delta_{\mathbf{1}}$ if $p > 1/2$. $\square$

6.4.2 Models that do not classify the density on $\mathbb{Z}$

The first natural idea is to consider the majority rule for some neighbourhood of odd size.

Recall the situation in $\mathbb{Z}^2$: with a symmetric neighbourhood, classification is impossible (Lemma 6.1); with a non-symmetric neighbourhood, classification is possible (Th. 6.1). In $\mathbb{Z}$, Lemma 6.1 still holds, so classification is impossible with a symmetric neighbourhood. We now show that it remains impossible even with a non-symmetric neighbourhood.

Below, we denote by $[x_0 \cdots x_n]_k$ the cylinder of all configurations $y \in \mathcal{A}^{\mathbb{Z}}$ satisfying $y_{k+i} = x_i$ for $0 \leq i \leq n$.

Lemma 6.2. *Consider a cellular automaton F performing the majority rule over a neighbourhood of odd size. Then there exists k, l such that $F([0^k]_0) \subset [0^k]_l$ and $F([1^k]_0) \subset [1^k]_l$. In particular, F does not classify the density.*

Proof. Let the neighbourhood be $\mathcal{N} = \{e_0, \cdots, e_{2n}\}$ with $e_i \in \mathbb{Z}$ and $e_0 < e_1 < \cdots < e_{2n}$. Assume for simplicity that $e_n = 0$ (the general case is treated similarly). Set $k = e_{2n} - e_0 + 1$ and consider $x \in [0^k]_{e_0}$. By definition, $F(x)_i = \texttt{maj}(x_{i+e_0}, \ldots, x_{i+e_{2n}})$, and

$$\begin{array}{rcl} \text{if } \ e_0 \leq i \leq 0, \ \ F(x)_i & = & \texttt{maj}(x_{i+e_0}, \ldots, x_{i+e_{n-1}}, 0, \ldots, 0) = 0, \\ \text{if } \ 0 < i \leq e_{2n}, \ \ F(x)_i & = & \texttt{maj}(0, \ldots, 0, x_{i+e_{n+1}}, \ldots, x_{i+e_{2n}}) = 0. \end{array}$$

So we have $F([0^k]_{e_0}) \subset [0^k]_{e_0}$. Similarly $F([1^k]_{e_0}) \subset [1^k]_{e_0}$. For $p \in (0,1)$, under the probability measure μ_p, an initial configuration will contain both patterns 0^k and 1^k with probability 1. Therefore, the CA cannot classify the density. □

Another natural idea consists in having a model in which the interfaces between monochromatic regions evolve like random walks, leading to an homogenisation of the configuration. Let us show that a direct implementation of this idea does not work.

Consider the PCA with neighbourhood $\mathcal{N} = \{-1, 1\}$, and local function $\varphi(x,y) = (1/2)\delta_x + (1/2)\delta_y$. In words, at each time step, the value of a cell is updated to the value of its left neighbour with probability 1/2 and to the value of its right neighbour with probability 1/2. This is the synchronous version of the Glauber dynamics associated with the Ising model at temperature 0. (In $\mathbb{Z}^2$, the analogous dynamics is conjectured to classify, see the discussion in Sec. 6.2.)

More generally, consider the PCA F with neighbourhood $\mathcal{N} = \{e_1, \ldots e_k\}$, $e_i \in \mathbb{Z}$, parameters $p_1, \ldots, p_k \in (0,1)$ such that $\sum_{i=1}^k p_i = 1$, and local function

$$\varphi(x_{e_1}, \ldots, x_{e_k}) = p_1 \delta_{x_{e_1}} + \cdots + p_k \delta_{x_{e_k}} .$$

Lemma 6.3. *The PCA F does not classify the density.*

Proof. Let $(U_n)_{n \in \mathbb{Z}}$ be a sequence of i.i.d. random variables valued in $\{e_1, \ldots, e_k\}$ with common law: $p_1 \delta_{e_1} + \cdots + p_k \delta_{e_k}$. Let μ be a probability measure on $\mathcal{A}^{\mathbb{Z}}$ and consider a sequence of random variables $(X_n)_{n \in \mathbb{Z}}$ distributed according to μ, and independent of $(U_n)_{n \in \mathbb{Z}}$. Define $Y_n = X_{n+U_n}$ for all $n \in \mathbb{Z}$. By construction, the sequence $(Y_n)_{n \in \mathbb{Z}}$ is distributed according to μF. Assume now that μ is shift-invariant. (The value of $\mu[x]_k$ does not depend on the position k and we denote it by $\mu[x]$.) We have

$$\begin{array}{rcl} \mu F[1] = \mathbb{P}\{Y_0 = 1\} & = & \displaystyle\sum_{i=1}^{k} \mathbb{P}\{Y_0 = 1, U_0 = e_i\} = \sum_{i=1}^{k} \mathbb{P}\{X_{e_i} = 1, U_0 = e_i\} \\ & = & \displaystyle\sum_{i=1}^{k} \mathbb{P}\{X_{e_i} = 1\}\mathbb{P}\{U_0 = e_i\} = \sum_{i=1}^{k} \mu[1] p_i = \mu[1] . \end{array}$$

So the density of 1 is preserved by the dynamics, and F does not classify the density. The expected behaviour is that homogenisation occurs leading to: $\mu_p F^n \xrightarrow[n \to \infty]{w} (1-p)\delta_{\mathbf{0}} + p\delta_{\mathbf{1}}$. □

The behaviour is thus the same as for the one-dimensional voter model IPS.

6.4.3 Density classifier candidates on $\mathbb{Z}$

We now propose three models, two CA (GKL and Kari-traffic) and one PCA (majority-traffic), that are candidates to classify the density on $\mathbb{Z}$.

All three of them perform well with respect to the density classification on finite rings. Figures 6.6 and 6.7 illustrate this point with space-time diagrams for the ring $\mathbb{Z}/149\mathbb{Z}$.

All three of them have the *eroder* property: if the initial configuration contains only a finite number of ones (resp. zeros), then it reaches **0** (resp. **1**) in finite time (almost surely for the PCA). Proofs have been given by Gonzaga de Sá and Maes [GdSM92] for GKL and by Kari and Le Gloannec [KLG12] for Kari-traffic. For majority-traffic, $\alpha < 1/2$, a proof could be worked out by considering the interfaces between regions (all-black, all-white, and checkerboard) as particles.

GKL cellular automaton. The Gács-Kurdyumov-Levin (GKL) cellular automaton [GKL78] is the CA with neighbourhood $\mathcal{N} = \{-3, -1, 0, 1, 3\}$ defined by: for $x \in \mathcal{A}^{\mathbb{Z}}, i \in \mathbb{Z}$,

$$\texttt{Gkl}(x)_i = \begin{cases} \texttt{maj}(x_i, x_{i+1}, x_{i+3}) & \text{if } x_i = 1 \\ \texttt{maj}(x_i, x_{i-1}, x_{i-3}) & \text{if } x_i = 0. \end{cases} \tag{6.6}$$

Kari-traffic cellular automaton. The Kari-Le Gloannec traffic rule [KLG12], that we shorten as Kari-traffic CA and denote by `Kari`, is the CA of neighbourhood $\mathcal{N} = \{-3, -2, -1, 0, 1, 2, 3\}$ defined by: for $x \in \mathcal{A}^{\mathbb{Z}}$,

$$\texttt{Kari}(x) = \Phi \circ \texttt{Traf}(x)\,,$$

where `Traf` is the traffic CA, that is the global function associated with `traf`, and where Φ is the CA defined by: for $x \in \mathcal{A}^{\mathbb{Z}}, i \in \mathbb{Z}$,

$$\Phi(x)_i = \begin{cases} 0 & \text{if } (x_{i-2}, x_{i-1}, x_i, x_{i+1}) = 0010 \\ 1 & \text{if } (x_{i-1}, x_i, x_{i+1}, x_{i+2}) = 1011 \\ x_i & \text{otherwise.} \end{cases} \tag{6.7}$$

The Kari-traffic rule is closely related to Kůrka's modified version of GKL [Kůr03].

Both GKL and Kari-traffic are symmetric when swapping 0 and 1 and right and left simultaneously.

Majority-traffic probabilistic cellular automaton. The majority-traffic PCA of parameter $\alpha \in (0, 1)$ is the PCA of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$ and local function:

$$(x, y, z) \mapsto \alpha\, \delta_{\texttt{maj}(x,y,z)} + (1 - \alpha)\, \delta_{\texttt{traf}(x,y,z)}.$$

In words, at each time step, we choose, independently for each cell, to apply the majority rule with probability α and the traffic rule with probability $1 - \alpha$ (see Fig. 6.7).

The majority-traffic PCA has been introduced by Fatès [Fat11, Fat13], who has proved the following: for any $n \in \mathbb{N}$ and any $\varepsilon > 0$, there exists a value $\alpha_{n,\varepsilon}$ of the parameter such that on $\mathbb{Z}_n$, the PCA converges to the right uniform configuration with probability greater than $1 - \varepsilon$.

Conjecture. The GKL CA, the Kari-traffic CA, and the majority-traffic PCA with $0 < \alpha < \alpha_c$ (for some $0 < \alpha_c \leq 1/2$) classify the density.

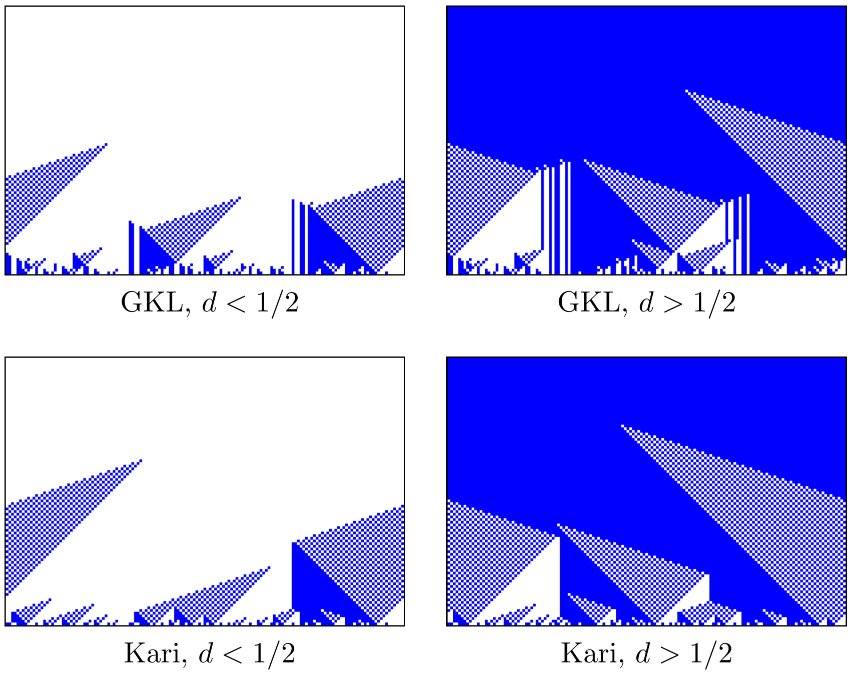


Figure 6.6: Two space-time diagrams of GKL (top) and Kari-traffic (bottom) on $\mathbb{Z}/149\mathbb{Z}$. The density of 1 in the initial condition is $70/149$ (left) and $77/149$ (right).

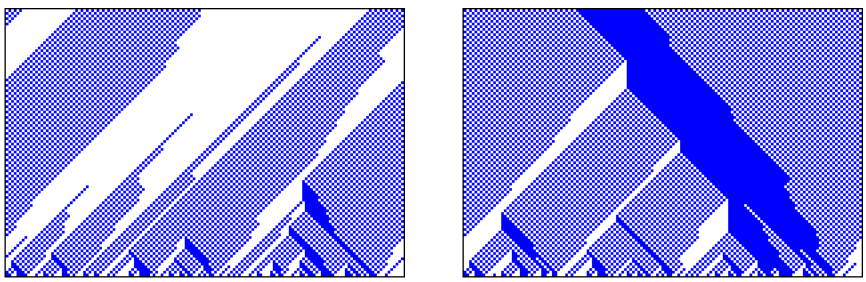


Figure 6.7: Two space-time diagrams of the majority-traffic PCA for $\alpha = 0.1$ on the ring $\mathbb{Z}/149\mathbb{Z}$. Both diagrams have the same initial condition with a density of 1 equal to $70/149$. The right diagram corresponds to a rare event: evolution towards a configuration with only 1's, starting from a majority of 0's.

6.4.4 Invariant Measures

Following ideas developed by Kůrka [Kůr03], we can give a precise description of the invariant measures of the three above models.

Let $x = (01)^{\mathbb{Z}}$ be the configuration defined by: $\forall n \in \mathbb{Z}$, $x_{2n} = 0$, $x_{2n+1} = 1$. The configuration $(10)^{\mathbb{Z}}$ is defined similarly.

Proposition 6.6. *For the majority-traffic PCA and for the Kari-traffic CA, the extremal invariant measures are $\delta_{\mathbf{0}}, \delta_{\mathbf{1}}$, and $(\delta_{(01)^{\mathbb{Z}}} + \delta_{(10)^{\mathbb{Z}}})/2$. For GKL, on top of these three measures, there exist extremal invariant measures of density p for any $p \in [1/3, 2/3]$.*

Proof. **Majority-traffic.** Let us consider the majority-traffic PCA P of parameter $\alpha \in (0,1)$. Let μ be any shift-invariant measure. An exhaustive search shows that if at time 1, we observe the cylinder $[100]_0$ then there are only eight possible cylinders of size 5 at time 0, that are:

$$[01100]_{-1}, [10000]_{-1}, [10001]_{-1}, [10010]_{-1},$$
$$[10100]_{-1}, [11000]_{-1}, [11001]_{-1}, [11100]_{-1}.$$

Since the measure μ is shift-invariant, the probability $\mu([x_0 \cdots x_n]_k)$ does not depend on k and we denote it by $\mu[x_0 \cdots x_n]$. If we weight each of the above cylinder by the probability to reach $[100]_0$ from it, we obtain the following expression:

$$\begin{aligned}\mu P[100] = {} & \alpha(1-\alpha)\mu[01100] + (1-\alpha)\mu[10000] + (1-\alpha)\mu[10001] + (1-\alpha)\mu[10010] \\ & + \alpha\mu[10100] + \alpha^2\mu[11000] + \alpha^2\mu[11001] + \alpha(1-\alpha)\mu[11100].\end{aligned}$$

Gathering the terms with the same coefficient, we have:

$$\begin{aligned}\mu P[100] &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha\mu[10100] + \alpha(1-\alpha)\mu[1100] + \alpha^2\mu[1100] \\ &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha\mu[10100] + \alpha\mu[1100].\end{aligned}$$

Some more rearrangements provide:

$$\begin{aligned}\mu P[100] &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha(\mu[100] - \mu[00100]) \\ &= \mu[100] - (1-\alpha)\mu[10011] - \alpha\mu[00100].\end{aligned}$$

This proves that the sequence $(\mu P^n[100])_{n \geq 0}$ is non-increasing. From now on, let us assume that $\mu P = \mu$. Then, $\mu[10011] = \mu[00100] = 0$.

Let us consider the cylinder $[10^n 0011]$ for some $n \geq 2$. If we apply the majority rule on each cell except on the second cell from the left, then after n iterations, we reach the cylinder $[10011]$. Since this occurs with a positive probability, we obtain that for any $n \geq 0, \mu[10^n 0011] = 0$. This provides: $\mu[0011] = \mu[00011] = \mu[000011] = \ldots = \mu[0^n 11]$ for any $n \geq 2$. Consequently, $\mu[0011] = 0$. From a cylinder of the form $[00(10)^n 11]$, if we choose to apply the majority rule on each cell, then we reach the cylinder $[0011]$ in n steps. Thus, $\mu[00(10)^n 11] = 0$ for any $n \geq 0$. It follows that μ can be written as the sum $\mu = \mu_0 + \mu_1$ of two invariant measures, where μ_0 charges only the subshift Σ_0 and μ_1 the subshift Σ_1 with

$$\Sigma_0 = \{x \in \mathcal{A}^{\mathbb{Z}} \mid \forall k \in \mathbb{Z}, x_k x_{k+1} \neq 00\}, \qquad \Sigma_1 = \{x \in \mathcal{A}^{\mathbb{Z}} \mid \forall k \in \mathbb{Z}, x_k x_{k+1} \neq 11\}. \quad (6.8)$$

Let us assume that $\mu[00] = 0$ (which is the case for μ_0). In the same way that we have computed $\mu P[100]$, we can compute $\mu P[11]$, and we obtain:

$$\begin{aligned}\mu P[11] &= \alpha\mu[0110] + \alpha\mu[1110] + \alpha\mu[1101] + \mu[1011] + \mu[0111] + \mu[1111] \\ &= \alpha\mu[110] + \alpha\mu[1101] + \mu[11] - \mu[0011] \\ &= \mu[11] + \alpha\mu[110] + \alpha\mu[1101].\end{aligned}$$

By hypothesis, $\mu P = \mu$, so that the last equality implies that $\mu[110] = 0$.

In all cases, if μ is a shift-invariant measure such that $\mu P = \mu$, then $\mu[00] = \mu(\mathbf{0}), \mu[11] = \mu(\mathbf{1})$ and $\mu[01] = \mu[10] = \mu((01)^{\mathbb{Z}}) = \mu((10)^{\mathbb{Z}})$.

Kari-traffic. If at time 1, we observe the pattern 100 at position 0, then, at time 0, this same pattern was present at position -1. This can be checked by systematic inspection. In the same way, if, at time 1, we observe the pattern 110 at position 0, then, at time 0, this same pattern was present at position 1.

Let μ be a shift-invariant measure such that $\mu K = \mu$, where $K = \texttt{Kari}$. A consequence of the above results on the patterns 100 and 110 is that: $\mu K[1100] = 0$ and $\mu K^{n+1}[110x100] = 0$ for any $n \geq 0$ and any $x \in \mathcal{A}^n$. But since $\mu K = \mu$, we obtain $\mu[110x100] = 0$ for any word x. Like for majority-traffic PCA, we can write $\mu = \mu_0 + \mu_1$ where μ_0 and μ_1 are two invariant measures defined on the subshifts Σ_0 and Σ_1, see equation (6.8).

Let us consider a configuration of Σ_0, that is, without the pattern 00. By the traffic rule, each 0 of the configuration will move one cell to the left. Then by rule Φ (see equation (6.7)), if a 0 is at distance greater than 2 from the next 0 on its right, it is erased. The result follows.

GKL. Any word $x \in \mathcal{A}^{\mathbb{Z}}$ which is a concatenation of the patterns $u = 001$ and $v = 011$ is a fixed point of the GKL cellular automaton: if $x_n = 0$, then either $x_{n-1} = 0$ or $x_{n-3} = 0$ so that $F(x)_n = 0$ and if $x_n = 1$, then either $x_{n+1} = 1$ or $x_{n+3} = 1$ so that $F(x)_n = 1$. As a consequence, GKL has extremal invariant measures of density p for any $p \in [1/3, 2/3]$. □

To summarise, majority-traffic and Kari-traffic have a simpler set of invariant measures. It does not rule out GKL as a candidate for solving the density classification task, but rather indicates that it could be easier to prove the result for majority-traffic or Kari-traffic.

The positive rates problem in $\mathbb{Z}$. Recall that the positive rates problem is defined in Sec. 6.2.3. On $\mathbb{Z}$, it had been a long standing conjecture that all positive-rate PCA and IPS are ergodic.

The GKL CA, see equation (6.6), was originally introduced as a candidate to solve the positive rates problem, with the conjecture that its perturbed version may be non-ergodic [GKL78]. It is still unknown if it is the case or not, although the belief seems now to be that it is ergodic [GdSM92, Par97].

Nevertheless, the positive rates conjecture is today known to be false. Gács suggested a counter-example in 1986 [Gác86], and published the full proof in 2001 [Gác01]. It is a very complex counter-example, with an alphabet of cardinality at least 2^{18} [Gác01, Gra01].

To summarise, in $\mathbb{Z}$, there is no known model that classifies the density, and there is no known "simple" model that solves the positive rates problem. This reflects the difficulty to build a model in $\mathbb{Z}$ with strong erasing properties.

6.4.5 Experimental results

Let us recall the arguments backing up the conjecture of Sec. 6.4.3. First, the three models have the eroder property. Second, they classify reasonably well on a finite ring.

To go further, we perform some numerical experimentations. Our approach is to test if the proportion of good classification on a finite ring converges to one as the size of the ring increases. Indeed, it is reasonable to believe that there is a relationship between this last property and the ability to classify on $\mathbb{Z}$.

More precisely, we proceed as follows. We fix a rule (GKL, Kari-traffic, or majority-traffic for $\alpha = 0.1$) and a parameter $p \in (0, 1/2)$. We consider different rings of odd sizes ranging from 101 to 2001. For each size, we perform 10^5 experiments, by choosing each time a new initial configuration according to the Bernoulli product measure μ_p, that is, we assign to each cell the value 1 with a probability p and the value 0 with probability $1 - p$. We record the proportion of good classifications among the 10^5 experiments. We denote this proportion by $Q(n)$ where n is the ring size. Let $d(x)$ be the proportion of 1 in the initial configuration x distributed according to μ_p. We may have $d(x) > 1/2$, although $E[d(x)] = p < 1/2$. We have a "good classification" for x if there is convergence to $\mathbf{0}$ when $d(x) < 1/2$ and to $\mathbf{1}$ when $d(x) > 1/2$.

The results are reported in Fig. 6.8. For each rule, we consider five different values for the parameter p, ranging from 0.45 to 0.49. For each rule and each value of the parameter, the plot is consistent with the hypothesis that $Q(n)$ converges to 1. However, when p approaches 1/2, the ring size n needed for $Q(n)$ to attain a certain quality level increases dramatically.

On each of the plots, we observe an initial decrease of $Q(n)$ followed by an increase for n large. For $p = 0.49$, the point of inflexion becomes hardly visible. Our explanation is that for small ring sizes, the dispersion of the actual density $d(x)$ is higher and covers values far from 1/2 for which the classification is easier.

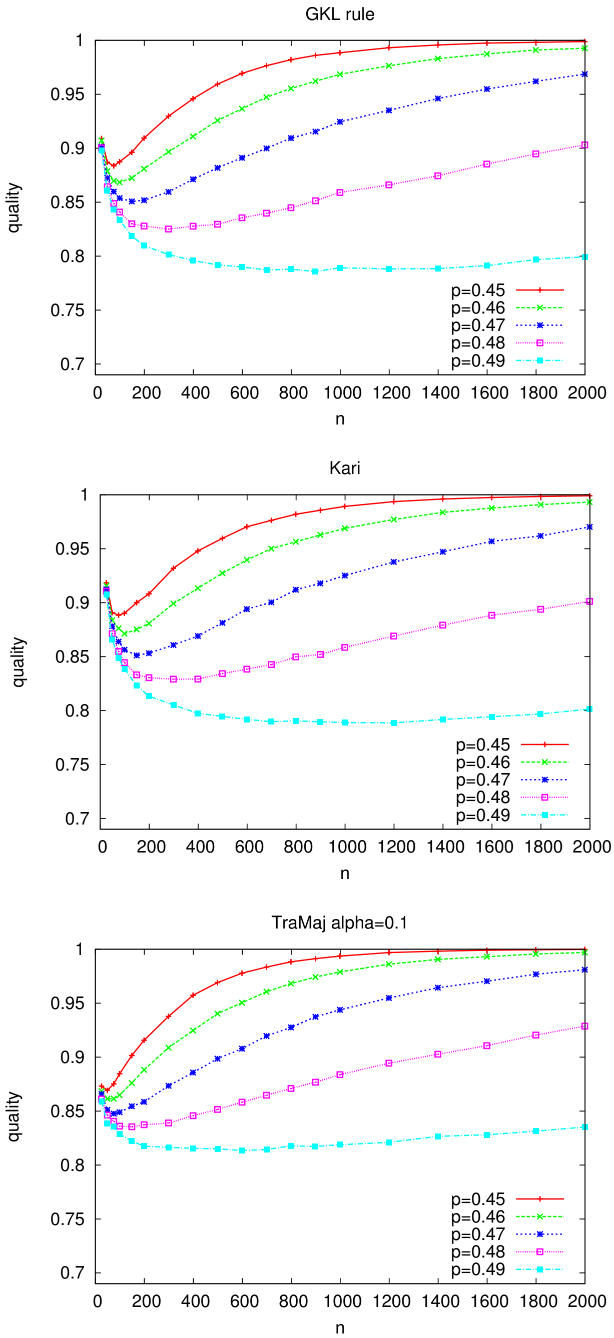


Figure 6.8: Experimental determination of the quality of classification $Q(n)$ as a function of ring size n. Cells are initialised with a probability p to be in state 1. Each point represents an average computed on 100 000 experiments (simulations made by Fatès with his software FiatLux).

Part III

Random walks and measures of maximal entropy

Chapter 7

Random walks and Markov-multiplicative measures

He walked with his shoulders very straight and kept his hands always stuffed down into his pockets. His grey eyes seemed to take in everything around him, and in his face there was still the look of peace that is seen most often in those who are very wise or very sorrowful. He was always glad to stop with anyone who wished his company. For after all he was only walking and going nowhere.

– Carson Mc Cullers, *The Heart is a Lonely Hunter*

Contents

We study random walks on infinite groups (or monoids) of free product type. The asymptotic behaviour of these random walks is described by the harmonic measure, giving the direction taken by the walk in its escape to infinity. Mairesse and Mathéus have proved that this measure has a Markov-multiplicative structure, and have given an in-depth study of the case of free product of groups [Mai05, MM07]. We give a general frame to describe the parameters of this measure, through a system of equations involving generating functions of weighted paths in each group. Our approach has some similarities with the work of Gilch, and we recover some of his results [Gil07]. The specificity is that we exploit the special combinatorial structure of the harmonic measure, known to be Markov-multiplicative.

7.1 Random walks on free products of groups

7.1.1 Free products of groups

Let $G_1, \ldots, G_n$ be n countable groups, pairwise disjoint but possibly isomorphic.

We denote by $\varepsilon_1, \ldots, \varepsilon_n$ the neutral elements of $G_1, \ldots, G_n$ and we set $\Sigma_i = G_i \setminus \{\varepsilon_i\}$, and $\Sigma = \cup_{i=1}^n \Sigma_i$.

Let us denote by $G = G_1 * \ldots * G_n$, the *free product* of $G_1, \ldots, G_n$. By definition, G is the set of words on Σ, equiped with the operation $*$ of concatenation with possible simplification within the groups G_i. Let us give a formal definition of this operation.

We denote by $\tau : \Sigma \to \{1, \ldots, n\}$ the application that maps an element $\alpha \in \Sigma$ to the unique integer $i \in \{1, \ldots, n\}$ such that $\alpha \in \Sigma_i$. We say that $\tau(\alpha)$ is the *type* of the element α. For $\alpha \in \Sigma$, the set of successors of α is given by:

$$\mathcal{S}(\alpha) = \{\beta \in \Sigma, \tau(\alpha) \neq \tau(\beta)\}.$$

We define the set of normal form words by:

$$L = \{u_1 \ldots u_k \in \Sigma^* \,; \forall i \in \{1, \ldots, k-1\}, u_{i+1} \in \mathcal{S}(u_i)\}.$$

The set L is a particulat subshift of finite type on the alphabet Σ.

The free group $G = G_1 * \ldots * G_n$ is the group with set of elements L, unit element ε (empty word), and group law $*$ defined recursively by:

$$u_1 \ldots u_k * v_1 \ldots v_l = \begin{cases} u_1 \ldots u_{k-1} u_k v_1 v_2 \ldots v_l & \text{if} \quad \tau(u_k) \neq \tau(v_1) \\ u_1 \ldots u_{k-1} (u_k \cdot v_1) v_2 \ldots v_l & \text{if} \quad \tau(u_k) = \tau(v_1), u_k \neq v_1^{-1} \\ u_1 \ldots u_{k-1} * v_2 \ldots v_l & \text{if} \quad u_k = v_1^{-1} \end{cases},$$

where in the second case, $u_k \cdot v_1$ is the product in $G_{\tau(u_k)}$ of u_k and v_1.

The *length* on an element $g \in G$ is given by:

$$|g|_\Sigma = \min\{k \in \mathbb{N} \,; g = u_1 * \ldots * u_k, u_i \in \Sigma\}.$$

The empty word ε is the only element of length 0.

Example 7.1. Let $G_1 = \mathbb{Z}/2\mathbb{Z} = \{1, a\}$ and $G_2 = \mathbb{Z}/3\mathbb{Z} = \{1, b, b^{-1}\}$. We have $\Sigma = \{a, b, b^{-1}\}$ and for example

$$ab^{-1}aba * ab = ab^{-1}ab^{-1}; \qquad aba * b^{-1}ab = abab^{-1}ab; \qquad (ab^{-1}ab)^{-1} = b^{-1}aba.$$

Example 7.2. Let $G_1 = \mathbb{Z}^2 = \langle a, b | ab = ba \rangle$ and $G_2 = \mathbb{Z} = \langle c | - \rangle$. We have $\Sigma = \{a^i b^j ; (i,j) \in \mathbb{Z}^2 \setminus \{(0,0)\}\} \cup \{c^k ; k \in \mathbb{Z} \setminus \{0\}\}$, and for example:

$$(a^2 b^{-3}) c^{12} (a^{-1}) * (a^2 b) c^{-2} = (a^2 b^{-3}) c^{12} (ab) c^{-2}.$$

An element of $\mathbb{Z}^2 * \mathbb{Z}$ can be represented by a heap of different pieces. Precisely, we introduce six different types of pieces: the pieces of type a or a^{-1} (resp. b or b^{-1}) have length 1 and can occupy a left (resp. right) position, and the pieces of type c or c^{-1} have length 2. If a piece lands directly on top of a piece of inverse type, the two pieces annihilate, see Fig. 7.1.

7.1.2 Random walks and the harmonic measure

We keep the same notations as in the previous subsection. Let us consider a probability distribution $p = (p_\alpha)_{\alpha \in \Sigma}$ on Σ, such that for any $i \in \{1, \ldots, n\}$, $\{\alpha \in \Sigma_i \,; p_\alpha > 0\}$ generates the group G_i. We consider the random walk (G, p) which consists, at each time-step, in jumping from $\omega \in G$ to $\omega\alpha \in G$ with probability p_α. Precisely, let $(x_k)_{k \geq 0}$ be an i.i.d sequence of random variables of distribution p. We set $X_0 = \varepsilon$ and

$$X_{k+1} = X_k * x_k = x_0 * x_1 * \ldots * x_k.$$

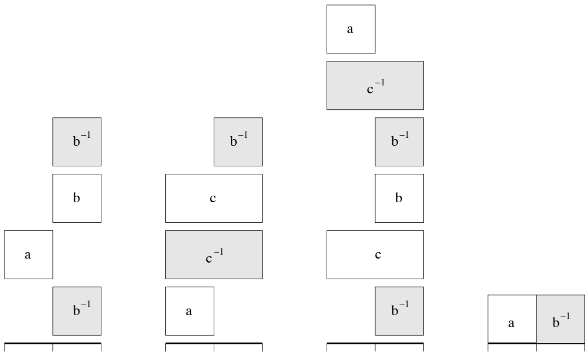


Figure 7.1: Different heaps representing the same element ab^{-1} of $\mathbb{Z}^2 * \mathbb{Z}$.

Then, $(X_k)_{k\geq 0}$ is a realisation of the random walk (G, p).

A random walk on the free product G corresponds to a particular random walk on the Cayley graph of G. For instance, if we consider the free product of cyclic groups $G = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$, as in Ex. 7.1, and set $p_b = p_{b^{-1}} \in (0, 1/2)$, with $p_a = 1 - 2p_b$, the walk consists in choosing independently at each time step to move with probability p_b along one of the two edges of the triangle on which the walker stands, or to follow the edge going out of that triangle with probability p_a, see Fig. 7.2.

In the same way, for Ex. 7.2, if we set $p_a + p_{a^{-1}} + p_b + p_{b^{-1}} + p_c + p_{c^{-1}} = 1$, the random walk can be interpreted in terms of random heaps: at each time step, one of the six pieces is chosen according to the probability p and falls onto the heap.

Since, $|u * v|_\Sigma \leq |u|_\Sigma + |v|_\Sigma$, Guivarc'h [Gui80] observed that a simple corollary of Kingman's subbadditive ergordic theorem [Kin73] is the existence of a constant $\gamma \geq 0$ such that almost surely and in L^p, for all $1 \leq p < \infty$,

$$\lim_{k\to\infty} \frac{|X_k|_\Sigma}{k} = \gamma.$$

The constant γ is called the *drift*. Intuitively, γ is the speed of escape to infinity of the walk.

In all the following, we assume that G is a non-trivial free product, different from $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$. In that case, any random walk living on the whole group is transient and has a drift γ that is strictly positive [Gui80, Woe00]. Moreover, we have the following theorem, which is proved for example in the survey of Ledrappier [Led01].

A measure μ on L^∞ is called p-stationary if it is invariant by left-multiplication by an element of Σ distributed according to p. This can be written:

$$\mu = \sum_{\alpha\in\Sigma} p_\alpha \cdot (\alpha\mu),$$

where $\alpha\mu$ is the measure obtained when left-multiplying by the letter α (with possible simplification) a word distributed according to μ.

Theorem 7.1 ([Led01]). *Let $L^\infty = \{u_1u_2\ldots \in \Sigma^{\mathbb{N}}; \forall i \in \mathbb{N}, u_{i+1} \in \mathcal{S}(u_i)\}$. There exists a random variable X_∞ valued in L^∞, such that, almost surely,*

$$\lim_{k\to\infty} X_k = X_\infty,$$

in the sense that the length of the common prefix of X_n and X_∞ tends to infinity.

Furthermore, the law μ^∞ of X_∞ is stationary and it is the only p-stationary measure on L^∞. It is called the harmonic measure *associated to the random walk (G, p).*

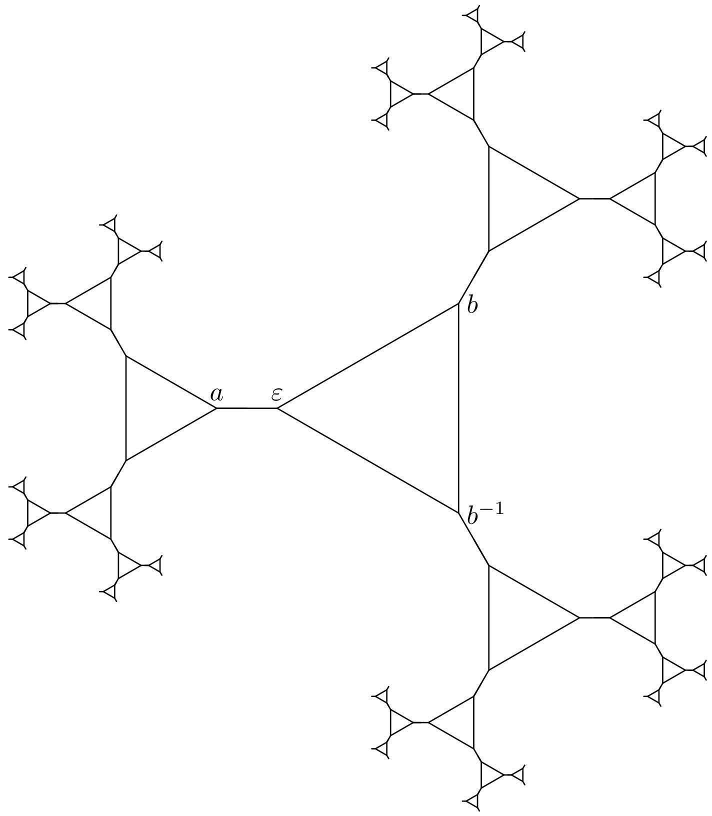


Figure 7.2: Cayley graph of $G = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$.

7.2 Description of the harmonic measure

7.2.1 Markov-multiplicative measures

To begin with, let us introduce the notion of Markov-multiplicative measure.

Definition 7.1. Let μ be a measure on $\Sigma^{\mathbb{N}}$. We say that μ is a *Markov-multiplicative measure*, if there exists a probability measure x on Σ, such that for any $u_1 \dots u_k \in L$,

$$\mu(u_1 \dots u_k \Sigma^{\mathbb{N}}) = \frac{x(u_1) \dots x(u_k)}{x(\mathcal{S}(u_1)) \dots x(\mathcal{S}(u_{k-1}))}. \tag{7.1}$$

A Markov-multiplicative measure is a Markov measure, given by the transition matrix P of dimension $\Sigma \times \Sigma$ given by:

$$P_{u,v} = \begin{cases} x(v)/x(\mathcal{S}(u)) \text{ if } v \in \mathcal{S}(u), \\ 0 \text{ otherwise.} \end{cases}$$

Observe that in general, we have $xP \neq x$, so that it is not stationary.

We want to describe the distribution μ_∞, which gives the direction in which $(X_k)_{k \geq 0}$ goes to infinity. Th. 7.1 tells us that it amounts to finding a p-stationary measure for the process. The following result reduces the search domain.

Proposition 7.1 ([Mai05])**.** *The harmonic measure of a random walk on a free product of groups is a Markov-multiplicative measure.*

7.2.2 Traffic equations

For $k \in \{1, \ldots, n\}$ and $\alpha \in \Sigma_k$, let us define the generating function counting the weighted paths of first visit to α from ε_k in the group G_k, that is:

$$f_k(\alpha, z) = \sum_{\substack{\ell \geq 1 \\ v_1, \ldots, v_\ell \in \Sigma_k \\ v_1 \ldots v_\ell = \alpha \\ v_i \ldots v_\ell \neq \varepsilon_k \ (1 \leq i \leq \ell)}} p_{v_1} \ldots p_{v_\ell} \, z^\ell.$$

We also define: $F_k(z) = \sum_{\alpha \in \Sigma_k} f_k(\alpha, z)$.

Let us set $p_k = p_{\Sigma_k} = \sum_{\alpha \in \Sigma_k} p_\alpha$.

Proposition 7.2. *The harmonic measure of the random walk (G, p) is the Markov-multiplicative measure associated to the distribution x on Σ given by:*

$$x(\alpha) = \frac{1}{1 + F_k(B_k)} f_k(\alpha, B_k),$$

where $(B_1, \ldots, B_n)$ is the unique positive solution of the system given by the following equations, for $1 \leq k \leq n$:

$$\frac{F_k(B_k)}{1 + F_k(B_k)} = \frac{B_k \, p_k - B_k \Big(\sum_{i \neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k - 1)}{(n-1)B_k} \Big)}{1 - B_k \Big(\sum_{i \neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k - 1)}{(n-1)B_k} \Big)}.$$

Proof. Let us consider a Markov-multiplicative measure μ, defined as in (7.1). Then the measure μ is p-stationary if and only if the distribution x satisfies the following traffic equations: for any $k \in \{1, \ldots, n\}$ and any $\alpha \in \Sigma_k$,

$$\begin{aligned} x(\alpha) &= p_\alpha \, x(\mathcal{S}(\alpha)) + \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u \, x(u^{-1}\alpha) + \sum_{u \in \mathcal{S}(\alpha)} p_u \frac{x(u^{-1})x(\alpha)}{x(\mathcal{S}(u^{-1}))}. \\ &= p_\alpha \, (1 - x(\Sigma_k)) + \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u \, x(u^{-1}\alpha) + \sum_{i \neq k} \sum_{u \in \Sigma_i} p_u \frac{x(u^{-1})x(\alpha)}{1 - x(\Sigma_i)}. \end{aligned}$$

For $j \in \{1, \ldots, n\}$, let us set

$$A_j = \sum_{u \in \Sigma_j} p_u \, x(u^{-1}), \qquad \text{and} \qquad B_j = \frac{1}{1 - \sum_{i \neq j} \frac{A_i}{1 - x(\Sigma_i)}}.$$

It follows from the definition of B_j, that:

$$\sum_{i \neq j} \frac{A_i}{1 - x(\Sigma_i)} = \frac{B_j - 1}{B_j},$$

so that:

$$\begin{pmatrix} 0 & \frac{1}{1-x(\Sigma_2)} & \frac{1}{1-x(\Sigma_3)} & \cdots & \frac{1}{1-x(\Sigma_n)} \\ \frac{1}{1-x(\Sigma_1)} & 0 & \frac{1}{1-x(\Sigma_3)} & \cdots & \frac{1}{1-x(\Sigma_n)} \\ \vdots & & & \ddots & \\ \frac{1}{1-x(\Sigma_1)} & \frac{1}{1-x(\Sigma_2)} & \cdots & \frac{1}{1-x(\Sigma_{n-1})} & 0 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} \frac{B_1 - 1}{B_1} \\ \frac{B_2 - 1}{B_2} \\ \vdots \\ \frac{B_n - 1}{B_n} \end{pmatrix}.$$

The left matrix can also be written:

$$\begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & & & & \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{1-x(\Sigma_1)} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{1-x(\Sigma_2)} & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 0 & \frac{1}{1-x(\Sigma_n)} \end{pmatrix}.$$

By inverting this product, we obtain:

$$A_j = (1 - x(\Sigma_j))\Big(\sum_{i \neq j} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_j - 1)}{(n-1)B_j}\Big). \tag{7.2}$$

For $\alpha \in \Sigma_k$, the traffic equations can be rewritten:

$$x(\alpha) = p_\alpha\,(1 - x(\Sigma_k)) + \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u\, x(u^{-1}\alpha) + \sum_{i \neq k} \frac{A_i}{1 - x(\Sigma_i)} x(\alpha).$$

Thus,

$$x(\alpha)\Big(1 - \sum_{i \neq k} \frac{A_i}{1 - x(\Sigma_i)}\Big) = p_\alpha\,(1 - x(\Sigma_k)) + \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u\, x(u^{-1}\alpha).$$

It follows that:

$$x(\alpha) = B_k\, p_\alpha (1 - x(\Sigma_k)) + B_k \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u\, x(u^{-1}\alpha). \tag{7.3}$$

Let us recall the notation $p_k = p_{\Sigma_k} = \sum_{\alpha \in \Sigma_k} p_\alpha$. We have:

$$\begin{aligned} \sum_{\alpha \in \Sigma_k} \sum_{u \in \Sigma_k \setminus \{\alpha\}} p_u\, x(u^{-1}\alpha) &= \sum_{u \in \Sigma_k} \sum_{\alpha \in \Sigma_k \setminus \{u\}} p_u\, x(u^{-1}\alpha) \\ &= \sum_{u \in \Sigma_k} p_u (x(\Sigma_k) - x(u^{-1})) \\ &= p_k x(\Sigma_k) - \sum_{u \in \Sigma_k} p_u\, x(u^{-1}) \\ &= p_k\, x(\Sigma_k) - A_k. \end{aligned}$$

Thus, summing equation (7.3) on $\alpha \in \Sigma_k$ provides:

$$\begin{aligned} x(\Sigma_k) &= B_k\, p_k\, (1 - x(\Sigma_k)) + B_k\, (p_k\, x(\Sigma_k) - A_k) \\ &= B_k\, (p_k - A_k), \end{aligned}$$

Using the above expression (7.2) of A_k, we obtain:

$$x(\Sigma_k) = B_k\Big(p_k - (1 - x(\Sigma_k)\Big)\Big(\sum_{i \neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k - 1)}{(n-1)B_k}\Big),$$

so that gathering the terms $x(\Sigma_k)$, we have finally:

$$x(\Sigma_k) = \frac{B_k\; p_k - B_k\Big(\sum_{i \neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k - 1)}{(n-1)B_k}\Big)}{1 - B_k\Big(\sum_{i \neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k - 1)}{(n-1)B_k}\Big)}.$$

If B_k and $x(\Sigma_k)$ are supposed to be known, the equations (7.3) for $\alpha \in \Sigma_k$ give a system of linear equations.

One can check that a solution of the system of equations (7.3) is given by $x(\alpha) = (1 - x(\Sigma_k))\, f_k(\alpha, B_k)$. By summing over $\alpha \in \Sigma_k$, we obtain: $x(\Sigma_k) = (1 - x(\Sigma_k))F_k(B_k)$, that is:

$$x(\Sigma_k) = \frac{F_k(B_k)}{1 + F_k(B_k)}.$$

Finally, we have:

$$x(\Sigma_k) = \frac{F_k(B_k)}{1 + F_k(B_k)} = \frac{B_k\, p_k - B_k\Big(\sum_{i\neq k} \frac{B_i - 1}{(n-1)B_i} - \frac{(n-2)(B_k-1)}{(n-1)B_k}\Big)}{1 - B_k\Big(\sum_{i\neq k} \frac{B_i-1}{(n-1)B_i} - \frac{(n-2)(B_k-1)}{(n-1)B_k}\Big)}.$$

This gives a system of k equations and k unknowns. If we are able to find $B_1, \dots, B_n$ then we have a description of the measure μ^∞, since for $\alpha \in \Sigma_k$, $x(\alpha) = (1 - x(\Sigma_k))\, f_k(\alpha, B_k)$ where $x(\Sigma_k)$ and B_k are known.

By Prop. 7.1, the harmonic measure of the random walk is Markov-multiplicative. So, this system has at least one solution, which provides the harmonic measure. Conversely, any positive solution provides a stationary measure on L^∞ and is thus the (unique) harmonic measure, by Th. 7.1. This proves the characterisation of the harmonic measure given in the proposition. □

For a free product of two groups, we obtain for example the following system of equations:

$$\begin{cases} \frac{F_1(B_1)}{1+F_1(B_1)} = \frac{p_1 + \frac{1}{B_2} - 1}{\frac{1}{B_1} + \frac{1}{B_2} - 1} \\ \frac{F_2(B_2)}{1+F_2(B_2)} = \frac{p_2 + \frac{1}{B_1} - 1}{\frac{1}{B_1} + \frac{1}{B_2} - 1}. \end{cases}$$

In the case where the n groups $G_1, \dots, G_n$ are all isomorphic, with same probabilities allocated to associated elements, then one has a single equation to solve, which is: $\frac{F(B)}{1+F(B)} = \frac{1}{n}$, where $F = F_1 = \ldots = F_n$. That is,

$$F(B) = \frac{1}{n-1}.$$

Then, we obtain for any $\alpha \in \Sigma$,

$$x(\alpha) = \frac{n-1}{n}\, f(\alpha, B).$$

7.2.3 Examples of computations of the generating functions

If the group G_k is commutative, then for $\alpha \in G_k$, the generating function $f(\alpha, z)$ is equal to:

$$f(\alpha, z) = \sum_{\substack{\ell \geq 1 \\ v_1, \dots, v_\ell \in \Sigma_k \\ v_1 \dots v_\ell = \alpha \\ v_i \dots v_\ell \neq \varepsilon_k\ (1 \leq i \leq \ell)}} p_{v_1} \dots p_{v_\ell}\, z^\ell = \sum_{\substack{\ell \geq 1 \\ v_1, \dots, v_\ell \in \Sigma_k \\ v_1 \dots v_\ell = \alpha \\ v_1 \dots v_i \neq \varepsilon_k\ (1 \leq i \leq \ell)}} p_{v_1} \dots p_{v_\ell}\, z^\ell,$$

since if $(v_1, \dots, v_\ell)$ is a path to α without loop around α, then $(v_\ell, \dots, v_1)$ is a path to α that does not return to ε_k. This provides a bijection that preserves the weights $p_{v_1} \dots p_{v_\ell}$ of the paths. The second sum can be easier to compute than the first one. We give below some concrete examples of computations, using the expression on the right.

- For $G = \mathbb{Z}/2\mathbb{Z} = \{\varepsilon, a\}$,
$$f(a,z) = F(z) = p_a z.$$

- For $G = \mathbb{Z}/3\mathbb{Z} = \{\varepsilon, a, a^2\}$,
$$f(a,z) = \frac{p_a z + p_{a^2}^2 z^2}{1 - p_a p_{a^2} z^2} \ ; \ f(a^2,z) = \frac{p_{a^2} z + p_a^2 z^2}{1 - p_a p_{a^2} z^2},$$
so that:
$$F(z) = \frac{(p_a + p_{a^2})z + (p_a^2 + p_{a^2}^2)z^2}{1 - p_a p_{a^2} z^2}.$$

- For $G = \mathbb{Z}/n\mathbb{Z}$, in order to find $f(a^k, z)$ for $k \in \{1, \dots, n\}$, one just has to solve the linear system of equations given by: $f(a^k, z) = zp_{a^k} + zp_{a^{k-1}} f(a,z) + zp_{a^{k-2}} f(a^2, z) + \dots + zp_{a^{k-(n-1)}} f(a^{n-1}, z)$.

- For $G = \mathbb{Z}$, with $p_a = p_{a^{-1}} = p$ and $p_{a^k} = 0$ for $k \notin \{-1, 1\}$, a similar infinite system of linear equations provides: $f(a^n, z) = r(z)^n$ where $r(z) = \frac{1-\sqrt{1-4p^2z^2}}{2pz}$, and $F(z) = \frac{2r(z)}{1-r(z)}$.

- For the monoid $B_n =< b | b^{n+1} = b^n >$, with $p_b = p$ and $p_{b^i} = 0$ for $i > 1$, we have for $k \in \{1, \dots n\}$,
$$f(b^k, z) = p^k z^k,$$
and
$$F(z) = \frac{pz}{1 - pz}(1 - p^{n+1} z^{n+1}).$$

7.2.4 Expression of the drift

The drift γ can be expressed as the expected change of length of an infinite normal form distributed according to the harmonic measure μ^∞, when left-multiplying by an element distributed according to p [Led01, Mai05].

Let us consider an infinite word in normal form. It begins by a letter of Σ_k with probability $x(\Sigma_k)$. Let us denote by $\alpha \in \Sigma_k$ the first letter. Then the increment of the length will be equal to $+1$ if we left-multiply by an element of $\Sigma \setminus \Sigma_k$ (probability $1 - p_k$), equal to -1 if we left-multiply by α^{-1}, and equal to 0 otherwise (multiplication by an element of $\Sigma_k \setminus \{\alpha\}$). We thus obtain:
$$\gamma = \sum_{k=1}^{n} x(\Sigma_k)(1 - p_k) - \sum_{k=1}^{n} \sum_{\alpha \in \Sigma_k} x(\alpha) p_{\alpha^{-1}}.$$
Moreover, with the notations of Sec. 7.2.2, we have: $\sum_{\alpha \in \Sigma_k} x(\alpha) p_{\alpha^{-1}} = A_k$. Since both $x(\Sigma_k)$ and A_k can be written as rational functions of the B_k, a consequence is that the drift can also be written as a rational function of the B_k.

7.3 The group $\mathbb{Z}^2 * \mathbb{Z}$: a case study

As a case study, let us consider the random walk on $\mathbb{Z}^2 * \mathbb{Z}$. We set $G_1 = \mathbb{Z}^2 = \langle a, b | ab = ba \rangle$ and $G_2 = \mathbb{Z} = \langle c | - \rangle$. With these notations, we have $\Sigma_1 = \{a^i b^j; (i,j) \in \mathbb{Z}^2 \setminus \{(0,0)\}\}$, $\Sigma_2 = \{c^k; k \in \mathbb{Z} \setminus \{0\}\}$ and $\Sigma = \Sigma_1 \cup \Sigma_2$, see Ex. 7.2. We assume that the distribution describing the steps of the random walk is concentrated on the three letters a, b, c and their inverses, with: $p_a = p_{a^{-1}} = p_b = p_{b^{-1}} = p$, and $p_c = p_{c^{-1}} = q$, so that $2p + q = 1/2$.

7.3.1 Equations for the harmonic measure

Let us describe the generating functions associated to $G_1 = \mathbb{Z}^2$ and $G_2 = \mathbb{Z}$ for this choice of weights.

For an element $a^i b^j$ of Σ_1, we have:

$$f_1(a^i b^j, z) = \sum_{\ell \geq 0} M_\ell(i,j)\, p^\ell z^\ell,$$

where $M_\ell(i,j)$ is the number of paths of length ℓ on the grid $\mathbb{Z}^2$ that begin at $(0,0)$ and arrive in (i,j), without having come back to $(0,0)$. The numbers $M_\ell(i,j)$ satisfy the following recursive formula:

$$M_{\ell+1}(i,j) = M_\ell(i-1,j) + M_\ell(i+1,j) + M_\ell(i,j-1) + M_\ell(i,j+1), \tag{7.4}$$

where we set $M_k(0,0) = 0$ for any $k \geq 0$. We have:

$$\begin{aligned} F_1(z) &= \sum_{(i,j)\in\mathbb{Z}^2\setminus\{(0,0)\}} f_1(a^i b^j, z) \\ &= \sum_{(i,j)\in\mathbb{Z}^2\setminus\{(0,0)\}} \sum_{\ell\geq 0} M_\ell(i,j)\, p^\ell z^\ell \\ &= \sum_{\ell \geq 0} M_\ell\, p^\ell z^\ell \end{aligned}$$

where $M_\ell = \sum_{(i,j)\in\mathbb{Z}^2\setminus\{(0,0)\}} M_\ell(i,j)$ is the number of paths of length ℓ that begin at $(0,0)$ and never return to $(0,0)$. The first terms of $F_1(z)$ are given by:

$$F_1(z) = 4pz + 12p^2z^2 + 48p^3z^3 + 172p^4z^4 + 688p^5z^5 + 2576p^6z^6 + 10304p^7z^7 + 39340p^8z^8 + \dots$$

For $G_2 = \mathbb{Z}$, we also have the interpretation:

$$f_2(c^k, z) = \sum_{\ell\geq 0} N_\ell(k)\, q^\ell z^\ell,$$

where for $k \in \mathbb{Z}\setminus\{0\}$, $N_\ell(k)$ is the number of paths of length ℓ on $\mathbb{Z}$ that begin at 0 and arrive in k, without having come back to 0. As mentioned in the previous section, the recurence relation on the coefficients provides a close expression:

$$f_2(c^k, z) = r(z)^k \text{ and } F_2(z) = \frac{2r(z)}{1-r(z)}, \text{ where } r(z) = \frac{1-\sqrt{1-4q^2z^2}}{2qz}.$$

The numbers $N_\ell = \sum_{k\in\mathbb{Z}\setminus\{0\}} N_\ell(k)$ are given by the central binomial coefficients. Precisely, we have: $N_\ell = 2\binom{\ell}{\lfloor \ell/2 \rfloor}$, and the beginning of the development of $F_2(z)$ is given by:

$$F_2(z) = 2pz + 2p^2z^2 + 4p^3z^3 + 6p^4z^4 + 12p^5z^5 + 20p^6z^6 + 40p^7z^7 + 70p^8z^8 + 140p^9z^9 + \dots$$

We are interested in finding B_1 and B_2 satisfying the system:

$$\begin{cases} \frac{F_1(B_1)}{1+F_1(B_1)} = \frac{4p+\frac{1}{B_2}-1}{\frac{1}{B_1}+\frac{1}{B_2}-1} \\ \frac{F_2(B_2)}{1+F_2(B_2)} = \frac{2q+\frac{1}{B_1}-1}{\frac{1}{B_1}+\frac{1}{B_2}-1}. \end{cases}$$

The second equation can be rewritten:

$$\frac{2r(B_2)}{1+r(B_2)} = \frac{2q+\frac{1}{B_1}-1}{\frac{1}{B_1}+\frac{1}{B_2}-1},$$

providing an expression of B_1 as a function of B_2. By inserting it in the first equation, we can find a numerical solution of the system.

By Prop. 7.2, the harmonic measure is given by:

$$x(a^i b^j) = \frac{1}{1+F_1(B_1)} f_1(a^i b^j, B_1), \qquad x(c^k) = \frac{1}{1+F_2(B_2)} f_2(c^k, B_2).$$

For $p = q = 1/6$, the numerical values we have obtained are $B_1 = 1.0727$ and $B_2 = 1.1719$, providing $x(\Sigma_1) = 0.662$ and $x(\Sigma_2) = 0.338$.

7.3.2 Different notions of drift

Let us recall the definition of the length of a word given in Sec. 7.1.1, that is:

$$|g| = \min\{k \in \mathbb{N}\,; g = u_1 * \ldots * u_k, u_i \in \Sigma\}.$$

For this notion of length, we have for example: $|(a^2b^{-3})c^{12}(ab)c^{-2}| = 4$.

By specialising the formula of Sec. 7.2.4 and using $x(\Sigma_1) + x(\Sigma_2) = 1$ and $1 - 2q = 4p$, we obtain:

$$\begin{aligned} \gamma &= (1-4p)\,x(\Sigma_1) + (1-2q)\,x(\Sigma_2) - 4p\,x(a) - 2q\,x(c) \\ &= (2q-4p)\,x(\Sigma_1) - 4p\,x(a) - 2q\,x(c) + 4p. \end{aligned}$$

Let us also define for an element $a^i b^j$ of Σ_1:

$$|a^i b^j|_1 = |i| + |j| \text{ and } |a^i b^j|_\infty = \max\{|i|, |j|\},$$

as well as for an element c^k of Σ_2: $|c^k|_1 = |c^k|_\infty = |k|$.

We can then define two different lengths by:

$$|g|_1 = \min\{|u_1|_1 + \ldots + |u_k|_1 \in \mathbb{N}\,; g = u_1 * \ldots * u_k, u_i \in \Sigma\}$$

$$|g|_\infty = \min\{|u_1|_\infty + \ldots + |u_k|_\infty \in \mathbb{N}\,; g = u_1 * \ldots * u_k, u_i \in \Sigma\}.$$

For these notions of lengths, we have respectively

$$|(a^2b^{-3})c^{12}(ab)c^{-2}|_1 = 21, \qquad \text{and} \qquad |(a^2b^{-3})c^{12}(ab)c^{-2}|_\infty = 18.$$

In terms of heaps (see Fig. 7.1), $|g|_1$ corresponds to the smallest number of pieces of a heap representing g, while $|g|_\infty$ is the smaller height of a heap representing g.

We denote by γ_1 and γ_∞ the drifts corresponding to these notions of length. To find an expression for γ_1 and γ_∞, let us study in both cases the increment of the length of an infinite normal form distributed according to the harmonic measure μ^∞, when left-multiplying by an element. We first consider γ_1, and then γ_∞.

Drift γ_1 (growing speed of the number of pieces in the heap). Let us look at the first letter of an infinite normal form word. This letter can be any letter of Σ.

1. If it is of the form a^i (resp. b^j), then the length with respect to $|\cdot|_1$ increases by 1 unless we left-multiply by a^{-1} if $i > 0$ or a if $i < 0$ (resp. b^{-1} if $j > 0$ or b if $j < 0$), in which case it decreases by 1. So, the increment is $+1$ with probability $1 - p = 3p + 2q$, and -1 with probability p, and the expected value of the increment is $1 - 2p = 2p + 2q$.

2. If it is another letter $a^i b^j$ of Σ_1, then the increment is $+1$ with probability $2p+2q$ and -1 with probability $2p$, so that the expected value of the increment is $2q$.

3. If it is of the form c^k, then the increment is $+1$ with probability $4p+q$ and -1 with probability q, and the expected value of the increment is $4p$.

By symmetry, $\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i) = \sum_{i\in\mathbb{Z}\setminus\{0\}} x(b^i)$, so that the probability of event 1 is $2\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i)$. We have thus:

$$\begin{aligned}
\gamma_1 &= 2(2p+2q)\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i) + 2q\Big(x(\Sigma_1) - 2\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i)\Big) + 4p\,x(\Sigma_2)\\
&= 4p\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i) + 2q\,x(\Sigma_1) + 4p\,x(\Sigma_2)\\
&= (2q-4p)\,x(\Sigma_1) + 4p\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i) + 4p.
\end{aligned}$$

The value $\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i)$ can be expressed more explicitly. Let us set

$$h_0(z) = 1 + \sum_{j\in\mathbb{Z}\setminus\{0\}} f_1(b^j, z),$$

and for $i \geq 1$,

$$h_i(z) = \sum_{j\in\mathbb{Z}} f_1(a^i b^j, z).$$

Using (7.4), we get, for $i \geq 1$,

$$h_i(z) = pz(h_{i+1}(z) + 2h_i(z) + h_{i-1}(z)).$$

It follows that: $h_i(z) = h_0(z)s(z)^i$, with:

$$s(z) = \frac{1 - 2pz - \sqrt{1-4pz}}{2pz},$$

and since $\sum_{i\in\mathbb{Z}} h_i(z) = F_1(z) + 1$, we obtain:

$$h_0(z) = (1 + F_1(z))\frac{1-s(z)}{1+s(z)},$$

and

$$\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^i) = \frac{1}{1+F_1(B_1)}(h_0(B_1) - 1) = \frac{1-s(B_1)}{1+s(B_1)} - \frac{1}{1+F_1(B_1)}.$$

Finally, since $x(\Sigma_1) = \frac{F_1(B_1)}{1+F_1(B_1)}$, we obtain:

$$\begin{aligned}
\gamma_1 &= (2q-4p)\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\Big(\frac{1-s(B_1)}{1+s(B_1)} - \frac{1}{1+F_1(B_1)}\Big) + 4p\\
&= 2q\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\frac{1-s(B_1)}{1+s(B_1)}\\
&= 2q\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\frac{-1+4pB_1+\sqrt{1-4pB_1}}{1-\sqrt{1-4pB_1}}.
\end{aligned}$$

Drift γ_∞ (growing speed of the height of the heap). Once again, to describe the increment of the length when left-multiplying by an element of Σ, we look at the first letter of the normal form word.

1. If it is of the form a^ib^i (resp. a^ib^{-i}) for some $i \in \mathbb{Z}^0$, then the increment (with respect to $|\cdot|_\infty$) is $+1$ with probability $2p+2q$ and 0 with probability $2p$, which gives a mean of $2p+2q$.

2. If it is another letter a^ib^j of Σ_1, then the increment is $+1$ with probability $p+2q$, -1 with probability p, and 0 with probability $2p$, so that the expected value of the increment is $2q$.

3. If it is of the form c^k, then the increment is $+1$ with probability $4p+q$ and -1 with probability q, and the expected value of the increment is $4p$.

By symmetry, the probability of event 1 is equal to $2\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i)$. We thus obtain:

$$\begin{aligned}
\gamma_\infty &= 2(2p+2q)\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i) + 2q\left(x(\Sigma_1) - 2\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i)\right) + 4p\,x(\Sigma_2)\\
&= 4p\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i) + 2q\,x(\Sigma_1) + 4p\,x(\Sigma_2)\\
&= (2q-4p)\,x(\Sigma_1) + 4p\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i) + 4p.
\end{aligned}$$

The value $\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i)$ can also be expressed more explicitly: if we introduce

$$\tilde{h}_0(z) = 1 + \sum_{j\in\mathbb{Z}\setminus\{0\}} f_1(a^jb^j, z)$$

and for $i \geq 1$,

$$\tilde{h}_i(z) = \sum_{j\in\mathbb{Z}} f_1(a^{i+j}b^i, z),$$

we indeed obtain for $i \geq 1$:

$$\tilde{h}_i(z) = 2pz(h_{i+1}(z) + h_{i-1}(z)),$$

which allows to find an expression for $\tilde{h}_0$. Precisely, we have $\tilde{h}_i(z) = \tilde{h}_0(z)\tilde{s}(z)^i$, with:

$$\tilde{s}(z) = \frac{1-\sqrt{1-4p^2z^2}}{2pz},$$

and since $\sum_{i\in\mathbb{Z}} h_i(z) = F_1(z)+1$, we obtain:

$$\tilde{h}_0(z) = (1+F_1(z))\frac{1-\tilde{s}(z)}{1+\tilde{s}(z)},$$

and

$$\sum_{i\in\mathbb{Z}\setminus\{0\}} x(a^ib^i) = \frac{1}{1+F_1(B_1)}(\tilde{h}_0(B_1)-1) = \frac{1-\tilde{s}(B_1)}{1+\tilde{s}(B_1)} - \frac{1}{1+F_1(B_1)}.$$

Finally, since $x(\Sigma_1) = \frac{F_1(B_1)}{1+F_1(B_1)}$, we obtain:

$$
\begin{aligned}
\gamma_\infty &= (2q-4p)\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\Big(\frac{1-\tilde{s}(B_1)}{1+\tilde{s}(B_1)} - \frac{1}{1+F_1(B_1)}\Big) + 4p \\
&= 2q\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\frac{1-\tilde{s}(B_1)}{1+\tilde{s}(B_1)} \\
&= 2q\frac{F_1(B_1)}{1+F_1(B_1)} + 4p\frac{-1+2pB_1+\sqrt{1-4p^2B_1^2}}{1+2pB_1-\sqrt{1-4p^2B_1^2}}.
\end{aligned}
$$

Numerical results. For $p = q = 1/6$, using the approximation $B_1 = 1.0727$, we obtain respectively $\gamma_1 = 0.576$ and $\gamma_\infty = 0.492$, these two values being consistent with the experimental results obtained when simulating the random walk. For the drift γ, we obtain experimentally $\gamma = 0.351$.

Chapter 8

Measures of maximal entropy of subshifts of finite type

Il leur avait semblé à tous les trois que c'était une bonne idée d'acheter ce cheval. Même si ça ne devait servir qu'à payer les cigarettes de Joseph. D'abord, c'était une idée, ça prouvait qu'ils pouvaient encore avoir des idées.

– Marguerite Duras, *Un barrage contre le Pacifique*

Contents

The study of measures of maximal entropy of SFT is motivated by the wish to be able to generate configurations as uniformly as possible, and to understand what do "typical" configurations look like. On $\mathbb{Z}$, it is well-known that a given SFT has a unique measure of maximal entropy, which is a Markov measure, known as the Parry measure of the SFT (these measures have been introduced by Shannon [Sha48], but Parry has proved the uniqueness [Par64]). We present alternative descriptions of Parry measures, allowing to sample them using i.i.d. random variables. On $\mathbb{Z}^2$, there can be in general several measures of maximal entropy, and little is known about these measures. We also present a contribution to the understanding of measures of maximal entropy for SFT defined on regular trees, providing a practical setting to the theory of the f-invariant developed by Bowen [Bow10]. Finally, we highlight a close connection between measures of maximal entropy of SFT and PCA.

8.1 SFT on $\mathbb{Z}$: the Parry measure

8.1.1 Definition and characterisation of the Parry measure

Let us consider the nearest-neighbour *subshift of finite type* Σ_A over the alphabet $\mathcal{A} = \{1, \dots, n\}$, defined by the adjacency matrix $A \in \mathcal{M}_n(\{0,1\})$, that is:

$$\Sigma_A = \{x \in \mathcal{A}^{\mathbb{Z}} \,;\forall k \in \mathbb{Z}, A_{x_k,x_{k+1}} = 1\},$$

meaning that

$$A_{i,j} = \begin{cases} 1 \text{ if } ij \text{ is an admissible pattern,} \\ 0 \text{ if } ij \text{ is a forbidden pattern.} \end{cases}$$

In all the following, we assume that A is irreducible and aperiodic.

Remark. Through this chapter, we will only consider nearest-neighbour SFT, for which the set of forbidden patterns is a subset of $\mathcal{A}^2$. But all the results can be extended to general SFT. Indeed, if a SFT is described by forbidden patterns of length k, for $k \geq 2$, then one can interpret it as a nearest-neighbour SFT on the alphabet $\mathcal{B} = \mathcal{A}^{k-1}$, of transition matrix B defined for $u, v \in \mathcal{B}$ by:

$$B_{u,v} = \begin{cases} 1 \text{ if } u_2 \dots u_{k-1} = v_1 \dots v_{k-2} \text{ and } u_1 \dots u_{k-1} v_{k-1} \text{ is an admissible pattern,} \\ 0 \text{ otherwise.} \end{cases}$$

Let us denote by $\mathcal{S}(i)$ the set of successors of i, that is, $\mathcal{S}(i) = \{j \in \mathcal{A} \,; A_{i,j} = 1\}$.

We denote by $\mathcal{W}(A, k)$ the set of admissible words of Σ_A of length k.

Since A is irreducible and aperiodic, by Perron-Frobenius theorem, there exists a real eigenvalue $\lambda > 0$ such that: λ has strictly positive right and left eigenvectors, the eigenvectors for λ are unique up to a multiplicative constant, and $\lambda > |\mu|$, where μ is any other eigenvalue. Moreover, the only eigenvectors whose components are all positive are those associated with the eigenvalue λ.

This eigenvalue λ is called the *Perron eigenvalue* of A.

Definition 8.1 (Parry measure)**.** Let λ be the Perron value of the matrix A, and let r be the right-eigenvector associated to λ, satisfying $\sum_{i=1}^n r(i) = 1$. The *Parry measure* is the Markov measure of transition matrix P defined, for any $i, j \in \mathcal{A}$, by

$$P_{i,j} = A_{i,j} \frac{r(j)}{\lambda r(i)}.$$

The vector r can be interpreted as a probability on $\mathcal{A}$. By definition of λ and r, we have: $\sum_{k=1}^n A_{i,k} r(k) = \lambda r(i)$, so that $\lambda r(i) = \sum_{k \in \mathcal{S}(i)} r(k) = r(\mathcal{S}(i))$. The matrix P can thus be written:

$$P_{i,j} = \begin{cases} r(j)/r(\mathcal{S}(i)) \text{ if } j \in \mathcal{S}(i) \\ 0 \text{ otherwise,} \end{cases}$$

meaning that it is a Markov-multiplicative measure, see Sec. 7.2.1.

The stationary measure of the Markov chain of transition matrix P is given by $\pi(i) = \ell(i) r(i)$ where ℓ is the left-eigenvector associated to λ satisfying $\sum_{i=1}^n \ell(i) r(i) = 1$. We have indeed:

$$\sum_{i \in \mathcal{A}} \ell(i) r(i) P_{i,j} = \sum_{i \in \mathcal{A}} \ell(i) r(i) A_{i,j} \frac{r(j)}{\lambda r(i)} = \frac{r(j)}{\lambda} \sum_{i \in \mathcal{A}} \ell(i) A_{i,j} = \ell(j) r(j).$$

We still denote by π the Parry measure, that is, the shift-invariant measure that is induced by P and π on Σ_A, so that the probability of occurence of a word $a_1 \dots a_k$ is given by:

$$\pi(a_1 \dots a_k) = \pi(a_1) P_{a_1,a_2} \dots P_{a_{n-1},a_n}.$$

Let us notice that for any word $w \in \mathcal{A}^k$ such that $awb \in \mathcal{W}(A, k+2)$, we have

$$\pi(awb) = \pi(a)\frac{r(w_1)}{\lambda r(a)}\frac{r(w_2)}{\lambda r(w_1)}\cdots\frac{r(w_k)}{\lambda r(w_{k-1})}\frac{r(b)}{\lambda r(w_k)} = \frac{\pi(a)r(b)}{\lambda^{k+1}r(a)}.$$

With the terminology of Chap. 1, Sec. 1.2, we will thus say that π is *Markov-uniform*: for any $k \geq 0$ and any $a, b \in \mathcal{A}$, the measure $\pi(awb)$ does not depend of the word $w \in \{1, \ldots, n\}^k$ such that $awb \in \mathcal{W}(A, k+2)$. In general, the uniform measures on the set $\mathcal{W}(A, k)$ of allowable words of Σ_A of length k are not consistent for different values of k, so that it is not possible to extend them with Kolmogorov consistency theorem, to define a measure on the whole subshift. But in some sense, the Parry measure distributes probabilities on paths as uniformly as possible.

The following characterisation of the Parry measure is a folk result [Kit98], that has been generalised by Burton and Steif [BS94].

Theorem 8.1. *Let $\mathcal{M}_{\Sigma_A}$ be the set of translation invariant measures on the SFT Σ_A, and let $\pi \in \mathcal{M}_{\Sigma_A}$. The following properties are equivalent.*

(i) The measure π is the Parry measure associated to Σ_A.

(ii) The measure π is Markov-uniform.

(iii) The measure-theoretic entropy of π satisfies $h(\pi) = \sup_{\mu \in \mathcal{M}_{\Sigma_A}} h(\mu)$.

The Parry measure π is thus the unique measure that achieves the supremum of the entropy. Its entropy is given by:

$$\begin{aligned}
h(\pi) &= -\sum_{ij \in \mathcal{W}(A,2)} \pi_i P_{i,j} \log P_{i,j} \\
&= -\sum_{ij \in \mathcal{W}(A,2)} \ell(i) r(i) \frac{r(j)}{\lambda r(i)} \log \frac{r(j)}{\lambda r(i)} \\
&= -\frac{1}{\lambda} \sum_{i,j \in \mathcal{A}} \ell(i) r(j) A_{i,j} (\log r(j) - \log \lambda - \log r(i)) \\
&= -\sum_{j \in \mathcal{A}} \ell(j) r(j) (\log r(j) - \log \lambda) + \sum_{i \in \mathcal{A}} \ell(i) r(i) \log r(i) \\
&= \log(\lambda),
\end{aligned}$$

which is equal to the topological entropy of Σ_A, defined by:

$$h(\Sigma_A) = \lim_{k \to \infty} \frac{\log(\text{Card } \mathcal{W}(k, \mathcal{A}))}{k}.$$

The following example will be used as an illustration through this chapter.

Example 8.1 (Fibonacci subshift)**.** The *Fibonacci* or *golden mean* subshift is the subshift over the binary alphabet $\mathcal{A} = \{0, 1\}$, defined by the matrix: $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. The constraints can also be represented by the automaton of Fig 8.1.

The Perron value of A is the golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$, satisfying $\varphi^2 = \varphi + 1$. The corresponding right-eigenvector is given by $r(0) = \frac{1}{\varphi}$ and $r(1) = \frac{1}{\varphi^2}$, so that the Parry measure of the subshift is given by the transition matrix: $P = \begin{pmatrix} \frac{1}{\varphi} & \frac{1}{\varphi^2} \\ 1 & 0 \end{pmatrix}$, and we have $\pi(0) = \frac{\varphi^2}{1+\varphi^2}$, $\pi(1) = \frac{1}{1+\varphi^2}$. The entropy of the subshift is equal to $\log(\varphi)$.

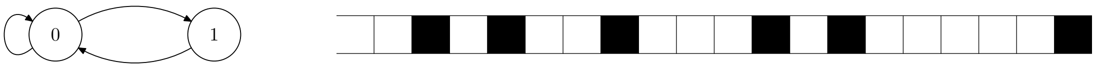


Figure 8.1: Fibonacci subshift: automaton representing the constraint, and example of configuration obtained.

In Chap. 7, we have seen how random walks on free products of groups give rise to Markov multiplicative measures on the SFT defined by the set of normal form words. The random walks on free products of groups that we have studied can be seen as the concatenation of i.i.d. elements of the different groups, with simplification when two consecutive elements belong to the same group. Conversely, the next two sections show how starting from a particular Markov multiplicative measure π corresponding to a Parry measure, we can define a probability measure $\mathcal{B}(p)$ on the alphabet $\mathcal{A}$, as well as simplification rules, such that from a sequence distributed according to $\mathcal{B}(p)^{\otimes\mathbb{Z}}$, after applying the possible simplifications, we recover a sequence distributed according to the measure π.

8.1.2 Realisations of the Parry measure with i.i.d random variables

We present here a very simple way to generate words of Σ_A distributed according to the Parry measure. The proposition follows from the fact that the Parry measure is the Markov-multiplicative measure associated to the probability r.

Proposition 8.1. *Let r be the right-eigenvector associated to the Perron value of A, and satisfying $\sum_{i=1}^n r(i) = 1$. A way to generate the Parry measure π of the SFT Σ_A consists in drawing the first letter according to π, and then choosing successively and independently letters of $\mathcal{A}$ according to the probability r, and rejecting letters that would provide a forbidden pattern.*

Precisely, the algorithm is the following.

Algorithm 5: Sampling the Parry measure with i.i.d. r.v.

Data: A sequence $(x_t)_{t\geq 1}$ of i.i.d. r.v. of probability r.
Result: A sequence $a_1 \dots a_n$ distributed according to the Parry measure.
begin
  $t = 0$; k=1 ;
  Choose the first letter a_1 according to the probability π;
  repeat
    If $A_{a_k,x_t} = 1$, then choose $a_{k+1} = x_t$ and set $k = k+1$ (else, do nothing) ;
    $t = t+1$.
  until $k{=}n$;
  return *The sequence* $a_1 \dots a_n$
end

If the last letter that has been chosen is the letter i, then the next letter will be j with probability:

$$A_{i,j}\frac{r(j)}{r(\mathcal{S}(i))}$$

so that we indeed recover the Parry measure.

Example 8.2. For the Fibonacci subshift, the algorithm consists in choosing the first letter to be a 0 with probability $\pi(0) = \frac{\varphi^2}{1+\varphi^2}$, and a one with probability $\pi(1) = \frac{1}{1+\varphi^2}$. And then for all the following letters, choosing to write a 0 with probability $r(0) = \frac{1}{\varphi}$ and a 1 with probability $r(1) = \frac{1}{\varphi^2}$, and rejecting the 1 that are not allowed. Here, the algorithm is not very efficient since once we have chosen a 1, there is no choice for the following letter: it is thus useless to wait until drawing a 0.

If there exists a particular symbol 0 such that for any letter $i \in \mathcal{A}$, $A_{0,i} = A_{i,0} = 1$, then the Parry measure can also be obtained by iterating a simple probabilistic cellular automaton.

Proposition 8.2. *Let Σ_A be an SFT on the alphabet $\mathcal{A} = \{0, 1, \ldots, n\}$ such that for any $i \in \mathcal{A}$, $A_{0,i} = A_{i,0} = 1$. We define $\bar{\mathcal{A}} = \{\bar{0}, \bar{1}, \ldots, \bar{n}\}$ and $\mathcal{B} = \mathcal{A} \cup \bar{\mathcal{A}}$. Let us consider the PCA F of alphabet $\mathcal{B}$ and neighbourhood $\mathcal{N} = \{-1, 0\}$ defined by the local function:*

$$f(i,j) = \begin{cases} \mathcal{B}(r) \text{ if } i \in \bar{\mathcal{A}}, j \in \mathcal{A} \text{ and } A_{i,j} = 0 \\ \delta_{\bar{j}} \text{ if } i \in \bar{\mathcal{A}}, j \in \mathcal{A} \text{ and } A_{i,j} = 1 \\ \delta_j \text{ otherwise.} \end{cases}$$

We choose a configuration according to the Bernoulli product measure $\mu_r = \mathcal{B}(r)^{\otimes \mathbb{Z}}$ on $\mathcal{A}^{\mathbb{Z}}$, and replace all the 0's by $\bar{0}$'s. From this initial configuration, the trajectories of the PCA converge to configurations of the SFT distributed according to the Parry measure (when forgetting the overlines of the letters).

Proof. We will define a coupling between the descriptions of the Parry measure of Prop. 8.1 and Prop. 8.2. Let us consider a sequence $((x_{s,t})_{s\in\mathbb{Z},t\in\mathbb{N}})$ of i.i.d. random variables of probability $\mathcal{B}(r)$. We construct a space-time diagram of the PCA F as follows: the initial configuration is $(x_{s,0})_{s\in\mathbb{Z}}$ (with 0's replaced by $\bar{0}$'s), and for each cell s, each time we need to draw a Bernoulli $\mathcal{B}(r)$ we take the first element of the sequence $(x_{s,t})_{t\geq 0}$ that has not been used.

Let us observe that once a cell is in a state of $\bar{\mathcal{A}}$ it always remains in the same state. Let $k \in \mathbb{Z}$ be such that $x_{k,0} = 0$, so that in the initial configuration, cell k is in state $\bar{0}$. Initially, cell $k+1$ is in a state $x_{k+1,0} = a_1$ distributed according to the Bernoulli $\mathcal{B}(r)$. Since $0\,a_1$ is an allowed pattern, a_1 becomes $\bar{a_1}$ and keeps this state forever. Cell $k+2$ is initially in a state $x_{k+2,0} = i$ distributed according to $\mathcal{B}(r)$. If $a_1\,i$ is an allowed pattern, then i becomes $\bar{i}$ and the cell remains in that state. If not, we look at $x_{k+2,1}$, and so on, until we read a value $x_{k+2,t} = a_2$ such that $a_1 a_2$ is an allowed pattern. Before fixing the value of cell $k+2$, the value of cell $k+3$ has not changed. The final value of cell $k+3$ will be the first of the $(x_{k+3,t})_{t\geq 0}$ that is allowed after a_2, and so on. The construction thus corresponds to a running of the Markov chain defining the Parry measure. By stationarity, the measure on $\mathcal{A}^{\mathbb{Z}}$ obtained is the Parry measure. □

When iterating the PCA, the configurations progressively stabilise on a fixed point that is distributed according to the Parry measure. This provides a parallelisation of the computation of the Parry measure. Another possible description of the process consists in allocating to each cell independently a 0 with probability $r(0)$ (and nothing otherwise). This divides the configuration into different sections separated by 0's. From each of the cells labeled 0, we run the Markov chain defining the Parry measure up to the last cell before the next 0. This can be made in parallel in the different sections of the configurations.

Example 8.3. For the Fibonacci SFT, one can also extend the neighbourhood instead of extending the alphabet. The initial state is chosen according to the Bernoulli product measure $\mu_r = \mathcal{B}(r)^{\otimes \mathbb{Z}}$, then we iterate the PCA of neighbourhood $\mathcal{N} = \{-2, -1, 0\}$ defined by:

$$f(i,j,k) = \begin{cases} \mathcal{B}(r) \text{ if } (i,j,k) = (0,1,1) \\ \delta_k \text{ otherwise} \end{cases}$$

or equivalently, the CA of local function

$$f(i,j,k) = \begin{cases} 0 \text{ if } (i,j,k) = (0,1,1) \\ k \text{ otherwise.} \end{cases}$$

In some sense, these PCA scan the configuration from left to right and correct patterns that are not allowed. It would be satisfying to obtain the Parry measure by iterating a PCA with a symmetric update rule. A candidate we could think of is the PCA defined by the neighbourhood $\mathcal{N} = \{-1, 0, 1\}$ and the local function:

$$f(i,j,k) = \begin{cases} \mathcal{B}(r) \text{ if } (i,j) = (1,1) \text{ or } (j,k) = (1,1), \\ \delta_j \text{ otherwise.} \end{cases}$$

But we point out that this PCA does not have the behaviour we wish, since from any Bernoulli product measure, the final values of two cells at distance greater than 2 are independent, which is not the case under the Parry measure.

8.1.3 The case of confluent SFT

Let us consider a nearest-neighbour SFT Σ_A on the alphabet $\mathcal{A} = \{1, \ldots, n\}$. We introduce the following terminology, that is specific to this thesis.

Definition 8.2 (Confluent SFT)**.** We say that the SFT Σ_A is *confluent* if the matrix A satisfies:

$$(A_{i,j} = 0 \text{ and } A_{j,k} = 0) \implies i = k.$$

A SFT is confluent if and only if for any sequence of letters, if we delete forbidden patterns occuring in the sequence until only admissible patterns remain, the word that is obtained does not depend on the order in which the forbidden patterns have been deleted.

Let Σ_A be a confluent SFT. Then for any $i \in \mathcal{A}$, there is at most one letter $j \in \mathcal{A}$ such that $A_{i,j} = A_{j,i} = 0$. Indeed, if $A_{i,j} = A_{j,i} = 0$ and $A_{i,k} = A_{k,i} = 0$, then we have in particular $A_{j,i} = A_{i,k} = 0$, so that by definition of confluence, $j = k$.

We partition the alphabet into two subsets:

$$S_1 = \{i \in \mathcal{A}\,;\ \exists \ell \in \mathcal{A}, A_{i,\ell} = A_{\ell,i} = 0\},$$

$$S_2 = \mathcal{A} \setminus S_1 = \{i \in \mathcal{A}\,;\ \forall \ell \in \mathcal{A}, A_{i,\ell} = 1 \text{ or } A_{\ell,i} = 1\}.$$

We set $s_1 = \text{Card } S_1$ and $s_2 = \text{Card } S_2$ (note that $s_1 + s_2 = n$).

Lemma 8.1. *For any $i \in S_1$ and $j \in S_2$, $A_{i,j} = A_{j,i} = 1$.*

Proof. Let $i \in S_1$. There exists ℓ such that $A_{i,\ell} = A_{\ell,i} = 0$. If $A_{i,j} = 0$, then $A_{\ell,i} = A_{i,j} = 0$, so that $j = \ell$. In particular, $j \notin S_2$. In the same way, if $A_{j,i} = 0$, then $A_{j,i} = A_{i,\ell} = 0$, meaning that $j = \ell \in S_1$. □

Example 8.4. Let us consider for example the SFT defined by $\mathcal{A} = \{1, \ldots, 10\}$, and

$$A = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

One can check that this matrix defines a confluent SFT. Fig. 8.2 represents the graph of forbidden transitions of Σ_A. We have: $S_1 = \{1, 2, 3\}$ and $S_2 = \{4, 5, 6, 7, 8, 9, 10\}$, so that $s_1 = 3$ and $s_2 = 7$.

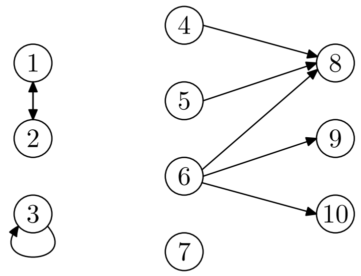


Figure 8.2: Graph of forbidden transitions for the SFT of Ex. 8.4.

Proposition 8.3. *Let λ be the Perron eigenvalue of the matrix A defining the confluent SFT Σ_A, and let $\sigma = \sum_{k \in S_2} \sum_{\ell=1}^{n} (1 - A_{k,\ell})$. We have:*

$$\lambda^3 + (1-n)\lambda^2 + (\sigma - s_2)\lambda + \sigma = 0, \tag{8.1}$$

and there exist $q_1, q_2 \geq 0$ satisfying $s_1 q_1 + s_2 q_2 = 1$ and

$$\begin{cases} 1 = \lambda q_1 - \frac{q_1}{\lambda} + \frac{q_1 s_1}{\lambda} + \frac{q_2 \sigma}{\lambda} \\ 1 = \lambda q_2 + \frac{q_1 s_1}{\lambda} + \frac{q_2 \sigma}{\lambda}. \end{cases}$$

Furthermore, a way to generate the Parry measure of Σ_A consists in drawing a bi-infinite i.i.d sequence of letters of $\mathcal{A}$ according to the distribution $\sum_{i \in S_1} q_1 \, \delta_i + \sum_{i \in S_2} q_2 \, \delta_i$, and erasing forbidden patterns.

We point out that the values of q_1 and q_2 are easy to find. One only has to identify the sets S_1 and S_2 by observing the matrix A, and to compute the value of σ, which corresponds to the number of edges among the vertices of S_2 in the graph of forbidden transitions (so that for Ex. 8.4, we have $\sigma = 5$). Then, the Perron value λ is a root of a polynomial of degree 3. And to determine the probabilities q_1 and q_2, one just has to solve a linear system of two equations and two unknowns.

Before proving the proposition, let us mention the result obtained for our favorite example.

Example 8.5. The Fibonacci subshift is a confluent SFT, with $S_1 = \{1\}$ and $S_2 = \{0\}$, so that $s_1 = s_2 = 1$. We have $\sigma = 0$, and equation (8.1) becomes $\lambda^3 - \lambda^2 - \lambda = 0$. We obtain $q_1 = \frac{1}{\varphi}$ and $q_2 = \frac{1}{\varphi^2}$, which means that if we draw independently 0's with probability $\frac{1}{\varphi^2}$ and 1's with probability $\frac{1}{\varphi}$ and then delete the pairs of consecutive 1's we see in the bi-infinite sequence obtained, we recover the Parry measure of the Fibonacci subshift. It is to compare with the result presented in Ex. 8.2.

Proof. Let p be a given probability on $\mathcal{A}$. We draw a sequence of $\mathcal{A}^{\mathbb{N}}$ according to the Bernoulli product measure $\mathcal{B}(p)^{\otimes \mathbb{N}}$ and delete forbidden patterns until reaching a word of the SFT. The understanding of the measure ν obtained on the admissible sequences of $\mathcal{A}^{\mathbb{N}}$ will then allow us to describe the measures obtained on bi-infinite sequences of Σ_A when starting from the Bernoulli product measure $\mathcal{B}(p)^{\otimes \mathbb{Z}}$.

Let P be the transition matrix of the Parry measure. We can look for parameters p_i for which ν would have the form:

$$\nu(a_1 \dots a_k) = \nu(a_1) P_{a_1,a_2} \dots P_{a_{k-1},a_k}, \tag{8.2}$$

for some distribution ν on $\mathcal{A}$. If it is the case, then by shift-stationarity, the measure obtained on $\mathcal{A}^{\mathbb{Z}}$ will be exactly the Parry measure. For readability, we will write p_i and r_i for $p(i)$

and $r(i)$ respectively, where r still denotes the normalised right-eigenvector of A associated to the Perron value, so that: $P_{i,j} = A_{i,j}\frac{r_j}{\lambda r_i}$.

By definition, the measure ν is invariant by left-multiplying by an element of $\mathcal{A}$ of probability p, and in the new word obtained, deleting the first two letters if they form a forbidden pattern. For one-dimensional marginals, the equation of this p-stationarity can be written:

$$\nu(i) = p_i \sum_{k=1}^{n} A_{i,k}\nu(k) + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell})\nu(\ell i).$$

For larger marginals, the traffic equation is:

$$\nu(i_1 i_2 \dots i_t) = p_{i_1} A_{i_1,i_2}\nu(i_2 \dots i_t) + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell})\nu(\ell i_1 i_2 \dots i_t).$$

If we require (8.2), it is sufficient to consider the equations for marginals of size one and two:

$$\nu(i) = p_i \sum_{k=1}^{n} A_{i,k}\nu(k) + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell})\nu(\ell)P_{\ell,i},$$

$$\nu(i)P_{i,j} = p_i A_{i,j}\nu(j) + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell})\nu(\ell)P_{\ell,i}P_{i,j}.$$

Finally, by multiplying the first line by $P_{i,j}$ and substracting it to the second line, we obtain the following system:

$$\left\{\begin{array}{rcl} \nu(i) & = & p_i \sum_{k=1}^{n} A_{i,k}\nu(k) + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell})\nu(\ell)P_{\ell,i} \\ A_{i,j}\nu(j) & = & P_{i,j}\sum_{k=1}^{n} A_{i,k}\nu(k). \end{array}\right.$$

The second equation is satisfied for $\nu = r$. The first equation then becomes:

$$r_i = p_i \sum_{k=1}^{n} A_{i,k} r_k + \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_k(1 - A_{k,\ell}) r_\ell P_{\ell,i}.$$

Since $\sum_{k=1}^{n} A_{i,k} r_k = \lambda r_i$ and $P_{\ell,i} = A_{\ell,i}\frac{r_i}{\lambda r_\ell}$, we can simplify the above expression by r_i and obtain:

$$1 = \lambda p_i + \frac{1}{\lambda}\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1 - A_{k,\ell})A_{\ell,i}p_k. \tag{8.3}$$

We will prove that there exist values $p_1, \dots, p_n \geq 0$ with $\sum_{i=1}^{n} p_i = 1$, satisfying the above equations (8.3).

Let us recall the partition of the alphabet into two subsets: $S_1 = \{i; \exists \ell, A_{i,\ell} = A_{\ell,i} = 0\}$, and $S_2 = \mathcal{A} \setminus S_1 = \{i; \forall \ell, A_{i,\ell} = 1 \text{ or } A_{\ell,i} = 1\}$, with $s_1 = \text{Card } S_1$ and $s_2 = \text{Card } S_2$.

Since the subshift is assumed to be confluent, for $i \in S_1$, there exists a unique letter ℓ (possibly the letter i itself) such that $A_{i,\ell} = A_{\ell,i} = 0$ (and then, we also have $\ell \in \Sigma_1$).

For a letter $i \in S_2$, three cases are possible:

- $\forall \ell \in \mathcal{A},\ A_{i,\ell} = 1$ and $A_{\ell,i} = 1$,
- $\exists \ell \in \mathcal{A},\ A_{i,\ell} = 0$ and $\forall k \in \mathcal{A},\ A_{k,i} = 1$,
- $\exists \ell \in \mathcal{A},\ A_{\ell,i} = 0$ and $\forall k \in \mathcal{A},\ A_{i,k} = 1$.

Let us now separate the terms $k = i$ and $k \neq i$ in the sum appearing in equation (8.3). We have:

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})A_{\ell,i}p_k = \sum_{\ell=1}^{n}(1-A_{i,\ell})A_{\ell,i}p_i + \sum_{k\neq i}\sum_{\ell=1}^{n}(1-A_{k,\ell})A_{\ell,i}p_k.$$

By confluence, if $k \neq i$, $(1-A_{k,\ell})(1-A_{\ell,i}) = 0$, so that: $(1-A_{k,\ell})A_{\ell,i} = 1 - A_{k,\ell}$. Thus,

$$\sum_{k\neq i}\sum_{\ell=1}^{n}(1-A_{k,\ell})A_{\ell,i}p_k = \sum_{k\neq i}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k = \sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k - \sum_{\ell=1}^{n}(1-A_{i,\ell})p_i,$$

and:

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})A_{\ell,i}p_k = \sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k - \sum_{\ell=1}^{n}(1-A_{i,\ell})(1-A_{\ell,i})p_i.$$

For $i \in S_1$, $\sum_{\ell=1}^{n}(1-A_{i,\ell})(1-A_{\ell,i}) = 1$, whereas for $i \in S_2$, $\sum_{\ell=1}^{n}(1-A_{i,\ell})(1-A_{\ell,i}) = 0$. So, for $i \in S_1$, equation (8.3) becomes:

$$1 = \lambda p_i + \frac{1}{\lambda}\Big(\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k - p_i\Big),$$

and for $i \in S_2$, it becomes:

$$1 = \lambda p_i + \frac{1}{\lambda}\Big(\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k\Big).$$

The quantity $\sum_{k=1}^{n}\sum_{\ell=1}^{n}(1-A_{k,\ell})p_k$ does not depend on i. Thus, if the equation has a solution, the p_i should have a common value q_1 for any $i \in S$, and a common value q_2 for any $i \in S$, with:

$$1 = \lambda q_1 - \frac{q_1}{\lambda} + \frac{q_1 s_1}{\lambda} + \frac{q_2\sigma}{\lambda},$$
$$1 = \lambda q_2 + \frac{q_1 s_1}{\lambda} + \frac{q_2\sigma}{\lambda},$$

where $\sigma = \sum_{k\in S_2}\sum_{\ell=1}^{n}(1-A_{k,\ell})$.

Our problem is thus equivalent to finding $q_1, q_2 \geq 0$ with $s_1q_1 + s_2q_2 = 1$ satisfying:

$$\begin{cases} 1 = \lambda q_1 - \frac{q_1}{\lambda} + \frac{q_1 s_1}{\lambda} + \frac{q_2\sigma}{\lambda}, \\ 1 = \lambda q_1 + \frac{q_1 s_1}{\lambda} + \frac{q_2\sigma}{\lambda}. \end{cases} \tag{8.4}$$

One can check that it is possible if and only if λ satisfies:

$$\lambda^4 - n\lambda^3 + (s_1 + \sigma - 1)\lambda^2 + s_2\lambda - \sigma = 0.$$

Since 1 is a root of that equation, we can simplify into:

$$\lambda^3 + (1-n)\lambda^2 + (\sigma - s_2)\lambda + \sigma = 0.$$

Let us denote by u_n the number of words of length n of our SFT ending with a letter of S_1, and by v_n the number of words of length n of the SFT ending with a letter of S_2. We claim that:

$$u_{n+1} = (s_1 - 1)u_n + s_1 v_n, \tag{8.5}$$

and

$$v_{n+1} = s_2(u_n + v_n) - \sigma(u_{n-1} + v_{n-1}). \tag{8.6}$$

Let us first explain relation (8.5). An admissible word of length $n+1$ ending by a letter of S_1 can be obtained either by taking an admissible word of length n ending by some letter i of S_1 (u_n choices) and adding at the end of that word any letter of S_1 different from the only letter ℓ such that $A_{i,\ell} = 0$ ($s_1 - 1$ choices), or by extending a word of length n ending with a letter of S_2 (v_n choices) by any letter of S_1 (s_1 choices). It is always possible since for any $i \in S_2$ and any $j \in S_1$, we have $A_{i,j} = 1$ (Lemma 8.1).

In order to explain equation (8.6), let us also introduce x_n, the number of words of length n ending by a given (fixed) letter i of S_2 such that $\exists \ell \in \mathcal{A},\ A_{i,\ell} = 0$ (we will see that the value of x_n does not depend of the choice of such an $i \in S_2$). If $i \in S_2$ satisfies $\exists \ell \in \mathcal{A},\ A_{i,\ell} = 0$, then: $\forall k \in \mathcal{A},\ A_{k,i} = 1$. It follows that $x_n = u_{n-1} + v_{n-1}$, since any word of length $n-1$ can be (uniquely) extended into a word of length n ending by i. Now, to obtain a word of length $n+1$ ending in any point of S_2, if we extend any admissible word of length n by a letter of S_2 (providing $s_2(u_n + v_n)$ words), we have counted exactly σx_n non-admissible words: the words ending by $i\ell$ for i as above, and ℓ such that $A_{i\ell} = 0$. The expression follows.

Equations (8.5) and (8.6) can be rewritten:

$$\begin{pmatrix} u_n \\ v_n \\ u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & s_1 - 1 & s_1 \\ -\sigma & -\sigma & s_2 & s_2 \end{pmatrix} \begin{pmatrix} u_{n-1} \\ v_{n-1} \\ u_n \\ v_n \end{pmatrix}.$$

The characteristic polynomial of the above matrix that is involved is given by:

$$X\ (X^3 + (1 - s_1 - s_2)X^2 + (\sigma - s_2)X + \sigma).$$

We have: $\lim_{n\to\infty} \frac{\log(u_n+v_n)}{n} = \log \mu$, where μ is a positive root of this polynomial. But we also know that the topological entropy of Σ_A is equal to the Perron eigenvalue λ of A. It follows that λ satisfies $\lambda^3 + (1 - s_1 - s_2)\lambda^2 + (\sigma - s_2)\lambda + \sigma = 0$. Thus, system (8.4) has a solution, from which we can define p satisfying (8.3). This concludes the proof. $\square$

8.2 SFT on $\mathbb{Z}^d$: characterisation of the measures of maximal entropy

In dimension $d \geq 2$, there are examples of SFT having several measures of maximal entropy, and these measures have no simple expression in general [BS94]. But we still have an analogue of Th. 8.1.

Let Σ be a nearest-neighbour SFT on $\mathbb{Z}^d$. We recall that a measure $\mu \in \mathcal{M}_\Sigma$ is *Markov-uniform* if it defines a Markov random field, and furthermore, the conditional distribution of μ on any finite set F given the configuration on its boundary ∂F is μ-a.s. uniform over all configurations on F which extend the configuration on ∂F.

Theorem 8.2 ([BS95]). *Let Σ be a (nearest-neighbour) SFT on $\mathbb{Z}^d$. There exists a measure $\mu \in \mathcal{M}_\Sigma$ such that $h(\mu) = \sup_{\nu\in\mathcal{M}_\Sigma} h(\nu)$. For $\mu \in \mathcal{M}_\Sigma$, the two following properties are equivalent.*

(i) The measure μ is Markov-uniform.

(ii) The measure-theoretic entropy of μ satisfies $h(\mu) = \sup_{\nu\in\mathcal{M}_\Sigma} h(\nu)$.

Furthermore, the topological entropy of Σ satisfies:

$$h(\Sigma) = \sup_{\nu\in\mathcal{M}_\Sigma} h(\nu).$$

Example 8.6. For the Fibonacci SFT on $\mathbb{Z}^2$, known also as the hard core (or hard square) model, the forbidden patterns are two consecutive ones, horizontally or vertically. We present an example of configuration in Fig. 8.3. It is known that this SFT has a unique measure of maximal entropy [vdBS94]. Nevertheless, this measure has no simple effective description and no close form is known for the value of the topological entropy. Approximating the entropy has lead to many research works, and is still an active research area [MP13].

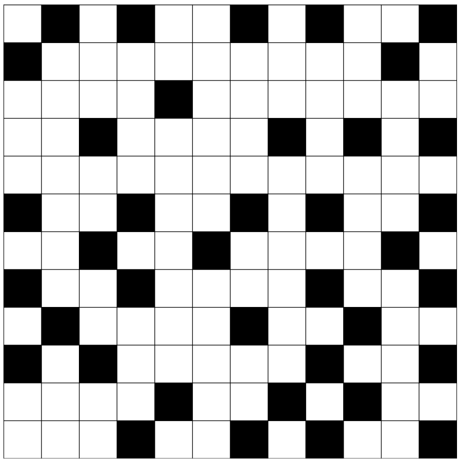


Figure 8.3: Example of configuration for the two-dimensional Fibonacci subshift.

In Sec. 8.4, we will go further into the interpretation of measures of maximal entropy in terms of Gibbs measures, and present a connection to PCA.

8.3 SFT on regular trees: generalising the Parry measure

We will now consider nearest-neighbour SFT on infinite regular trees, defined by some adjacency matrix A. Let us denote by F_d the free monoid with d generators. It can be represented as an infinite rooted tree of degree $d+1$ (with a root of degree d). The empty word, corresponding to the root, is denoted by ε. The SFT Σ_A^d is the set of labelings of the nodes of this tree such that if a node is labeled by the letter i, its children are labeled by letters that are in the set $\mathcal{S}(i) = \{j \in \mathcal{A};\ A_{i,j} = 1\}$. Formally, the alphabet is still $\mathcal{A} = \{1, \dots, n\}$, and if the generators of F_d are denoted by $a_1, \dots, a_d$, then:

$$\Sigma_A^d = \{x \in \mathcal{A}^{F_d};\ \forall w \in F_d, \forall i \in \{1 \dots, d\}, A_{x_w, x_{wa_i}} = 1\}.$$

In Sec. 8.3.3, we will assume that the matrix A is symmetric. In that case, the orientation of the tree can be forgotten, and instead of working on F_d, it also makes sense to consider the finitely presented group $T_{d+1} = \langle g_1, \dots, g_d, g_{d+1} \mid g_i^2 = 1 \rangle$.

SFT on trees have already been studied from a theoretical computer science point of view [AB12]. Here, the questions we address are the following. How to construct Markov-uniform measures for such SFT? Do the Markov-uniform measures maximise the entropy, for some "good" notion of entropy?

8.3.1 Markov chains on regular trees and the f-invariant

Let Q be a stochastic matrix with state space $\mathcal{A}$, and let π be a probability measure on $\mathcal{A}$.

A natural way to label the elements of F_d by letters of $\mathcal{A}$ (while taking in consideration the constraints given by the matrix A) is to do it in a Markovian way: we first choose the label of the root according to the distribution π, and then, if a node w is labeled by the letter α_0, its children wa_i, for $i \in \{1,\dots,d\}$, are labeled independently by a β with probability $P_{\alpha,\beta}$. This leads us to define the notion of Markov chains on trees.

For an element $w \in F_d$, let us denote by $\mathcal{C}(w)$ the set of children of w. Precisely, $\mathcal{C}(w) = \{wh; h \in F_d \setminus \{\varepsilon\}\}$.

Definition 8.3 (Markov chains on trees)**.** The *Markov chain* over F_d of transition matrix Q and initial distribution π is the set of random variables $(X_w)_{w\in F_d}$ such that the distribution of X_ε equals π, and for any $w \in F_d$ and any generator a_i,

$$\forall k \geq 1, \forall v_1,\dots,v_k \in F_d \setminus \mathcal{C}(w), \forall \alpha,\beta,\alpha_1,\dots,\alpha_k \in \mathcal{A},$$

$$\mathbb{P}(X_{wa_i} = \beta | X_w = \alpha, X_{v_1} = \alpha_1, \dots, X_{v_k} = \alpha_k) = \mathbb{P}(X_{wa_i} = \beta | X_w = \alpha) = Q_{\alpha,\beta}.$$

If π is an invariant measure of Q, the Markov chain is said to be invariant.

A Markov chain on F_d induces a *Markov measure* on $\mathcal{A}^{F_d}$. We now present the f-invariant of Bowen, that has been introduced in order to generalise the theory of entropy to free group actions [Bow10].

Definition 8.4 (f-invariant)**.** For any stochastic matrix Q of invariant measure π inducing a Markov measure μ on $\mathcal{A}^{F_d}$, we define the f-invariant of μ by:

$$f(\mu) = d\sum_{i=1}^{n} \pi(i)\log(\pi(i)) - \frac{d+1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \pi(i)Q_{i,j}\log \pi(i)Q_{i,j}, \tag{8.7}$$

or equivalently by

$$f(\mu) = \frac{d-1}{2}\sum_{i=1}^{n} \pi(i)\log(\pi(i)) - \frac{d+1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \pi(i)Q_{i,j}\log Q_{i,j}.$$

For $d = 1$, we recover the usual definition of the entropy of a Markov measure on $\mathbb{Z}$.

Like the entropy, the f-invariant can be defined for other measures on F_d than Markov measures, and it is a measure-conjugacy invariant for action of free groups. We do not present the f-invariant in its general setting. In our context, considering only Markov measures is justified by the fact that they maximise the f-invariant [Bow10, Cor. 11.2].

Remark. Curisously, the formula for the f-invariant given in (8.7), which is the one that seems to make sense in our context, happens to be a very slight modification of the one of Bowen [Bow10, Cor. 7.6]. This is something we would like to better understand. For a free group generated by r elements, the formula of Bowen is the same with d replaced by $2r-1$, which is coherent in some sense since the degree of the tree is then equal to $2r$. But in what follows, we do not assume that d is odd.

In the next section, we will present a construction for a SFT Σ_A^d of a particular measure having the property to be Markov-uniform. We will then study this measure in the light of the f-invariant, and show that there is a strong connection between being Markov-uniform and maximising the f-invariant.

8.3.2 Construction of Markov-uniform measures

In the context of the free monoid F_d with d generators, we define the boundary ∂S of a set $S \subset F_d \setminus \{\varepsilon\}$ by:

$$\partial S = \{w \in F_d \setminus S\,; \exists u \in S, \exists i \in \{1,\dots,d\}, w = ua_i \text{ or } u = wa_i\}.$$

For a subset $S \subset T_{d+1} = \langle g_1, \dots, g_d, g_{d+1} \mid g_i^2 = 1 \rangle$, one can also define:

$$\partial S = \{w \in T_{d+1} \setminus S\,; \exists u \in S, \exists i \in \{1, \dots, d, d+1\}, w = ug_i\}.$$

We will focus on measures that are Markov-uniform with respect to these boundaries. As already mentioned, to construct such a Markov-uniform measure π for a SFT Σ_A^d defined on F_d, it is natural to consider a measure defined by a stochastic matrix P, such that the nodes of the tree are labeled successively in a Markovian way, using the transition matrix P. Let us see if there exists a matrix P inducing a Markov measure on Σ_A^d that would be Markov-uniform. Like for Parry measure, we can expect some strong independence property and search P under a Markov multiplicative form, that is, under the form:

$$P_{i,j} = A_{i,j} \frac{r(j)}{\sum_{k=1}^n A_{i,k} r(k)} = A_{i,j} \frac{r(j)}{r(\mathcal{S}(i))}$$

for some probability vector r.

Furthermore, in order for the measure obtained to be Markov-uniform, for any $i, k_1, \dots, k_d$, the value $P_{i,j} \prod_{t=1}^d P_{j,k_t}$ should not depend on the letter j such that $A_{i,j} \prod_{t=1}^d A_{j,k_t} = 1$. In particular, if we take $k_1 = \ldots = k_d = k$, we should have:

$$P_{i,j} P_{j,k}^d = A_{i,j} A_{j,k} \frac{r(j)}{r(\mathcal{S}(i))} \Big(\frac{r(k)}{r(\mathcal{S}(j))} \Big)^d,$$

so that the quantity

$$\frac{r(j)}{r(\mathcal{S}(j))^d} = \frac{r(j)}{(\sum_{k=1}^n A_{j,k} r(k))^d}$$

should not depend on the letter j such that $A_{i,j} A_{j,k} = 1$. It is thus natural to search r satisfying $\sum_{s=1}^n A_{j,s} r(s) = \lambda r(j)^{1/d}$ for some constant λ.

We will use the following extension of the weak form of Perron-Frobenius theorem to prove that it is always possible to find a suitable probability vector r.

Proposition 8.4. *Let A be an irreducible non-negative matrix, and let $d \geq 1$. There exists $\lambda > 0$ and $r_1, \dots, r_n > 0$ satisfying $\sum_{i=1}^n r_i = 1$ and:*

$$A \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \lambda \begin{pmatrix} r_1^{1/d} \\ \vdots \\ r_n^{1/d} \end{pmatrix}.$$

Proof. Let us consider the set $S = \{x \in \mathbb{R}_+^n; \sum_{i=1}^n x_i = 1\}$. The set S is a convex compact set of $\mathbb{R}^n$. We define a function $\mathcal{F} : S \to S$ by

$$\mathcal{F}(x) = \frac{1}{\left\| A \begin{pmatrix} x_1^d \\ \vdots \\ x_n^d \end{pmatrix} \right\|_1} A \begin{pmatrix} x_1^d \\ \vdots \\ x_n^d \end{pmatrix}.$$

One can check that the function $\mathcal{F}$ is well-defined and continuous. Consequently, by Brouwer fixed point theorem, there exists $x \in S$, such that $\mathcal{F}(x) = x$. Let us set $\alpha = \left\| A \begin{pmatrix} x_1^d \\ \vdots \\ x_n^d \end{pmatrix} \right\|_1$ and

$y_i = x_i^d$ for $1 \leq i \leq d$. We obtain:

$$A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \alpha \begin{pmatrix} y_1^{1/d} \\ \vdots \\ y_n^{1/d} \end{pmatrix},$$

so that for $\lambda = \alpha \|y\|_1^{\frac{1}{d}-1}$ and $r = \frac{1}{\|y\|_1} y$, we have $\sum_{i=1}^n r_i = 1$ and:

$$A \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \lambda \begin{pmatrix} r_1^{1/d} \\ \vdots \\ r_n^{1/d} \end{pmatrix}.$$

As we will see in Sec. 8.3.4, unlike the one-dimensional case, the map $\mathcal{F}$ is not contractive in general for $d \geq 2$.

Proposition 8.5. *Let r be a probability vector satisfying*

$$A \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \lambda \begin{pmatrix} r_1^{1/d} \\ \vdots \\ r_n^{1/d} \end{pmatrix}. \tag{8.8}$$

Then the Markov measure on F_d defined by the transition matrix

$$P_{i,j} = A_{i,j} \frac{r_j}{\sum_{k=1}^n A_{i,k} r_k} = A_{i,j} \frac{r_j}{\lambda r_i^{1/d}}$$

*is a Markov-uniform measure for Σ_A^d. We will call such a measure a d-*Parry *measure.*

Proof. The value $\frac{r_j}{\lambda r_i^{1/d}} \prod_{t=1}^d \frac{r_{k_t}}{\lambda r_j^{1/d}}$ does not depend on j, so that it is easy to check that for any finit set, the measure is Markov-uniform. □

8.3.3 The f-invariant of d-Parry measures

Let us consider a symmetric nearest neighbour system on F_d, that is, a SFT on F_d defined by a symmetric adjacency matrix A. Let P be a transition matrix defined as in the previous subsection by:

$$P_{i,j} = A_{i,j} \frac{r_j}{\lambda r_i^{1/d}}, \quad \text{where } A \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \lambda \begin{pmatrix} r_1^{1/d} \\ \vdots \\ r_n^{1/d} \end{pmatrix}.$$

Then, the invariant measure of the Markov chain P is given by:

$$\pi_i = \alpha \lambda r_i^{1+1/d}, \quad \text{where } \alpha = \frac{1}{\lambda \sum_{i=1}^n r_i^{1+1/d}},$$

and the Markov chain in reversible.

Proposition 8.6. *Let $\mathcal{M}_{\Sigma_A^d}$ be the set of translation invariant measures on the symmetric SFT Σ_A^d, and let $\mu \in \mathcal{M}_{\Sigma_A^d}$. The measure μ maximising the f-invariant is a d-Parry measure. Conversely, d-Parry measures are (local) extrema of the f-invariant.*

Proof. As already mentioned, Markov chains maximise the f-invariant [Bow10, Cor. 11.2].

We follow the steps used on $\mathbb{Z}$ by Kitchens [Kit98, Paragraph 6.2]. A one-step Markov measure ν, stationary measure of Q, is uniquely defined by a set of weights $x_{i,j} = \nu(i)Q_{i,j}$ satisfying:

- $x_{i,j} \geq 0$,
- $\sum_{i=1}^n \sum_{j=1}^n x_{i,j} = 1$,
- $\forall j \in \{1, \dots, n\}, \sum_{i=1}^n x_{i,j} = \sum_{k=1}^n x_{j,k}$.

The f-invariant is then given by:

$$f(x) = d\sum_{i=1}^n (\sum_{j=1}^n x_{i,j}) \log \sum_{j=1}^n x_{i,j} - \frac{d+1}{2} \sum_{i=1}^n \sum_{j=1}^n x_{i,j} \log x_{i,j}.$$

We want to maximise the function f under the above constraints:

$$g_j(x) = \sum_{i=1}^n x_{i,j} - \sum_{k=1}^n x_{j,k} = 0 \text{ for all } j \in \{1, \dots, n\}$$

$$g(x) = 1 - \sum_{i=1}^n \sum_{j=1}^n x_{i,j} = 0.$$

Making use of Lagrange multipliers, we introduce the function:

$$F(x, \kappa, \eta) = f(x) + \sum_{s=1}^n \kappa_s g_s(x) + \eta g(x),$$

and compute, for i, j such that $A_{i,j} = 1$,

$$\frac{\partial F}{\partial x_{i,j}} = d(\log \sum_{k=1}^n x_{i,k} + 1) - \frac{d+1}{2}(\log x_{i,j} + 1) + \kappa_i(\delta_{i,j} - 1) + \kappa_j(1 - \delta_{i,j}) - \eta.$$

$$= \frac{-1}{2} \log \frac{x_{i,j}^{d+1}}{(\sum_{k=1}^n x_{i,k})^{2d}} + \frac{d-1}{2} + \kappa_j - \kappa_i - \eta.$$

This is equal to 0 if and only if:

$$\frac{x_{i,j}^{d+1}}{(\sum_{k=1}^n x_{i,k})^{2d}} = \exp(d - 1 - 2\kappa_i + 2\kappa_j - 2\eta).$$

Let us define: $\alpha_i = e^{\frac{2\kappa_i}{d+1}}$ for $i \in \{1, \dots, n\}$, and $\beta = e^{\frac{d-1-2\eta}{d+1}}$. In order to simplify the notations, we also define: $p_i = \sum_{j=1}^n x_{i,j}$, so that:

$$\frac{x_{i,j}}{p_i^{\frac{2d}{d+1}}} = \beta \frac{\alpha_j}{\alpha_i}. \tag{8.9}$$

If $x_{i,j} = x_{j,i}$ (symmetry condition), we obtain:

$$\frac{p_i^{\frac{2d}{d+1}}}{\alpha_i^2} = \frac{p_j^{\frac{2d}{d+1}}}{\alpha_j^2},$$

so that there exists a constant γ such that $p_i^{\frac{d}{d+1}} = \gamma\alpha_i$. Consequently, if $A_{i,j} = 1$, we have:

$$x_{i,j} = \beta\gamma^2\alpha_i\alpha_j.$$

This leads to:

$$p_i = \gamma^{1+\frac{1}{d}}\alpha_i^{1+\frac{1}{d}} = \sum_{j=1}^{n} x_{i,j} = \sum_{j=1}^{n} A_{i,j}\beta\gamma^2\alpha_i\alpha_j.$$

Thus,

$$\sum_{j=1}^{n} A_{i,j}\alpha_j = \beta^{-1}\gamma^{\frac{1}{d}-1}\alpha_i^{\frac{1}{d}} = \lambda\alpha_i^{1/d},$$

where $\lambda = \beta^{-1}\gamma^{\frac{1}{d}-1}$ and the transition matrix is given by:

$$Q_{i,j} = \frac{x_{i,j}}{p_i} = \frac{\alpha_j}{\lambda\alpha_i^{1/d}},$$

meaning that the measure is a d-Parry measure. □

8.3.4 Examples

Example 8.7. Let us consider the SFT on F_d corresponding to the Fibonacci constraint: the alphabet is $\mathcal{A} = \{0, 1\}$ and the matrix is $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, meaning that it is forbidden for two consecutive nodes to be both in state 1. A d-Parry measures will be given by a transition matrix

$$P = \begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix}, \quad \text{with} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix} = \lambda\begin{pmatrix} \alpha^{1/d} \\ (1-\alpha)^{1/d} \end{pmatrix},$$

where α, λ are some positive real numbers. We thus obtain: $\alpha^{1+1/d} = (1-\alpha)^{1/d}$, that is:

$$\alpha^{d+1} = 1 - \alpha.$$

For any $d \geq 1$, there exists a unique d-Parry measure, which is the Markov-multiplicative measure defined by $r(0) = \alpha$ and $r(1) = 1 - \alpha$, where α is the unique positive solution of the equation

$$\alpha^{d+1} = 1 - \alpha.$$

For $d = 1$, we recover $r(0) = \frac{1}{\varphi}$ and $r(1) = \frac{1}{\varphi^2}$.

In the case of the Fibonacci constraint, the fixed point of Prop. 8.4 is thus unique. But the application $\mathcal{F}$ defined in the proof of Prop. 8.4 is not contractive, and it can have for example orbits of period 2.

Let $\beta \in (0, 1)$ satisfy $\beta = (1-\beta)(1+\beta^d)^d$, and let $\alpha = 1/(1+\beta^d)$. If we set $\lambda_1 = \beta^{-1/d}$ and $\lambda_2 = \alpha^{-1/d}$, then by construction, we have:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix} = \lambda_1\begin{pmatrix} \beta^{1/d} \\ (1-\beta)^{1/d} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} = \lambda_2\begin{pmatrix} \alpha^{1/d} \\ (1-\alpha)^{1/d} \end{pmatrix}.$$

For $d \leq 4$, the equation $\beta = (1-\beta)(1+\beta^d)^d$ has only one root, and we find $\alpha = \beta$, so that we recover the d-Parry measure. But for $d \geq 5$, this provides periodic orbits of period 2. Let $P = \begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix}$ and $\tilde{P} = \begin{pmatrix} \beta & 1-\beta \\ 1 & 0 \end{pmatrix}$ be the two transition matrices, and let $\pi = \begin{pmatrix} \frac{\beta}{\alpha+\beta-\alpha\beta} & \frac{\alpha-\alpha\beta}{\alpha+\beta-\alpha\beta} \end{pmatrix}$, as well as $\tilde{\pi} = \begin{pmatrix} \frac{\alpha}{\alpha+\beta-\alpha\beta} & \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta} \end{pmatrix}$. These two probability measures satisfy: $\pi P = \tilde{\pi}$ and $\tilde{\pi}\tilde{P} = \pi$. Let us choose the label of a given node according to the probability π, and label its neighbours using the transition matrix P, and the next nodes with $\tilde{P}$, and so on, using alternatively P and $\tilde{P}$. The measure obtained is Markov-uniform: indeed, since we have both: $P_{0,1} = P_{0,0}\tilde{P}^d_{0,0}$ and $\tilde{P}_{0,1} = \tilde{P}_{0,0}P^d_{0,0}$, conditionally to a given boundary, the probability to have some pattern is equal to the probability to have the "all zero" pattern.

More generally, each time we have a periodic orbit, it provides a Markov-uniform measure, which is not stationary in the sense that it is only left invariant by some power of the shift, not by a single action of the shift.

Example 8.8. Let us consider the subshift on $\mathcal{A} = \{1, 2, 3\}$ defined by the transition matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

This matrix is not symmetric, but we can look at d-Parry measures anyway, and it will give us Markov-uniform measures. For any $d \geq 1$, an elementary solution of (8.8) is given by the vector $r = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$, so that the transition matrix $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$ of invariant probability $\pi = \begin{pmatrix} 1/3 & 1/6 & 1/2 \end{pmatrix}$, defines a d-Parry measure for any $d \geq 1$. Nevertheless, the measure obtained is not an extremum of the f-invariant as defined by (8.7). Indeed, the condition of (8.9) cannot be satisfied. If $A_{i,j} = 1$, then $x_{i,j} = \pi_i/2$, so that:

$$\frac{x_{i,j}}{\pi_i^{\frac{2d}{d+1}}} = \frac{1}{2}\pi_i^{\frac{1-d}{d+1}}.$$

If there were real β and $\alpha_1, \alpha_2, \alpha_3$ such that:

$$\frac{1}{2}\pi_i^{\frac{1-d}{d+1}} = \beta\frac{\alpha_j}{\alpha_i}$$

as soon as $A_{i,j} = 1$, it would imply $\alpha_1 = \alpha_2 = \alpha_3$ and $\pi_1 = \pi_2 = \pi_3$, which is not the case.

8.4 Fundamental link with PCA

8.4.1 SFT on $\mathbb{Z}$

Let us consider the nearest-neighbour SFT Σ_A over the alphabet $\mathcal{A} = \{1, \ldots, n\}$, defined by the adjacency matrix $A \in \mathcal{M}_n(\{0, 1\})$. We consider a PCA F_A on $\mathcal{A}^{\mathbb{Z}}$ of neighbourhood $\mathcal{N} = \{0, 1\}$ and local function f_A satisfying, for $i, j \in \mathcal{A}$ such that $(A^2)_{i,j} \geq 1$,

$$f_A(i, j)(k) = \frac{1}{(A^2)_{i,j}} A_{i,k}A_{k,j}$$

(for i, j such that $(A^2)_{i,j} = 0$, we do not assume anything on $f_A(i, j)$). The value $(A^2)_{i,j}$ is equal to the number of letters k such that $A_{i,k} = A_{k,j} = 1$, that is, the number of letters k such that ikj is an allowed pattern. By definition, the measure $f_A(i, j)$ is thus uniform on all letters k such that ikj is allowed.

Like in Chap. 4, for symmetry, we can represent the space-time diagram of F_A on a triangular lattice. Two consecutive steps of time of the evolution of the PCA then correspond to a labeling of the graph Γ represented on the right of Fig. 8.4. With the terminology of Chap. 1, Sec. 1.4, we call this graph Γ the doubling graph of F_A. It is one-to-one with $\mathbb{Z}$.

Let $\overline{\overline{\mu}}$ be the measure on $\mathcal{A}^\Gamma$ corresponding to one iteration of the PCA from the measure μ, so that in particular, the projection of $\overline{\overline{\mu}}$ on the top line of Γ is equal to μF_A. We say that $\overline{\overline{\mu}}$ corresponds to a reversible behaviour of the PCA if when reversing time (symmetry of Γ of horizontal axis), the measure obtained still corresponds to an iteration of the PCA F_A.

The next proposition indicates that the Parry measure π of Σ_A, seen as a measure on Γ by "folding the graph $\mathbb{Z}$", as suggested in Fig. 8.4, corresponds to a reversible behaviour of the PCA F_A.

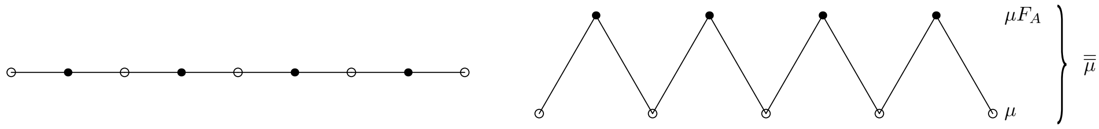
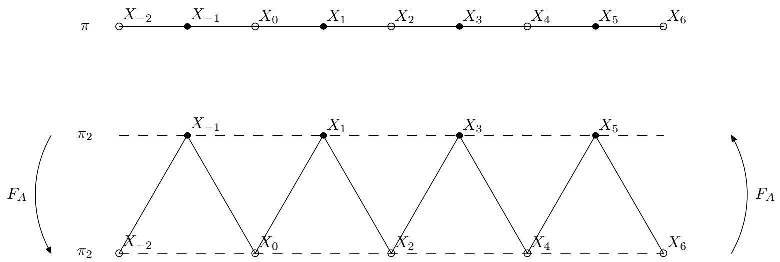
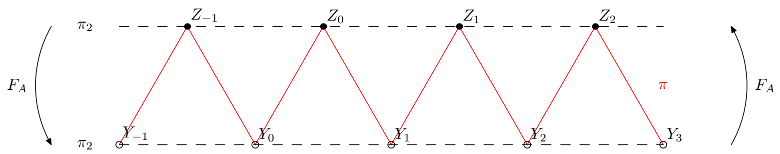


Figure 8.4: The lattice $\mathbb{Z}$, interpreted on the right as a doubling graph Γ.

Proposition 8.7. *Let $(X_i)_{i\in\mathbb{Z}}$ be distributed according to the Parry measure π of Σ_A. Then, the two sequences $(X_{2i})_{i\in\mathbb{Z}}$ and $(X_{2i+1})_{i\in\mathbb{Z}}$ have the same distribution π_2, which is an invariant measure of the PCA F_A.*



Furthermore, if $(Y_i)_{i\in\mathbb{Z}}$ is distributed according to π_2, and if $(Z_i)_{i\in\mathbb{Z}}$ is the image of $(Y_i)_{i\in\mathbb{Z}}$ by the PCA F_A, then the sequence $(Y_i, Z_i)_{i\in\mathbb{Z}}$ is distributed according to π.



Proof. Since the Parry measure is shift-invariant, the two sequences $(X_{2i})_{i\in\mathbb{Z}}$ and $(X_{2i+1})_{i\in\mathbb{Z}}$ have the same distribution π_2. Moreover, by Thm. 8.1, the Parry measure is a Markov random field, so that for any $a \le b$, we have:

$$\mathbb{P}((X_{2i})_{a\le i\le b} = (x_{2i})_{a\le i\le b} | (X_{2i+1})_{a-1\le i\le b}) = (x_{2i+1})_{a-1\le i\le b}))$$

$$= \prod_{i=a}^{b} \mathbb{P}(X_{2i} = x_{2i} | X_{2i-1} = x_{2i-1}, X_{2i+1} = x_{2i+1}).$$

And this random field being Markov-uniform (still by Thm 8.1), the probabilities $\mathbb{P}(X_{2i} = x_{2i} | X_{2i-1} = x_{2i-1}, X_{2i+1} = x_{2i+1})$ are exactly given by $f_A(x_{2i-1}, x_{2i+1})(x_{2i})$. It proves that π_2 is an invariant measure of F_A, and that if $(Y_i)_{i\in\mathbb{Z}}$ is distributed according to π_2, and if $(Z_i)_{i\in\mathbb{Z}}$ is the image of $(Y_i)_{i\in\mathbb{Z}}$ by the PCA F_A, then the sequence $(Y_i, Z_i)_{i\in\mathbb{Z}}$ is distributed according to π. □

Let us come back to the fact that the Parry measure π of Σ_A, seen as a measure on Γ, corresponds to a reversible behaviour of the PCA F_A. We index the sites of Γ from left to right by the integers (as represented with the X_i in Prop. 8.7), and define a pair potential on Γ by

$$\varphi_{n,n+1}(i,j) = \begin{cases} +\infty \text{ if } A_{i,j} = 0, \\ 0 \text{ otherwise.} \end{cases}$$

The following properties are satisfied.

- A Markov-uniform measure on the SFT Σ_A is a Gibbs measure of potential φ.
- The Gibbs measures $\overline{\overline{\mu}}$ of potential φ, seen as measures on Γ, correspond to a reversible behaviour of the PCA F_A, see Sec. 1.4.3.

Since the Parry measure π is a Markov-uniform measure, by Thm. 8.1 (or Thm. 8.2 for the generalisation to any dimension $d \geq 1$), we obtain that seen as a measure on Γ, π corresponds to a reversible behaviour of the PCA F_A.

Example 8.9. With the notations of Chap. 4, for the Fibonacci SFT, the PCA F_A obtained is the PCA of parameters:

$$\theta_{00} = 1/2 \text{ and } \theta_{01} = \theta_{10} = \theta_{11} = 0.$$

Observe that we recover the directed animal PCA of Sec. 2.3. One can check that it satisfies the conditions of Th. 4.2 for having a Markov invariant measure.

Instead of considering Markov-uniform measures, one can also attach a weight $\gamma(\alpha) > 0$ to each letter α of $\mathcal{A}$, and look at the Markov measure π on $\mathcal{A}^{\mathbb{Z}}$ such that for $a, b \in \mathcal{A}$, the measure $\pi(awb)$ of the words $w \in \mathcal{W}(A,k)$ such that $awb \in \mathcal{W}(A, k+2)$ would be proportional to

$$\prod_{i=1}^{k} \gamma(w_i) = \prod_{\alpha \in \mathcal{A}} \gamma(\alpha)^{|w|_\alpha}.$$

It corresponds to replacing the matrix A by a matrix B defined by $B_{i,j} = \gamma(i) A_{i,j}$, and the analogue of the Parry measure is the Markov-multiplicative measure defined by the transition matrix:

$$P_{i,j} = B_{i,j} \frac{r(j)}{\lambda r(i)} = A_{i,j} \frac{\gamma(i) r(j)}{\lambda r(i)},$$

where $Br = \lambda r$ and $\sum_{i=1}^{n} r(i) = 1$.

In terms of Gibbs measures, it amounts to consider the pair potential defined by:

$$\varphi_{n,n+1}(i,j) = \begin{cases} +\infty \text{ if } A_{i,j} = 0, \\ -\log \sqrt{\gamma(i)\gamma(j)} \text{ otherwise.} \end{cases}$$

and the PCA involved satisfies

$$f(i,j)(k) = \frac{1}{\sum_{\ell \in \mathcal{A}} A_{i,\ell} A_{\ell,j}\, \gamma(\ell)} \gamma(k) = \frac{1}{B^2(i,j)} B_{i,k} B_{k,j}.$$

8.4.2 SFT on $\mathbb{Z}^d$ and on regular trees

For the generalisation to $\mathbb{Z}^d$ and regular trees, let us consider for simplicity a SFT defined by a symmetric adjacency matrix A. Once again, the Markov-uniform measures are the Gibbs measures corresponding to the pair potential defined for any edge v of the graph by:

$$\varphi_v(i,j) = \begin{cases} +\infty \text{ if } A_{i,j} = 0, \\ 0 \text{ otherwise.} \end{cases}$$

The graph $\mathbb{Z}^d$ as well as regular trees are bipartite graphs. If we color the sites that are at an odd distance of the origin in black and the sites that are at an even distance of the origin in white, we obtain two isomorphic subgraphs. Like in the case of $\mathbb{Z}$, we can deform the graph by shifting up all the black sites of one unit, and see the graph obtained as the doubling graph Γ of a PCA (for $\mathbb{Z}^2$, this is represented on Fig. 8.5). The neighbourhood of a (black) cell is the set of (white) cells to which it is connected, and the local function is such that $f((x_v)_{v\in\mathcal{N}})$ is the uniform measure on all letters providing an allowed pattern. Again, Markov-uniform measures of the SFT correspond to a reversible behaviour of the PCA.

Example 8.10. For the Fibonacci SFT on $\mathbb{Z}^2$, the neighbourhood of the PCA is of size 4, and

$$f(x,y,z,t)(1) = \begin{cases} 1/2 \text{ if } x=y=z=t=0, \\ 0 \text{ otherwise.} \end{cases}$$

We recover the PCA introduced by Eloranta [Elo96]. It is the two-dimensional analogue of the directed animals PCA. No explicit form of the invariant measure is known.

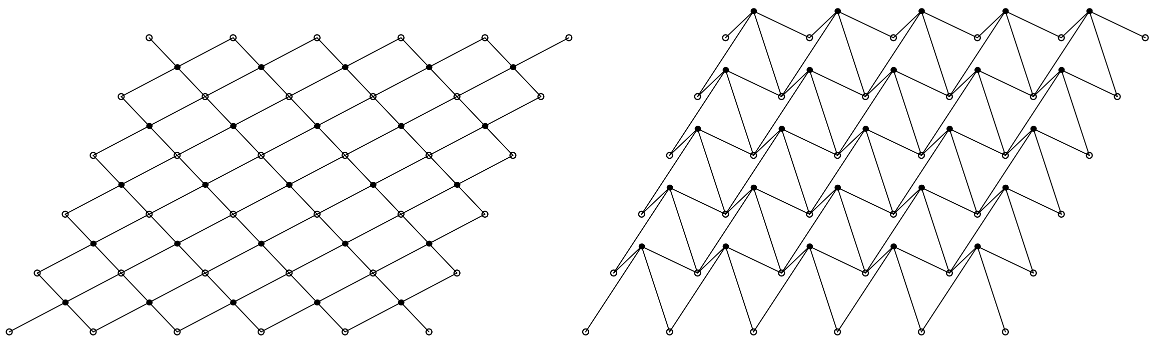


Figure 8.5: The lattice $\mathbb{Z}^2$, interpreted on the right as a doubling graph Γ.

If we add weights $\gamma(i)$ (also known as *activity parameters*) on the letters of the alphabet, it amounts to considering the pair potential

$$\varphi_v(i,j) = \begin{cases} +\infty \text{ if } A_{i,j}=0 \\ -\log\gamma(i)^{1/k}\gamma(j)^{1/k} \text{ otherwise,} \end{cases}$$

where k is the degree of the graph (so that $k=2d$ for $\mathbb{Z}^d$, and $k=d+1$ for F_d or T_{d+1}), and the corresponding PCA is such that $f((x_v)_{v\in\mathcal{N}})(y)$ is proportional to $\gamma(y)$ for the letters y providing allowed patterns.

Gibbs measures on $\mathbb{Z}^d$ are extensively studied in probability theory. Since the work of Zachary [Zac83], Gibbs measures on trees have also given rise to many works, one motivation being to model computer communication systems [Kel85].

Conclusion and future work

Des guirlandes d'étoiles descendaient du ciel noir au-dessus des palmiers et des maisons.

– Albert Camus, *L'exil et le royaume*

In the different chapters of this thesis, we have illustrated how particular measures on symbolic spaces with strong combinatorial structure are involved in the study of probabilistic cellular automata and of other stochastic dynamics.

Bernoulli product measures appear naturally at the forefront. We have seen that PCA having Bernoulli product invariant measures give rise to interesting space-time diagrams having special properties, with weak and non-trivial correlation structure, every line being constituted of i.i.d. random variables and PCA appearing in different directions. The discussion on Bernoulli invariant measures was carried on in Chap. 5 for deterministic CA. In Chap. 6, Bernoulli product measures are also the natural initial measures for defining the density classification problem on infinite lattices and trees. And of course, we meet again Bernoulli product measures when defining the random walks of Chap. 7, and also in the different ways to generate measures of maximal entropy presented in Chap. 8.

After Bernoulli product measures, when considering the next level of complexity, we come across Markov measures. Space-time diagrams of PCA are always Markov random fields, as mentioned in Chap. 1. When studying the conditions of reversibility for PCA, Markov fields on the doubling graph corresponding to two successive time steps also appear to be the adapted tool. In Chap. 4, we have generalised to Markov measures the approach used to compute the image of a Bernoulli product measure by a PCA, and given a characterisation of simple PCA having a Markov invariant measure. Some of them are related to the counting of directed animals, which make them processes of particular interest. In Chap. 7 and Chap. 8, a specialisation of Markov measure plays an important role: Markov-multiplicative measures, which can be seen as product measures conditioned to avoid some patterns. More generally, the measures that we have introduced under the name of Markov-uniform measures are the central object of Chap. 8: they are Markov measures that are uniform on all allowed patterns, conditionally to any fixed value of the boundary.

To summarise, these particular measures play a fundamental role. But for general PCA, we have no simple description of the equilibrium behaviour. For example, if we consider a PCA of alphabet $\mathcal{A} = \{0,1\}$ and neighbourhood $\mathcal{N} = \{0,1\}$, defined by four parameters $\theta_{00}, \theta_{01}, \theta_{10}, \theta_1$ giving the probabilities to obtain a 1 for the different values of the neighbourhood, there is no general tool to express the invariant(s) measure(s) of the PCA as a function of these parameters.

The PCA defined by $\theta_{00} = \theta_{01} = \theta_{10} = a$, and $\theta_{11} = 1 - a$, for some $a \in (0,1)$, has drawn a specific interest, due to a connection with directed animals and percolation theory first noticed by Dhar [Dha83]. More specifically, determining explicitly the invariant measure for the above PCA would enable to compute the area and perimeter generating function of directed animals in the square lattice, and to compute the directed site-percolation threshold in the square lattice [BM98, LBM07, Mar12].

The theoretical understanding of the swarming model presented in Chap. 2 is also a challenge, as is the one of the majority-flip PCA of Chap. 3. Generally, such phase transition phenomena appear be very difficult to analyse. And the simulations are to be handled with care, since it is extremely difficult to deal experimentally with both the infinite set of cells and the infinite time, corresponding to the equilibrium behaviour. In particular, there are very few known results relating the asymptotic behaviour of the restriction of a PCA to finite windows and its asymptotic behaviour on an infinite lattice. The perfect sampling procedure of Chap. 3 takes all its importance in that context.

In fact, if we consider a PCA defined as above by four parameters $\theta_{00}, \theta_{01}, \theta_{10}, \theta_1$, it is not even known if such a PCA is ergodic as soon as these four values are strictly between 0 and 1 (positive-rate PCA). Since the work Gács, it is known that there exist one-dimensional PCA with positive rates that are non ergodic [Gác86, Gác01]. But the known examples being very complex, the positive rates problem is still open if we restrict it to elementary PCA of alphabet and neighbourhood of size 2.

In the domain of deterministic CA, Chap. 5 leaves many open questions. We have mentioned the existence of surjective and state-conserving CA having no direction of equicontinuity. But above all, a great challenge would be to be able to prove randomisation (or to begin with, randomisation in Cesáro mean) for a non-affine CA.

Concerning the density classification problem of Chap. 6, the central open question is the existence of a CA, probabilistic or deterministic, that would classify the density on $\mathbb{Z}$. Let us also mention another open problem: there is no known family of PCA able to classify the density with an arbitrary precision on two-dimensional finite grids. We are thus in a surprising situation:

- in one dimension, we know a simple PCA that classifies the density with an arbitrary precision on finite rings, namely the majority-traffic PCA [Fat11, Fat13]; but the density classification seems to be a difficult problem on the infinite lattice $\mathbb{Z}$;
- in two dimensions, we know a simple CA that classifies the density on the infinite lattice $\mathbb{Z}^2$, namely Toom CA; but it seems difficult to design PCA classifying the density with an arbitrary precision on finite grids.

In Chap. 7, we have developed tools to describe the harmonic measure of a random walk on free products of groups. Let us relax slightly that framework and consider for example the group defined by the presentation $G =< a, b, c, d | ac = ca, ad = da, bd = db >$. The elements of the group can be represented by heaps of pieces, as suggested in Fig. 8.6. The group G is not a free product, but it is an algamated free product, which still allows to exploit some Markov-multiplicative structure. But we are not anymore able to have a satisfactory formula for the drift, corresponding to the growth rate of the height of the heap.

Finally, in Chap. 8, the work on subshifts of finite type defined on trees is still in progress. In particular, we would like to have a better understanding of the f-invariant, and see if it is related with some analogue of the topological entropy on trees.

We also wish to analyse further the connection with PCA. The initial motivation for the work of Chap. 8 was to develop some tools that could possibly be used to obtain some information on the measures of maximal entropy of SFT defined on $\mathbb{Z}^d, d \geq 2$. Let us recall that already for the Fibonacci on $\mathbb{Z}^2$, very little is known on the measure of maximal entropy, and on the topological entropy. This remains an ambitious challenge.

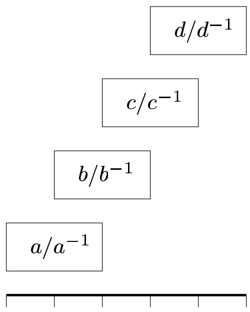


Figure 8.6: Representation of the group $G =< a, b, c, d \,|\, ac = ca, ad = da, bd = db >$ by a heap of pieces.

Bibliography

[AB12] N. Aubrun and M.-P. Béal. Tree-shifts of finite type. *Theoret. Comput. Sci.*, 459:16–25, 2012.

[Alb09] M. Albenque. A note on the enumeration of directed animals via gas considerations. *Ann. Appl. Probab.*, 19(5):1860–1879, 2009.

[AST13] P. Arrighi, N. Schabanel, and G. Theyssier. Stochastic cellular automata: Correlations, decidability and simulations. *Fund. Inform.*, 126(2-3):121–156, 2013.

[BF05] V. Belitsky and P. A. Ferrari. Invariant measures and convergence properties for cellular automaton 184 and related processes. *J. Stat. Phys.*, 118(3-4):589–623, 2005.

[BFC11] O. Bouré, N. Fatès, and V. Chevrier. Observation of novel patterns in a stressed lattice-gas model of swarming. Research report, October 2011.

[BFC13] O. Bouré, N. Fatès, and V. Chevrier. A robustness approach to study metastable behaviours in a lattice-gas model of swarming. In J. Kari, M. Kutrib, and A. Malcher, editors, *Cellular Automata and Discrete Complex Systems*, volume 8155 of *Lecture Notes in Computer Science*, pages 84–97. Springer Berlin Heidelberg, 2013.

[BFMM12] A. Bušić, N. Fatès, J. Mairesse, and I. Marcovici. Density classification on infinite lattices and trees. In D. Fernández-Baca, editor, *LATIN*, volume 7256 of *Lecture Notes in Computer Science*, pages 109–120. Springer, 2012.

[BFMM13] A. Bušić, N. Fatès, J. Mairesse, and I. Marcovici. Density classification on infinite lattices and trees. *Electron. J. Probab.*, 18:51, 1–22, 2013.

[BGM69] J. K. Beljaev, J. I. Gromak, and V. A. Malyšev. Invariant random Boolean fields. *Mat. Zametki*, 6:555–566, 1969.

[BM98] M. Bousquet-Mélou. New enumerative results on two-dimensional directed animals. In *Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995)*, volume 180, pages 73–106, 1998.

[BMM11] A. Bušić, J. Mairesse, and I. Marcovici. Probabilistic cellular automata, invariant measures, and perfect sampling. In *28th International Symposium on Theoretical Aspects of Computer Science*, volume 9 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 296–307. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2011.

[Bow10] L. Bowen. Non-abelian free group actions: Markov processes, the Abramov-Rohlin formula and Yuzvinskii's formula. *Ergodic Theory Dynam. Systems*, 30(6):1629–1663, 2010.

[BS94] R. Burton and J. E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14(2):213–235, 1994.

[BS95] R. Burton and J. E. Steif. New results on measures of maximal entropy. *Israel J. Math.*, 89(1-3):275–300, 1995.

[Bur56] P. J. Burke. The output of a queuing system. *Operations Res.*, 4:699–704 (1957), 1956.

[CD91] J. T. Cox and R. Durrett. Nonlinear voter models. In *Random walks, Brownian motion, and interacting particle systems*, volume 28 of *Progr. Probab.*, pages 189–201. Birkhäuser Boston, Boston, MA, 1991.

[CEP96] H. Cohn, N. Elkies, and J. Propp. Local statistics for random domino tilings of the Aztec diamond. *Duke Math. J.*, 85(1):117–166, 1996.

[CM11] P. Chassaing and J. Mairesse. A non-ergodic probabilistic cellular automaton with a unique invariant measure. *Stochastic Process. Appl.*, 121(11):2474–2487, 2011.

[CPY89] K. Culik, II, J. Pachl, and S. Yu. On the limit sets of cellular automata. *SIAM J. Comput.*, 18(4):831–842, 1989.

[CT10] C. F. Coletti and P. Tisseur. Invariant measures and decay of correlations for a class of ergodic probabilistic cellular automata. *J. Stat. Phys.*, 140(1):103–121, 2010.

[Daw77] D. A. Dawson. Stable states of probabilistic cellular automata. *Information and Control*, 34(2):93–106, 1977.

[Daw75] D. A. Dawson. Synchronous and asynchronous reversible Markov systems. *Canad. Math. Bull.*, 17(5):633–649, 1974/75.

[Dha83] D. Dhar. Exact solution of a directed-site animals-enumeration problem in three dimensions. *Phys. Rev. Lett.*, 51(10):853–856, 1983.

[DKT90] R. L. Dobrushin, V. I. Kryukov, and A. L. Toom. *Stochastic cellular systems: ergodicity, memory, morphogenesis.* Nonlinear science. Manchester University Press, 1990.

[DMO05] M. Draief, J. Mairesse, and N. O'Connell. Queues, stores, and tableaux. *J. Appl. Probab.*, 42(4):1145–1167, 2005.

[DP92] P. Dai Pra. *Space-time large deviations for interacting particle systems.* 1992. Thesis (Ph.D.)–Rutgers University.

[DPLR02] P. Dai Pra, P.-Y. Louis, and S. Rœlly. Stationary measures and phase transition for a class of probabilistic cellular automata. *ESAIM Probab. Statist.*, 6:89–104 (electronic), 2002.

[Elo96] K. Eloranta. *Golden Mean Subshift Revised.* Research reports / Helsinki University of Technology, Institute of Mathematics. Helsinki University of Techn., Institute of Mathematics, 1996.

[Fas02] H. Fukś. Nondeterministic density classification with diffusive probabilistic cellular automata. *Phys. Rev. E*, 66:066106, Dec 2002.

[Fat09] N. Fatès. Asynchronism induces second-order phase transitions in elementary cellular automata. *J. Cellular Automata*, 4(1):21–38, 2009.

[Fat11] N. Fatès. Stochastic cellular automata solve the density classification problem with an arbitrary precision. In *28th International Symposium on Theoretical Aspects of Computer Science*, volume 9 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 284–295. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2011.

[Fat13] N. Fatès. Stochastic cellular automata solutions to the density classification problem. *Theory of Computing Systems*, 53(2):223–242, 2013.

[FM05] N. Fatès and M. Morvan. An experimental study of robustness to asynchronism for elementary cellular automata. *Complex Systems*, 16(1):1–27, 2005.

[FMMN00] P. A. Ferrari, A. Maass, S. Martínez, and P. Ney. Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory Dynam. Systems*, 20(6):1657–1670, 2000.

[FSS02] L. R. Fontes, R. H. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Comm. Math. Phys.*, 228(3):495–518, 2002.

[Gác86] P. Gács. Reliable computation with cellular automata. *J. Comput. System Sci.*, 32(1):15–78, 1986.

[Gác90] P. Gács. A Toom rule that increases the thickness of sets. *J. Statist. Phys.*, 59(1-2):171–193, 1990.

[Gác01] P. Gács. Reliable cellular automata with self-organization. *J. Statist. Phys.*, 103(1-2):45–267, 2001.

[GdSM92] P. Gonzaga de Sá and C. Maes. The Gacs-Kurdyumov-Levin automaton revisited. *J. Statist. Phys.*, 67(3-4):507–522, 1992.

[Geo11] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2011.

[Gil07] L. A. Gilch. Rate of escape of random walks on free products. *J. Aust. Math. Soc.*, 83(1):31–54, 2007.

[GKL78] P. Gács, G. L. Kurdyumov, and L. A. Levin. One-dimensional homogeneous media dissolving finite islands. *Problems of Information Transmission*, 14(3):92–96, 1978.

[GKLM89] S. Goldstein, R. Kuik, J. L. Lebowitz, and C. Maes. From PCAs to equilibrium systems and back. *Comm. Math. Phys.*, 125(1):71–79, 1989.

[GR08] P. Guillon and G. Richard. Nilpotency and limit sets of cellular automata. In *Mathematical foundations of computer science 2008*, volume 5162 of *Lecture Notes in Comput. Sci.*, pages 375–386. Springer, Berlin, 2008.

[GR12] F. García-Ramos. Product decomposition for surjective 2-block ncca. *DMTCS Proceedings*, 0(01), 2012.

[Gra87] Lawrence Gray. The behavior of processes with statistical mechanical properties. In H. Kesten, editor, *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, volume 8 of *The IMA Volumes in Mathematics and Its Applications*, pages 131–167. Springer New York, 1987.

[Gra01] L. 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] *J. Statist. Phys.*, 103(1-2):1–44, 2001.

[Gri99] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[Gui80] Y. Guivarc'h. Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire. In *Conference on Random Walks (Kleebach, 1979) (French)*, volume 74 of *Astérisque*, pages 47–98, 3. Soc. Math. France, Paris, 1980.

[Hed69] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory*, 3:320–375, 1969.

[HMM03] B. Host, A. Maass, and S. Martínez. Uniform Bernoulli measure in dynamics of permutative cellular automata with algebraic local rules. *Discrete Contin. Dyn. Syst.*, 9(6):1423–1446, 2003.

[HN98] O. Häggström and K. Nelander. Exact sampling from anti-monotone systems. *Statist. Neerlandica*, 52(3):360–380, 1998.

[How00] C. D. Howard. Zero-temperature Ising spin dynamics on the homogeneous tree of degree three. *J. Appl. Probab.*, 37(3):736–747, 2000.

[Hub04] M. Huber. Perfect sampling using bounding chains. *Ann. Appl. Probab.*, 14(2):734–753, 2004.

[JPS98] W. Jockusch, J. Propp, and P. Shor. Random domino tilings and the arctic circle theorem, 1998.

[Kar92] J. Kari. The nilpotency problem of one-dimensional cellular automata. *SIAM J. Comput.*, 21(3):571–586, 1992.

[Kel85] F. P. Kelly. Stochastic models of computer communication systems. *J. Roy. Statist. Soc. Ser. B*, 47(3):379–395, 415–428, 1985. With discussion.

[Kin73] J. F. C. Kingman. Subadditive ergodic theory. *Ann. Probability*, 1:883–909, 1973. With discussion by D. L. Burkholder, D. Daley, H. Kesten, P. Ney, F. Spitzer and J. M. Hammersley, and a reply by the author.

[Kit98] B. P. Kitchens. *Symbolic dynamics*. Universitext. Springer-Verlag, Berlin, 1998. One-sided, two-sided and countable state Markov shifts.

[KLG12] J. Kari and B. Le Gloannec. Modified traffic cellular automaton for the density classification task. *Fund. Inform.*, 116(1-4):141–156, 2012.

[KM11] Y. Kanoria and A. Montanari. Majority dynamics on trees and the dynamic cavity method. *Ann. Appl. Probab.*, 21(5):1694–1748, 2011.

[KST13] J. Kari, V. Salo, and I. Törmä. Surjective two-neighbor cellular automata on prime alphabets. In *Proceedings 19th International Workshop on Cellular Automata and Discrete Complex Systems (AUTOMATA 2013) – Exploratory Papers*, IFIG Research Report 1302. Institut für Informatik, Universität Gießen, 2013.

[KT12] J. Kari and S. Taati. Conservation laws and invariant measures in surjective cellular automata. In *Automata 2011—17th International Workshop on Cellular Automata and Discrete Complex Systems*, Discrete Math. Theor. Comput. Sci. Proc., AP, pages 113–122. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.

[Kůr03] P. Kůrka. Cellular automata with vanishing particles. *Fund. Inform.*, 58(3-4):203–221, 2003.

[KV80] O. Kozlov and N. Vasilyev. Reversible Markov chains with local interaction. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 451–469. Dekker, New York, 1980.

[LB95] M. Land and R. K. Belew. No perfect two-state cellular automata for density classification exists. *Physical Review Letters*, 74(25):5148–5150, 1995.

[LBM07] Y. Le Borgne and J.-F. Marckert. Directed animals and gas models revisited. *Electron. J. Combin.*, 14(1):Research Paper 71, 36 pp. (electronic), 2007.

[Led01] F. Ledrappier. Some asymptotic properties of random walks on free groups. In *Topics in probability and Lie groups: boundary theory*, volume 28 of *CRM Proc. Lecture Notes*, pages 117–152. Amer. Math. Soc., Providence, RI, 2001.

[Lig05] T. M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[Lin84] D. A. Lind. Applications of ergodic theory and sofic systems to cellular automata. *Phys. D*, 10(1-2):36–44, 1984. Cellular automata (Los Alamos, N.M., 1983).

[LMS90] J. L. Lebowitz, C. Maes, and E. R. Speer. Statistical mechanics of probabilistic cellular automata. *J. Statist. Phys.*, 59(1-2):117–170, 1990.

[Lou02] P.-Y. Louis. *Automates cellulaires probabilistes: mesures stationnaires, mesures de Gibbs associées et ergodicité*. 2002. Thesis (Ph.D.)–Université des Sciences et Technologies de Lille & Politecnico di Milano.

[Mai05] J. Mairesse. Random walks on groups and monoids with a Markovian harmonic measure. *Electron. J. Probab.*, 10:1417–1441 (electronic), 2005.

[Mar12] J.-F. Marckert. Directed animals, quadratic systems and rewriting systems. *Electron. J. Combin.*, 19(3):Paper 45, 31, 2012.

[Miy79] M. Miyamoto. An equilibrium state for a one-dimensional life game. *J. Math. Kyoto Univ.*, 19(3):525–540, 1979.

[Miy94] M. Miyamoto. Stationary measures for automaton rules 90 and 150. *J. Math. Kyoto Univ.*, 34(3):531–538, 1994.

[MM98] A. Maass and S. Martínez. On Cesàro limit distribution of a class of permutative cellular automata. *J. Statist. Phys.*, 90(1-2):435–452, 1998.

[MM07] J. Mairesse and F. Mathéus. Random walks on free products of cyclic groups. *J. Lond. Math. Soc. (2)*, 75(1):47–66, 2007.

[MP13] B. Marcus and R. Pavlov. Approximating entropy for a class of $\mathbb{Z}^2$ Markov random fields and pressure for a class of functions on $\mathbb{Z}^2$ shifts of finite type. *Ergodic Theory Dynam. Systems*, 33(1):186–220, 2013.

[Pac88] N. H. Packard. Adaptation toward the edge of chaos. In *Dynamic patterns in complex systems (Fort Lauderdale, FL, 1987)*, pages 293–301. World Sci. Publ., Teaneck, NJ, 1988.

[Par64] W. Parry. Intrinsic Markov chains. *Trans. Amer. Math. Soc.*, 112:55–66, 1964.

[Par97] K. Park. *Ergodicity and mixing rate of one-dimensional cellular automata.* ProQuest LLC, Ann Arbor, MI, 1997. Thesis (Ph.D.)–Boston University.

[Pip94] N. Pippenger. Symmetry in self-correcting cellular automata. *J. Comput. System Sci.*, 49(1):83–95, 1994.

[PW96] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

[Reg08] D. Regnault. Directed percolation arising in stochastic cellular automata analysis. In *Mathematical foundations of computer science 2008*, volume 5162 of *Lecture Notes in Comput. Sci.*, pages 563–574. Springer, Berlin, 2008.

[Rei57] E. Reich. Waiting times when queues are in tandem. *Ann. Math. Statist.*, 28:768–773, 1957.

[Sab10] M. Sablik. Recherche de mesures invariantes pour l'action conjointe d'un automate cellulaire et du décalage. In *École de Théorie Ergodique*, volume 20 of *Sémin. Congr.*, pages 207–251. Soc. Math. France, Paris, 2010.

[Sal12] V. Salo. On nilpotency and asymptotic nilpotency of cellular automata. In E. Formenti, editor, *DCM*, volume 90 of *EPTCS*, pages 86–96, 2012.

[Sch09] N. Schabanel. *Systèmes complexes & algorithmes.* Thesis (mémoire d'habilitation à diriger des recherches), Université Paris Diderot - Paris 7, 2009.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[SOS09] M. Schüle, T. Ott, and R. Stoop. Computing with probabilistic cellular automata. In Cesare Alippi, M. M. Polycarpou, C. G. Panayiotou, and G. Ellinas, editors, *ICANN (2)*, volume 5769 of *Lecture Notes in Computer Science*, pages 525–533. Springer, 2009.

[Spi71] F. Spitzer. *Random fields and interacting particle systems.* Mathematical Association of America, Washington, D.C., 1971. Notes on lectures given at the 1971 MAA Summer Seminar, Williams College, Williamstown, Mass.

[Sud90] A. Sudbury. The branching annihilating process: an interacting particle system. *Ann. Probab.*, 18(2):581–601, 1990.

[Too80] A. Toom. Stable and attractive trajectories in multicomponent systems. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 549–575. Dekker, New York, 1980.

[Too00] A. Toom. Algorithmical unsolvability of the ergodicity problem for binary cellular automata. *Markov Process. Related Fields*, 6(4):569–577, 2000.

[Too01] A. Toom. *Contours, convex sets, and cellular automata.* Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2001. 23o Colóquio Brasileiro de Matemática. [23rd Brazilian Mathematics Colloquium].

[Vas78] N. B. Vasilyev. Bernoulli and Markov stationary measures in discrete local interactions. In *Developments in statistics, Vol. 1*, pages 99–112. Academic Press, New York, 1978.

[vdBS94] J. van den Berg and J. E. Steif. Percolation and the hard-core lattice gas model. *Stochastic Process. Appl.*, 49(2):179–197, 1994.

[vdBS99] J. van den Berg and J. E. Steif. On the existence and nonexistence of finitary codings for a class of random fields. *Ann. Probab.*, 27(3):1501–1522, 1999.

[Ver76] A. M. W. Verhagen. An exactly soluble case of the triangular Ising model in a magnetic field. *J. Statist. Phys.*, 15(3):219–231, 1976.

[Wal82] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1982.

[Woe00] W. Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 2000.

[Zac83] S. Zachary. Countable state space Markov random fields and Markov chains on trees. *Ann. Probab.*, 11(4):894–903, 1983.