

# Mémoire de Magistère de Mathématiques

Jimmy PAROT,  
université Paris-Sud XI.

★

Responsable : Frédéric PAULIN,  
département de mathématiques d'Orsay.

5 janvier 2016

# Table des matières

<b>1</b>	<b>Première année</b>	<b>2</b>
1.1	UE spécifique de magistère . . . . .	2
1.2	UE supplémentaire de L3 . . . . .	3
1.3	Informatique théorique . . . . .	3
1.4	Stage en lycée . . . . .	3
<b>2</b>	<b>Deuxième année</b>	<b>4</b>
2.1	UE spécifiques de magistère . . . . .	4
2.2	Travail encadré de recherche de M1 . . . . .	5
<b>3</b>	<b>Troisième année</b>	<b>6</b>
3.1	UE supplémentaire de M2 . . . . .	6
3.2	Mémoire de M2 . . . . .	7
3.3	Présentation d'un domaine de recherche . . . . .	8
3.4	Mémoire de magistère . . . . .	8
<b>A</b>	<b>Rapport du stage en lycée</b>	<b>9</b>
<b>B</b>	<b>TER de M1</b>	<b>23</b>
<b>C</b>	<b>Mémoire de M2</b>	<b>40</b>
C.1	Présentation du domaine de recherche : « Les algorithmes génétiques simples » . . . . .	40
C.2	Mémoire : « Évaluation Route Royale et sélection par rang pour les algorithmes génétiques simples, sans croisements » . . . . .	49

# Chapitre 1

## Première année

Ce chapitre détaille la première année que j'ai effectuée en magistère de mathématiques, qui correspond à la troisième année de licence et coïncide avec mon arrivée à Orsay. Ce fut l'année scolaire 2011/2012, où je m'inscrivis donc en L3 Mathématiques Fondamentales et Appliquées (L3 MFA) pour le magistère.

### 1.1 UE spécifique de magistère

Cette première année de magistère comportait un enseignement spécifique s'ajoutant aux autres cours du L3 MFA, de 3 heures par semaine, qui eut lieu au deuxième semestre.

Cet enseignement fut dispensé par deux enseignants-chercheurs, Wendelin WERNER pour le premier quart et le dernier, et Dominique HULIN pour les deux quarts restant du semestre.

Dans le cours de M. WERNER, nous avons parlé de thèmes variés, comme :

- le théorème d'Arzela-Ascoli,
- les ensembles de Cantor,
- les dimensions fractales,
- et les théorèmes du point fixe dans la boule unité en dimension 2 et 3.

Avec Mme HULIN, le cours abordait cette fois les espaces topologiques, avec :

- les théorèmes de Hahn-Banach, analytique et géométrique,
- les enveloppes et polyèdres convexes,
- les espaces compacts et faiblement compacts,
- le théorème de Markov-Kakutani (point fixe),
- et le lemme de Baire.

Ces enseignements furent très intéressants et fort utiles pour la préparation à l'agrégation deux ans plus tard.

## 1.2 UE supplémentaire de L3

La formation du magistère demandait par ailleurs de choisir deux options facultatives du parcours L3 MFA au lieu d'une seule pour les non-magistériens inscrits dans ce parcours.

J'ai personnellement choisi les enseignements *Programmation, algorithmique et complexité* au premier semestre, et *Graphes, algèbre linéaire* au second, qui m'intéressèrent tous les deux beaucoup.

## 1.3 Informatique théorique

Lors de la première année de magistère, il y avait également un projet à réaliser, sauf pour les étudiants inscrits dans l'UE *Informatique théorique* au second semestre.

J'ai fait le choix de suivre cette UE qui m'a paru vraiment passionnante, dans laquelle j'ai appris notamment les notions de décidabilité, non-déterminisme (problèmes NP-complet) et autres, et je n'ai donc pas réalisé de projet cette première année.

## 1.4 Stage en lycée

Enfin, pour clore cette année pourtant bien remplie, le magistère prévoyait de plus un stage de découverte, appelé plutôt « apprentissage hors-murs », d'une durée de 3 semaines au sein d'une entreprise, ou éventuellement d'un lycée pour ceux qui se destinaient préférentiellement à l'enseignement. Ce stage devait avoir lieu pendant les vacances scolaires d'été, à la fin en fait de la première ou de la deuxième année de magistère au choix.

J'ai pu réaliser mon stage en lycée pendant le mois de juin de la première année, puisque les semestres du supérieur se terminent plus tôt que dans le secondaire. Le rapport de stage correspondant, intitulé *L'enseignement des mathématiques dans le secondaire à travers 3 professeurs de lycée*, constitue l'annexe A de ce mémoire<sup>1</sup>.

---

1. Notons par contre que le rapport fourni ici ne comporte que 13 pages alors qu'il en annonce 17, car j'avais probablement inséré manuellement les 4 pages d'annexes dans mon insouciante jeunesse, et je ne les ai donc malheureusement plus aujourd'hui.

# Chapitre 2

## Deuxième année

Ce nouveau chapitre présente ma deuxième année de magistère, en parallèle au parcours M1 MFA, pendant l'année scolaire 2012/2013. Heureusement, celle-ci fut moins chargée en travaux supplémentaires que l'année de L3.

### 2.1 UE spécifiques de magistère

Cette nouvelle année de magistère exigeait tout de même de suivre un enseignement spécifique de plus par semestre par rapport au parcours M1 MFA standard. Ils furent dispensés par les enseignants-chercheurs Frédéric PAULIN pour le premier, et Pierre PANSU pour le deuxième.

Avec M. PAULIN, nous avons eu un cours très dense sur la théorie spectrale et les fonctions harmoniques planes, avec plus précisément :

- les espaces de Hilbert,
- le spectre d'opérateurs, continus (i.e. bornés), compacts, auto-adjoints,
- le calcul fonctionnel et théorème spectral,
- les fonctions harmoniques et théorème de Poisson,
- le problème de Dirichlet et théorème de représentation conforme,
- la décomposition spectrale du laplacien,
- et l'analyse harmonique des sphères.

Dans le cours de M. PANSU, nous avons étudié la notion d'entropie, fortement apparentée à la théorie de l'information, et notamment :

- l'entropie d'une distribution de probabilité,
- l'entropie relative et information mutuelle, entropie conditionnelle,
- la compression de données et un théorème de Shannon,
- l'entropie des processus stationnaires,
- l'équipartition asymptotique,
- la capacité d'un canal de transmission et autre théorème de Shannon,
- l'entropie métrique, puis topologique.

Ces deux cours ne manquèrent pas d'intérêt, et le premier eut certainement une utilité pour l'agrégation également. Cependant, bien que portant sur la branche des probabilités que j'affectionne, je dirai simplement que le deuxième fut un peu trop éloigné des notions généralement étudiées dans le cursus MFA ou même pour l'agrégation, et qu'il ne m'a donc servi que pour la culture générale, ce qui est toutefois bien aussi.

## 2.2 Travail encadré de recherche de M1

Les cours spécifiques sont tout ce que le magistère prévoyait pour cette deuxième année, avec l'apprentissage hors-murs pour certains. Je souhaite cependant noter ici le premier projet de recherche que j'ai réalisé durant ces trois années, qui est le Travail Encadré de Recherche (TER) prévu par le parcours M1 MFA. Ce fut d'ailleurs la première fois que j'utilisai le logiciel  $\text{\LaTeX}$  pour rédiger un devoir.

J'ai réalisé ce travail dans le domaine des Probabilités, sous la direction de l'enseignant-chercheur Jean-François LE GALL, sur les « Marches aléatoires et arbres de Galton-Watson », et j'ai adoré étudier ce thème. Il constitue la deuxième annexe de ce mémoire.

## Chapitre 3

# Troisième année

En 2013/2014, j'ai effectué une année d'interruption dans le parcours MFA pour préparer l'agrégation de Mathématiques, en M2 Formation des Professeurs Agrégés (M2 FPA) à Orsay.

Après avoir réussi le concours de l'agrégation, j'ai fait le choix de demander un report de stage pour terminer ma formation au sein du magistère. En 2014/2015, je me suis donc inscrit en M2 MFA, spécialité Probabilités et Statistiques (l'un des M2 recherche proposés à Orsay), ainsi qu'en troisième année de magistère.

### 3.1 UE supplémentaire de M2

Comme les magistériens se trouvaient dans divers cursus, il n'y eut pas de cours spécifique cette dernière année, et il fut simplement demandé de suivre un enseignement supplémentaire de niveau M2. Cet enseignement pouvait cependant ne pas appartenir au même M2 recherche, ni même se trouver à l'université d'Orsay... Il s'agissait donc d'un enseignement d'ouverture, afin d'élargir une fois de plus nos connaissances.

Ayant personnellement axé fortement mon cursus vers les Probabilités cette année de M2, j'ai choisi comme UE supplémentaire un enseignement de Statistiques de mon cursus. Celui-ci fut dispensé par l'enseignant-chercheur Pascal MASSART, et intitulé *Choix (ou sélection) de modèles*. C'était en réalité la suite du cours probabiliste *Concentration de la mesure*, que j'ai suivi pour la validation de mon M2.

Dans ce cours de Statistiques, nous avons donc étudié le problème important de sélection de modèles, et abordé également une introduction à l'apprentissage statistique. J'ai suivi tous les cours de cette UE, mais ne me suis pas présenté à l'examen, qui m'était facultatif.

## 3.2 Mémoire de M2

Au second semestre, j'ai réalisé mon mémoire de recherche dans le domaine des Probabilités, sous la direction de l'enseignant-chercheur Raphaël CERF. Mon mémoire s'intitulait alors « Évaluation Route Royale et sélection par rang pour les algorithmes génétiques simples, sans croisements », et je le poursuivis jusqu'au mois de septembre 2015, dans l'espoir d'aboutir un peu plus loin dans mes recherches.

Ce mémoire de M2 se trouve dans la troisième annexe de ce mémoire de magistère, et je ne parlerai pas du contenu mathématique ici, car c'est l'objet de la partie suivante. J'explique en revanche en quelques lignes le contexte et l'étendue du travail que j'ai réalisé.

J'ai eu la chance d'avoir le choix entre deux types de mémoires : l'un consistant à étudier et traduire un article récent de Percolation, comme j'en avais initialement le projet, et l'autre consistant à réaliser de vraies recherches sur les algorithmes génétiques. Le premier pouvait se réaliser en quelques mois, mais le deuxième était bien entendu beaucoup plus incertain, et plus excitant en un sens ! C'est donc lui que j'ai choisi...

Les algorithmes génétiques sont un domaine de recherche sur lequel Raphaël CERF a déjà écrit plusieurs articles, et mon rôle fut tout d'abord d'étudier son dernier article<sup>1</sup> qui démontre, en utilisant des méthodes de grandes déviations, l'existence de paramètres pour lesquels des algorithmes génétiques simples très généraux réalisent leur rôle d'optimisation.

Durant les premières semaines de mon mémoire, j'ai donc étudié la grande démonstration de son article, en comblant également mon retard en grandes déviations, car c'était le seul enseignement probabiliste que je n'avais pas suivi au premier semestre.

Ces théorèmes d'existence, qui se révèlent précieux pour les mathématiciens, ne le sont en fait pas autant pour les praticiens. L'espoir de Raphaël CERF, qui fut alors le mien, était que sa démonstration puisse être adaptée pour des algorithmes très particuliers afin de fournir des paramètres concrets, et donc utilisables en simulation.

Après l'étude de l'article, mon travail consistait donc à adapter la démonstration aux cas particuliers des algorithmes génétiques simples avec une fonction d'évaluation Route Royale pour la sélection par rang, et sans croisements, ce qui est a priori un des cas les plus simples à considérer dans ce domaine. Or, en fin de compte, la tâche s'avéra trop longue et compliquée pour moi, et je n'ai pu réinvestir qu'une petite partie de la démonstration dans mon propre mémoire avec le temps imparti, ce qui explique pourquoi il se termine si abruptement.

Si je n'ai pas abouti personnellement à un résultat exploitable, j'ai pu apprendre beaucoup de choses grâce à ce mémoire de M2, et notamment sur les

---

1. De février 2015 ! (Non encore publié à ce moment...), intitulé *The quasispecies regime for the simple genetic algorithm with ranking selection*.



algorithmes génétiques qui peuvent facilement être expliqués même au lycée. De plus, comme je n'ai pas continué moi-même en thèse à la fin de mon année de M2, j'espère que mon mémoire puisse faciliter le travail de la personne qui souhaitera poursuivre dans la même voie.

### 3.3 Présentation d'un domaine de recherche

Enfin, la formation du magistère demandait à la fin de cette troisième année une soutenance, qui eut lieu pendant les vacances de la Toussaint 2015. Il s'agissait de présenter un domaine de recherche actuel en mathématiques, notamment en rapport avec notre propre sujet de thèse. Il fallait en faire une présentation orale et écrite de quelques pages.

Comme j'ai fait le choix d'entrer dans l'éducation nationale cette même année, je devais donc présenter le domaine de recherche lié à mon mémoire de M2, à savoir les algorithmes génétiques simples. Ma rédaction écrite de présentation constitue donc une partie de la dernière annexe de ce mémoire de magistère, que j'ai choisi de placer juste avant mon mémoire de M2 pour cette raison.

### 3.4 Mémoire de magistère

Pour conclure, la rédaction du présent mémoire est la dernière exigence du magistère, hormis la validation du cursus standard MFA. Ce mémoire avait pour but, je pense, de rendre compte de l'ensemble de la formation suivie individuellement, et de regrouper toutes les rédactions réalisées pendant cette période par chacun. J'espère donc avoir rempli ces conditions.

Je remercie pour finir le responsable du magistère de mathématiques, Frédéric PAULIN, pour avoir proposé une telle formation, à la fois exigeante et intéressante, ainsi que tous les enseignants étant intervenus dans celle-ci.

Annexe A

Rapport du stage en lycée

## **Rapport de stage - Magistère**

Stagiaire : Jimmy PAROT,  
L3 MFA - première année de Magistère.

Établissement d'origine : Université Paris-Sud 11 – UFR Sciences,  
15 rue Georges Clemenceau,  
91440 Orsay Cedex.

Responsable de stage : Charles FRANCES,  
Département des Mathématiques d'Orsay.

Sujet :

### **L'enseignement des Mathématiques dans le secondaire à travers trois professeurs de lycée.**

Établissement d'accueil : Cité Scolaire André CHAMSON,  
1 Boulevard Pasteur,  
30120 Le Vigan.

Responsable dans l'établissement d'accueil : Michel BARTHEZ,  
professeur de Mathématiques.

Période du stage : Du 21 mai 2012 au 8 juin 2012,  
soit 3 semaines, ou 15 jours ouvrés.

Fait à Orsay, septembre 2012.

## Sommaire

<u>I – Introduction</u> .....	p.3
1 – Motivations	
2 – Présentation du stage	
<u>II – Mon stage au quotidien</u> .....	p.4
1 – Mr BOURDIN	
2 – Mr GALTIER	
3 – Mr BARTHEZ	
<u>III – Bilan classe par classe</u> .....	p.7
1 – Les Scientifiques et les "spéMath"	
2 – Les Secondes	
3 – Les autres filières	
<u>IV – Autres expériences</u> .....	p.9
1 – Préparation d'un cours de Seconde avec Mr GALTIER	
2 – Excursion au collège avec Mr SCHNEIDER	
a – Observations	
b – Compte-rendu	
<u>V – Conclusion et Apport</u> .....	p.11
<u>VI – Annexes : Deux cours de Seconde sur l'échantillonnage</u> .....	p.12
1 – Cours préparé personnellement à la maison	
2 – Cours préparé par Mr GALTIER pour sa classe	
<u>VII – Remerciements</u> .....	p.17

## **I – Introduction**

### 1 – Motivations

Les différentes applications des Mathématiques dans la vie courante sont naturellement très intéressantes mais un stage dans une entreprise quelconque présentait peu d'intérêts pour moi, au sens où celui-ci ne me permettrait de découvrir au mieux qu'une seule de ces applications, et aucune ne m'intéresse particulièrement plus que les autres.

Il y avait également la possibilité d'un stage dans un laboratoire de recherche et l'informatique m'aurait beaucoup intéressée, et en particulier la programmation, mais j'ai pensé qu'un stage dans l'enseignement me serait plus profitable puisque j'ai pour ambition d'obtenir l'Agrégation et d'enseigner ensuite.

De plus, la recherche et l'enseignement à l'université m'intéressent moins. J'ai toujours préféré m'intéresser au Mathématiques dans leur généralité et c'est pourquoi je préférerais sans doutes enseigner en classe préparatoire, si possible, et sinon dans le secondaire, au niveau lycée. Voilà pourquoi j'ai choisi d'effectuer ce stage dans mon ancien lycée-collège.

### 2 – Présentation du stage

Étant donnée que le stage devait durer trois semaines, j'ai pu m'arranger pour suivre trois professeurs, ce qui m'a permis de pouvoir comparer différentes méthodes d'enseignement et de collecter plusieurs opinions de choix quant à l'enseignement des Mathématiques au lycée.

J'ai pu consacrer mon temps à l'observation, mais aussi à la pratique avec quelques participations occasionnelles en classe, et surtout avec la préparation d'un cours de seconde sur l'échantillonnage, que j'ai pu confronter à celui préparé par un des professeurs, afin de comparer les deux.

Toujours dans l'objectif de rendre mon étude la plus complète possible, bien qu'irréremédiablement limitée au même lycée régional, j'ai assisté à deux cours de collège, en 6ième et en 5ième, afin de me faire également une idée des difficultés rencontrées par les professeurs à ce niveau.

Ainsi, je vais pouvoir exposer ce qu'a été mon quotidien auprès de ces différents professeurs, durant lequel j'ai pu découvrir leurs propres méthodes d'enseignement, et observer également le comportement de chaque classe vis-à-vis des Mathématiques.

## II – Mon stage au quotidien

Dans cette partie, je vais présenter ce qu'a été mon quotidien au cours de ce stage. Je donnerai aussi les remarques que j'ai pu récolter auprès des personnes qui m'ont entouré, et les différents constats que j'ai pu faire sur les professeurs et leurs méthodes, les élèves et leur comportement, et le lycée et son organisation.

J'ai choisi pour cela de présenter les choses dans un ordre chronologique, d'une semaine à l'autre.

### 1 – Mr BOURDIN

Mon stage débute avec Mr BOURDIN, professeur de lycée que j'ai moi-même eu comme professeur en classe de Première. C'est donc avec lui que j'effectue ma première semaine de stage. Je m'adapte naturellement à son emploi-du-temps, et assiste à ses propres cours jour après jour.

*Lundi 21/05/2012*

T STG : La classe des Terminales STG est agitée tout au long du cours, mais Mr BOURDIN préfère opter pour une politique pacifique. Le cours avance plutôt lentement mais dans une ambiance qui semble bien convenir aux élèves.

2nde5 : C'est une séance de TD, et j'en profite pour participer : je passe dans les rang à l'instar du professeur pour aider les élèves dans leur recherche. Je constate qu'en Seconde, les Mathématiques ne suscitent pas le même intérêt chez tout le monde, et les niveaux sont très inégales, c'est pourquoi le professeur préfère favoriser les séances de TD et ne fait que très peu de cours, ou bien des TD-cours.

*Mardi 22/05/2012*

T ES spéMath : C'est un petit groupe d'élèves beaucoup plus intéressés. Le professeur fait travailler sur des sujets de bac, privilégiant la recherche et s'arrêtant sur la rédaction, le cours se passe bien.

T ES : En classe entière, il y a beaucoup de dissipation, le cours est chargé d'exemples pour faire participer la classe au maximum et capter l'attention.

1èreS2 : Comme on s'y attend pour une classe Scientifique, elle est beaucoup plus calme et un cours magistral se déroule aisément. Il termine par quelques activités, pour apprendre à utiliser la calculatrice.

*Jeudi 24/05/2012*

1èreS2 : On continue comme lors du dernier cours, mais je note qu'il y a beaucoup d'utilisations de la calculatrice, pour les statistiques, mais aussi pour les primitives, dérivées... Les élèves semblent s'en passer de moins en moins.

*Vendredi 25/05/2012*

Les classes de T ES spéMath, T ES, T STG, 2nde5 et T ES de nouveau se succéderont sans réels changements par rapport au reste de la semaine. Sauf pour la classe de T STG où certains élèves ont décidé de participer et ont montré beaucoup de bonnes volontés. C'est une classe très irrégulière au niveau du comportement.

## 2 – Mr GALTIER

J'entame ma deuxième semaine avec Mr GALTIER, également professeur de lycée dans cet établissement, que j'ai peu connu par le passé mais qui m'aidera beaucoup aussi pour mon étude durant ce stage.

*Mardi 29/05/2012*

2nde3 : L'ambiance est visiblement plus stricte, et les exercices maisons sont relevés, puis notés ! Mais la classe est à peu près calme, il y a de l'attention et de la participation, de la part des meilleurs peut-être.

TS1 : Même ambiance assez décontractée mais stricte si nécessaire. Beaucoup d'insistance sur la rédaction lors des corrections. Mr GALTIER préfère donner des exercices à la maison plutôt qu'en classe pour gagner du temps.

T S spéMath : TD avec un très bon groupe d'élèves motivés. Je participe ici aussi en passant au tableau pour corriger les exercices.

Mr GALTIER m'apprend par ailleurs que les programmes sont généralement difficiles à boucler. Mais le professeur d'expérience à sa propre stratégie : cibler tous les cours importants dès le début d'année, afin de prévenir les nombreuses pertes de temps telles que les grèves ou les sorties en classe entière dans les autres matières.

*Jeudi 31/05/2012*

Je profite d'un blanc dans l'emploi-du-temps pour passer la matinée au collège avec Mr SCHNEIDER. (voir IV.2.)

L'après-midi se déroule avec les 1èreS1, les 2nde3 et les 1èreS1 de nouveau, et la particularité porte sur les 2nde3. Plusieurs élèves, qui ont l'intention de poursuivre en 1èreL, sont totalement dépourvus de bonnes volontés. Et Mr GALTIER à donc pris des mesures pour la fin d'année :

Les 2 demi-groupes ont été modifiés de sorte que les élèves qui ont l'intention de continuer à travailler cette matière se trouvent dans l'un, et les futurs Littéraires se trouvent dans l'autre. De cette façon, Mr GALTIER prend le premier demi-groupe pendant 2h et le professeur de français prend l'autre en parallèle, et cela à la place de 1h l'un puis 1h l'autre.

*Vendredi 01/06/2012*

TS1 : Il s'agit d'une des fois où beaucoup d'élèves sont absents pour une sortie, le professeur n'a plus qu'à meubler avec des exercices de révisions.

Mr GALTIER me propose aussi un peu de devoir maison pour me faire une meilleure idée du métier, et me propose donc de préparer un cours pour la semaine prochaine, et de comparer avec celui qu'il aura lui-même préparé. (voir IV.1.)

### 3 – Mr BARTHEZ

Je fais donc ma troisième et dernière semaine avec Mr BARTHEZ, professeur de lycée que j'ai également eu moi-même, en Seconde et en Terminale.

*Lundi 04/06/2012*

TS2 : L'heure est aux révisions pour le bac, et comme souvent dans ces cas là, je passe moi aussi dans les rangs pour aider si besoin.

2nde4 : De même, je passe pour aider dans les exercices. Mais le problème avec les futurs Littéraires est le même.

2nde1 : Même remarque, mais quelques élèves sont très demandeurs et certains futurs Littéraires aussi jouent le jeu, malgré l'absence d'évaluations.

Mr BARTHEZ n'a pas pris de dispositions particulières avec les Secondes et a préféré garder les demi-groupes comme prévu, mais le problème persiste et nuit au bon déroulement des cours. Il n'y a plus de Mathématiques en 1èreL (sauf option) mais Mr BARTHEZ tente de lancer l'idée vers les hautes instances d'un contrôle-continue pour les futurs Littéraires, comme solution à ce problème.

*Mardi 05/06/2012*

TS2 : Pour la première heure de la journée, le professeur me propose de tester un DS de 2h que les élèves devons faire l'après-midi même. Comme il faut aussi créer des sujets de devoirs quand on enseigne, cela aura donc aussi permis de compléter ma vision du métier.

T L spéMath : Seulement 4 élèves, mais ce n'est pas très étonnant pour une filière Littéraire, et cela offre des facilités de travail appréciables.

*Mercredi 06/06/2012*

TS2 : C'est la fin de l'année, quelques révisions tranquilles.

2nde1 : 12 absents ! Et à peine 7 élèves resteront pour la dernière heure d'Aide Personnalisée.

*Jeudi 07/06/12*

Mon stage se termine avec l'année scolaire. Il n'y a pas grand chose à ajouter pour cette semaine avec Mr BARTHEZ, mais c'est ce jour ci que j'ai assister au cours de Mr GALTIER sur l'échantillonnage avec les 2nde3. (voir IV.1.)



### III – Bilan classe par classe

#### 1 - Les Scientifiques et les "spéMath"

Il y a peu à dire concernant les classes Scientifiques et les options spéMath, si ce n'est qu'elles sont souvent synonyme de bonnes classes. La principale raison est que la plupart des élèves présents dans ces classes ont un goût prononcé pour les Mathématiques, ou les sciences en générale.

Par contre, il est vrai qu'on peut aussi y trouver des élèves qui ne sont pas particulièrement intéressés, et qui sont là simplement parce que la filière Scientifique est la plus générale ou celle qui offre le plus de débouchés. Mais cela ne pose aucuns problèmes parce que les Mathématiques est la matière qui possède le plus fort coefficient, donc tous les élèves de la filière travaillent de leur mieux.

Quant aux spéMath des autres filières, il est évident que leur intérêt pour la matière n'est pas à prouver, suite à leur choix.

#### 2 – Les Secondes

Ce que l'on peut retenir des classes de Secondes, en revanche, c'est que ce sont en général des classes bien plus délicates.

En effet, comme il s'agit de la seule année de lycée encore générale, le goût des élèves pour la matière est beaucoup plus mitigé. Et le problème majeur vient du fait que les élèves qui souhaitent poursuivre en 1èreL n'auront absolument plus de Mathématiques l'année suivante. Et il n'y a évidemment pas d'examens en fin de Seconde, donc cette matière n'a souvent aucune importance pour ces élèves.

C'est un problème auquel tous les professeur que j'ai pu suivre ont dû faire face, puisque l'absence de motivation de ces quelques élèves suffit à perturber et ralentir tout le reste de la classe.

De ce fait, j'ai pu récolter les différentes réactions et propositions des enseignants qui m'ont encadré, comme je l'ai également signalé en énonçant le déroulement de mon stage dans la partie précédente :

- Il y a par exemple les dispositions prises par Mr GALTIER, qui consistent à trier les élèves en formant 2 demi-groupes, avec dans l'un les élèves qui n'ont aucun intérêt pour la matière. Et le professeur, n'ayant pas l'intention de s'encombrer avec ces élèves là bien longtemps, a pu trouver un arrangement avec celui de Français : chacun prendra le groupe qui convient 2h durant, au lieu de 1h l'un puis 1h l'autre.
- Il y a aussi la proposition faite par Mr BARTHEZ, qui serait d'instaurer un contrôle-continue en Seconde pour les futurs Littéraires, de sorte que ces élèves soient plus motivés et travaillent cette matière.

Finalement, on peut conclure que ce problème des Secondes est un obstacle majeur dans l'enseignement des Mathématiques au lycée, que certains professeurs comme Mr GALTIER cherchent à minimiser, mais ce genre de dispositions n'est possible qu'avec une entente entre les deux professeurs.

Pour ma part, je pense vraiment que l'idée d'un contrôle-continue, moins pénalisant qu'un examen, présente un grand intérêt. Cela permettrait aux futurs Littéraires de pouvoir gagner quelques points, ou d'essayer de ne pas trop en perdre, en travaillant un peu cette matière. Le principale objectif étant qu'ils fournissent un minimum de travail, et prêtent un peu d'attention à ce cours.

### 3 – Les autres filières

Pour les filières non-Scientifique telles que STG ou ES, les choses sont encore différentes. Puisque les Mathématiques ne font pas partie des matières principales, on ne peut pas en demander autant aux élèves qu'en S, et la difficulté est finalement de trouver un bon rythme de travail, qui soit ni trop exigeant, ni trop lent.

Il peut y avoir plusieurs façons d'aborder les choses.

On peut par exemple imposer un rythme aux élèves, et dérouler son cours de la façon dont on l'entend. Mais je pense que ça serait aussi prendre le risque de décrocher bon nombre d'élèves, qui ne pourraient ou ne voudraient pas suivre un tel rythme, et renonceraient rapidement.

Une autre façon d'aborder les choses, comme Mr BOURDIN l'a montré, serait de s'adapter, dans une certaine mesure, au rythme des élèves. Cela présente plusieurs inconvénients et ne permet peut-être pas d'aller aussi vite qu'on le voudrait, mais de cette façon on peut réussir à entraîner la majorité des élèves dans une bonne ambiance, et constater qu'ils acceptent plus volontiers de participer à la recherche de quelques exercices tous ensemble.

Il n'est pas sûr qu'une des méthodes vaut mieux que l'autre, et je crois finalement que ça n'est qu'une question de choix, où chaque professeur fait ce qu'il pense être le mieux, pour lui comme pour la classe.

## IV – Autres expériences

### 1 – Préparation d'un cours de Seconde avec Mr GALTIER

Mr GALTIER, au cours de la semaine que j'ai passé dans sa classe, a eu l'idée de me faire préparer un cours niveau Seconde, qu'il devait lui-même préparer pour la semaine d'après.

Il s'agit d'un petit cours de Statistiques, sur l'échantillonnage, et cela s'avérait être une très bonne opportunité d'expérimenter le métier d'enseignant. Puis comme le professeur allait lui aussi préparer ce cours pour sa classe, j'avais la possibilité de comparer mon cours avec le sien, et nous pouvions ensuite discuter de ce qui est bien ou moins bien dans mon travail.

Pour cela, j'ai commencé par sélectionner parmi les 9 ou 10 livres disponibles à la bibliothèque ceux qui me paraissent les plus intéressants quant à la façon d'aborder les notions. Et avec les 4 ou 5 livres que j'ai retenus, j'ai piocher de quoi construire mon propre cours, empruntant quelques exemples et remarques dans les uns, et cherchant des idées d'activités dans les autres.

Le cours que j'ai préparé est disponible à la fin de se rapport, est constitue les deux premières feuilles des **Annexes**.

→ voir VI.1., correspondant aux pages 13 et 14.

Et celui de Mr GALTIER est également disponible à la suite, à titre de comparaison, il s'agit des deux dernières feuilles des **Annexes**.

→ voir VI.2., correspondant aux pages 15 et 16.

Le jeudi 7 juin, je suis donc retourné dans la classe de 2nde3 pour assister au cours sur l'échantillonnage, et présenter moi-même mon travail au professeur. La première chose que l'on peut constater est qu'ils sont visiblement très différents, sur la manière d'aborder le chapitre et d'énoncer les notions essentielles.

J'avais dans l'idée de présenter les notions et les résultats de manière formelle, avec toutefois un ou deux exemples simples, puis de cibler les principaux savoirs-faire avec deux activités principales.

Finalement, Mr GALTIER me fait remarquer que mon cours est trop structuré pour une classe de Seconde, où le niveau est beaucoup plus faible que je ne le pensai. J'apprends donc qu'il vaut sans doutes mieux présenter les choses directement sous forme d'activités, en glissant les principaux résultats au milieu, et les faire travailler sur des exemples et beaucoup d'exercices.

Conclusion, on peut retenir que ce fut une très bonne expérience pour moi car elle a su me rappeler l'évidence, à savoir qu'il ne suffit pas qu'un cours soit bien rédigé pour être compris par une classe...

## 2 – Excursion au collège avec Mr SCHNEIDER

### a – Observations

Le Mercredi 31 mai, je profite d'une matinée de libre pour assister à deux cours de Mr SCHNEIDER, professeur de collège dans l'établissement, avec les classes de 6ème6 et 5ème1.

Je remarque tout de suite une profonde différence avec le lycée, et à commencé par la disposition de la salle. Les tables sont regroupées pour former plusieurs groupes de 4, 5 ou 6 élèves en cercle.

A ce niveau, il n'y a pas de cours à proprement parlé, la stratégie du professeur est de faire travailler les élèves en groupe sur de petits exercices pour qu'ils puissent s'entre-aider, et de faire participer tout le monde au maximum pour la correction. Il faut sans cesse interroger la classe ou envoyer des élèves au tableau pour capter l'attention des autres, car je remarque que tous les élèves se dissipent instantanément dans le cas contraire.

Enfin, la particularité du professeur est de ne pas faire de contrôle à ce niveau, il n'y a pas d'évaluations notées jusqu'en 4ème, et Mr SCHNEIDER offre à tous ses élèves la possibilité de rattraper à la maison un devoir raté à n'importe quel moment de l'année.

### b – Compte-rendu

Au collège, il est évident que le principale problème vient de dissipation des élèves, qui sont simplement incapables de se concentrer si on ne les y force pas un peu. Du coup les méthodes d'enseignement des Mathématiques sont très différentes du lycée.

En fait, le travail d'enseignant au collège n'est donc pas du tout le même qu'au lycée, et s'apparenterait presque à celui d'un animateur pour les petites classes, où le but est plus de capter l'attention et de faire participer les élèves.

Par ailleurs, je sais que le fait de ne pas noter les élèves jusqu'en 4ème est peu approuvé par les collègues de lycée de Mr SCHNEIDER, et aussi par les inspecteurs, car cela livre les élèves à eux-même et beaucoup ne travail pas suffisamment, puis accumule des lacunes qu'il sera difficile de combler ensuite.

En revanche, je considère que l'idée de regrouper les élèves est intéressante car même si cela semble inciter au bavardage, comme on pourrait le penser, j'ai pu constater que cela permettait surtout aux bons élèves d'aider tous le reste de la table, et donc même les élèves en difficulté finissent souvent par réussir, ce qui ne serait pas le cas deux par deux.

## V – Conclusion et Apport

Le principale objectif de mon stage était d'étudier les différentes méthodes d'enseignement des Mathématiques au lycée, et éventuellement au collège.

Trois semaines passées à arpenter toutes les classes de lycée et côtoyer les professeurs m'ont permis de réaliser à quel point ces méthodes peuvent différer, d'une filière à l'autre, puis selon les enseignants et leurs choix.

Notamment, j'ai pu discuter avec les trois professeurs qui m'ont encadré sur les différentes difficultés auxquelles ils sont confrontés chaque année, et que j'ai moi-même pu constater dans les classes. Parmi celles-là, nous retiendrons essentiellement le manque de pression sur les élèves de Seconde qui poursuivront en L, et qui se désintéressent donc totalement de cette matière, ce qui ralentit considérablement la progression de toute la classe.

Par ailleurs, j'ai également pu me rendre compte des exigences variables qu'un enseignants peut se permettre d'avoir avec ses différentes classes. En particulier, il est par exemple normal de ne pas en demander autant aux classes des filières non-Scientifiques, qu'à une classe Scientifique. Et la façon d'aborder les choses change aussi, avec entre autres plus d'exemples et d'exercices.

Ce stage m'aura donc permis d'étudier en détail chacune des classes de lycée, mais aussi d'avoir une idée de ce qui se passe au collège, grâce à une petite excursion lors d'une matinée. J'ai été surpris de constater que les choses sont encore très différentes à ce niveau, où la principale difficulté consiste à capter l'attention de ces élèves en jeune âge, plus qu'autre chose.

Enfin, j'ai également approché au plus près le métier d'enseignant auquel je me destine grâce à la préparation d'un cours qu'il m'a fallu réaliser du début à la fin, à la maison. Et l'expérience m'a été largement profitable puisque j'ai pu exposer mon travail à l'un des enseignants qui m'encadrerait, et profiter de ses conseils et remarques tout en comparant avec son propre cours, qu'il avait dû lui aussi préparer pour sa classe.

Pour conclure, je suis satisfait d'avoir pu réaliser ce stage en lycée, et d'avoir pu réaliser cette étude dans de pareilles conditions, car cela m'a permis d'élargir ma conception du métier d'enseignant, et d'avoir un meilleur aperçu de ce qui m'attend si je persévère dans cette voie, comme j'en ai l'intention.

## **VI – Annexes : Deux cours de Seconde sur l'échantillonnage**

Suite à la proposition de Mr GALTIER, j'ai préparé un cours de Seconde sur l'échantillonnage, afin de mieux me rendre compte de ce qu'est le métier d'enseignant au lycée.

Parallèlement, Mr GALTIER à lui aussi préparé le même cours pour sa classe, ce qui m'a permis de pouvoir comparer mon travail avec le sien, et d'apprendre sur les attentes d'une classe de Seconde.

Je fourni donc ces deux cours dans ces **Annexes**, à titre de comparaison, et d'illustration pour la partie IV.1. de mon rapport.

### 1 – Cours préparé personnellement à la maison

Voir les deux premiers documents, correspondant aux pages 13 et 14.

### 2 – Cours préparé par Mr GALTIER pour sa classe

Voir les deux derniers documents, correspondants aux pages 15 et 16.

## VII – Remerciements

Il me faut maintenant remercier plusieurs personnes qui ont joué un rôle essentiel dans le bon déroulement de ce stage, ou sans qui il n'aurait tout simplement pas pu avoir lieu.

Je remercie donc Mr FRANCES, mon responsable de stage, qui à finalement accepter de donner son accord pour ce stage en lycée malgré quelques réticences initiales, puisque conventionnellement ce genre de stage n'entre pas dans les critères souhaités, et auquel on préfère les stage en entreprise ou dans un laboratoire de recherche non-Mathématique.

Je remercie également Mme TOUZALIN, proviseur du lycée-collège qui m'a accueilli, d'avoir accepté dans son établissement un stagiaire de niveau seulement L3, et d'avoir fait le nécessaire d'un point de vue administratif pour que mon stage, qui devait débiter peu de temps après, puisse se dérouler en temps et en heure.

Puis je remercie Mr BOURDIN, Mr GALTIER et Mr BARTHEZ, les trois professeurs qui ont accepté de m'encadrer dans mon stage à tour de rôle, pour m'avoir apporté leur aide précieuse dans cette étude, puisqu'ils n'ont pas hésiter à me faire part de leur quotidien, leurs remarques et leurs conseils, et m'ont finalement permis de découvrir le métier d'enseignant de la meilleur manière qui soit.

Enfin, je remercie Mr SCHNEIDER, pour m'avoir lui aussi accueilli dans sa classe, lors d'une matinée, et permis d'entrevoir ce que peut être l'enseignement des Mathématiques au collège, complétant ainsi mon étude sur les différentes méthodes d'enseignements.

Quoique je pense pouvoir également remercier tous les élèves de toutes les classes dans lesquelles je suis passé, pour avoir contribué, bien qu'indirectement, à cette étude.

Annexe B

TER de M1



TER M1 MFA :

Marches aléatoires et arbres de Galton-Watson.

Réalisé par Jimmy Parot,

Encadré par Jean-François Le Gall.

19 juin 2013

Le but de ce TER est d'explorer quelques propriétés sur les arbres de Galton-Watson et de montrer le lien naturel avec les marches aléatoires.

Nous mettrons en évidence dans le cadre qui nous occupe des propriétés évidente sur les arbres de Galton-Watson, nous introduirons les notions d'excursions et de ponts qui peuvent se déduire d'un tel arbre et remarquerons ainsi le lien avec une marche aléatoire sur  $\mathbb{Z}$ .

## Table des matières

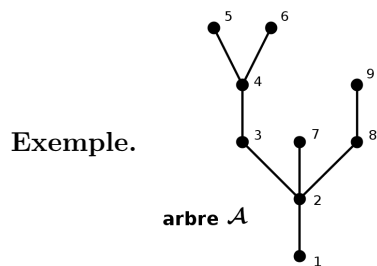
<b>1 Arbres, numérotation et excursion associées</b>	<b>2</b>
<b>2 Liens entre arbres et marches aléatoires</b>	<b>6</b>
2.1 Interprétation à numérotation fixée . . . . .	6
2.2 Arbre généré à partir d'une excursion . . . . .	7
2.3 Arbre généré à partir d'un pont . . . . .	9
2.4 Probabilité de sortie d'un arbre . . . . .	11
<b>3 Exemples d'arbres de Galton-Watson</b>	<b>13</b>
3.1 Le cas d'une distribution binaire . . . . .	13
3.2 Le cas d'une distribution géométrique . . . . .	16

## 1 Arbres, numérotation et excursion associées

On considère d'abord pour cette partie un arbre, que l'on notera  $\mathcal{A}$ , vérifiant les propriétés très générales d'être fini (de taille  $N$ ), dans le plan, possédant une racine et tel que tous les nœuds (dont la racine) sont numérotés de façon décrite juste après.

Par commodité, on interprétera tout arbre comme "arbre généalogique d'une famille" dont tous les individus sont les nœuds, et où la racine est l'unique ancêtre commun.

Avec cette interprétation, on fait l'hypothèse que les individus sont numérotés de 1 à  $N$  de sorte que **tout individu possède un numéro plus grand que son parent**. En particulier la racine portera toujours le numéro 1.



Dans cet exemple,  $N=9$ , et les individus ont été numéroté de gauche à droite comme si, partant de la racine, on avait longé les arêtes du côté gauche pour faire le tour de l'arbre. Il est clair que la numérotation choisie ici vérifie bien la propriété demandée ci-dessus.

**Remarque.** Comme  $\mathcal{A}$  est un arbre plan, on peut choisir d'ordonner les enfants d'un même individu, disons de gauche à droite. On imagine désormais  $\mathcal{A}$  comme un **arbre ordonné** de la sorte.

$\mathcal{A}$  est un **arbre de Galton-Watson** (G-W) si les nombres d'enfants des différents individus sont indépendants et suivent la même loi qu'une variable aléatoire  $X$  dont la loi de probabilité est donnée. Dans la suite, on notera  $X_i$ , pour  $i \in \{1, \dots, N\}$ , le nombre d'enfants de l'individu  $i$ .  $X_1, \dots, X_N$  est donc une suite de variables aléatoires i.i.d., que l'on suppose finie dans ce TER.

**Exemple.** Pour l'arbre  $\mathcal{A}$  ci-dessus, on a :

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
1	3	1	2	0	0	0	1	0

**Proposition 1.1.** Pour tout arbre  $\mathcal{A}$  comme avant, on a toujours :

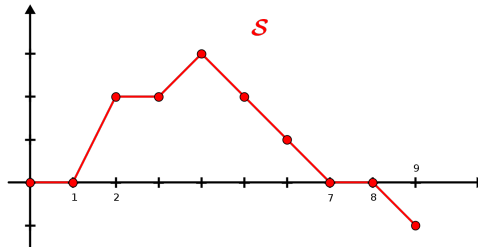
$$X_1 + \dots + X_N = N - 1.$$

**Preuve 1.1.** Par définition, sommer les  $X_i$  revient à compter tous les individus qui sont enfant d'un autre, et exactement une fois car aucun individu ne peut avoir 2 parents. On compte alors tout le monde sauf la racine, qui est la seule à n'avoir aucun parent, donc  $N - 1$  individus.

À tout arbre  $\mathcal{A}$  on peut associer un **chemin**, noté  $\mathcal{S} = (S_0, \dots, S_N)$ , défini par :

$$\begin{cases} S_0 = 0 \\ \forall i \in \{1, \dots, N\}, S_i = S_{i-1} + X_i - 1. \end{cases}$$

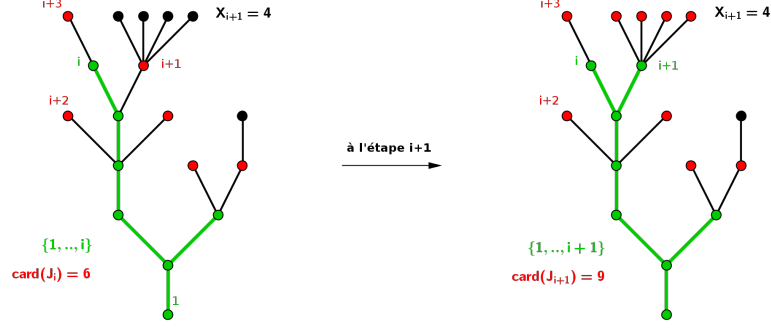
**Exemple.** Pour l'arbre  $\mathcal{A}$  comme avant, on a  $\mathcal{S} = (0, 0, 2, 2, 3, 2, 1, 0, 0, -1)$ , d'où la représentation graphique :



**Lemme 1.2.** Le chemin  $\mathcal{S}$  associé à l'arbre  $\mathcal{A}$  comme avant vérifie :  $\forall i \in \{1, \dots, N\}$ ,  $S_i = -1 + \text{card}\{j \in \{i+1, \dots, N\} / \text{le parent de l'individu } j \text{ à un numéro dans } \{1, \dots, i\}\}$ .

**Preuve 1.2.** Pour  $i \in \{1, \dots, N\}$ , notons  $J_i = \{j \in \{i+1, \dots, N\} / \text{le parent de l'individu } j \text{ à un numéro dans } \{1, \dots, i\}\}$ . Démontrons par récurrence sur  $i \in \{1, \dots, N\}$  que  $S_i = \text{card}(J_i) - 1$ .

- Pour  $i = 1$ ,  
 $J_1 = \{j \in \{2, \dots, N\} / \text{le parent de l'individu } j \text{ est la racine}\}$   
 $= \{\text{enfants de la racine}\}.$   
Donc  $\text{card}(J_1) = X_1 = S_1 + 1$  car  $S_0 = 0$ .
- Supposons que  $S_i = \text{card}(J_i) - 1$  jusqu'au rang  $i \in \{1, \dots, N - 1\}$ . Comme tous les individus ont un numéro plus grand que leur parent, il est clair que le sous-arbre engendré par  $\{1, \dots, i\}$  est connexe, comme sur la figure :



De même, l'individu  $i + 1$  a un parent dans  $\{1, \dots, i\}$ , et donc  $i + 1 \in J_i$ , mais  $i + 1 \notin J_{i+1}$  par définition de  $J_{i+1}$ . Par ailleurs, les enfants de  $i + 1$  ne sont pas dans  $J_i$ , puisque leur parent n'est pas dans  $\{1, \dots, i\}$ , mais ils sont dans  $J_{i+1}$ . Donc  $J_{i+1}$  contient les individus déjà présent dans  $J_i$ , plus les enfants de  $i + 1$ , mais plus  $i + 1$ . D'où on déduit :

$$\begin{aligned}
\text{card}(J_{i+1}) &= \text{card}(J_i) + X_{i+1} - 1 \\
&= (S_i + 1) + X_{i+1} - 1 \text{ par hypothèse de récurrence,} \\
&= S_i + X_{i+1} \\
&= S_{i+1} + 1.
\end{aligned}$$

**Proposition 1.3.** Le chemin  $\mathcal{S}$  défini comme avant vérifie alors :

$$\begin{cases} S_0 = 0 \\ \forall i \in \{1, \dots, N - 1\}, S_i \geq 0 \\ S_N = -1. \end{cases}$$

**Remarque.** Un tel chemin est appelé une **excursion**, et s'apparente à une marche aléatoire sur  $\mathbb{Z}$  qui part de 0 et reste positive jusqu'au moment d'arriver en  $-1$ .

**Preuve 1.3.** Avec les notations de la preuve précédente, on a :

- $S_0 = 0$  par définition de  $\mathcal{S}$ .
- Comme chaque individu a un numéro plus grand que son parent,  $i + 1$  a un parent dans  $\{1, \dots, i\}$ , donc  $\forall i \in \{1, \dots, N - 1\}$   $\text{card}(J_i) \geq 1$ , d'où :  $\forall i \in \{1, \dots, N - 1\}$   $S_i = \text{card}(J_i) - 1 \geq 0$ .
- Par définition,  $J_N = \{j \in \emptyset / \dots\} = \emptyset$ , donc  $S_N = \text{card}(J_N) - 1 = -1$ .

**Proposition 1.4.**  $\mathcal{S}$  est une excursion  $\iff \forall i \in \{1, \dots, N - 1\} X_1 + \dots + X_i \geq i$ .

**Preuve 1.4.**

- On suppose que  $\mathcal{S}$  est une excursion. Alors  $\forall i \in \{1, \dots, N-1\}$ ,  

$$X_1 + \dots + X_i = (S_1 - S_0 + 1) + (S_2 - S_1 + 1) + \dots + (S_i - S_{i-1} + 1)$$

$$= i + S_i \geq i \text{ car } S_i \geq 0 \text{ pour } i \in \{1, \dots, N-1\}.$$
  - On suppose que  $\forall i \in \{1, \dots, N-1\} X_1 + \dots + X_i \geq i$ .  
 $S_0 = 0$  par définition. De plus,  $\forall i \in \{1, \dots, N\}$ ,  

$$S_i = S_{i-1} + X_i - 1$$

$$= (S_{i-2} + X_{i-1} - 1) + X_i - 1$$

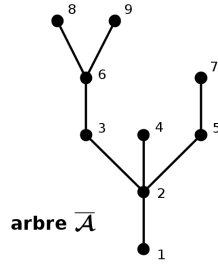
$$= \dots$$

$$= (S_0 + X_1 - 1) + (X_2 - 1) + (X_3 - 1) + \dots + (X_i - 1)$$

$$= X_1 + \dots + X_i - i \begin{cases} \geq 0 & \text{si } i \in \{1, \dots, N-1\} \text{ par hypothèse.} \\ = -1 & \text{si } i = N \text{ par la proposition 1.1.} \end{cases}$$
- Donc  $\mathcal{S}$  est une excursion.

**Remarque.** Tous les résultats vus dans cette première partie ne dépendent pas de la numérotation choisie pourvu que celle-ci vérifie la propriété demandée, à savoir que tout individu possède un numéro plus grand que son parent.

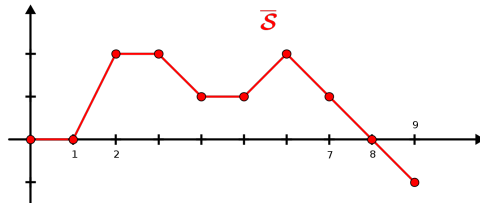
**Autre exemple.** On considère le même arbre  $\mathcal{A}$ , mais que l'on note  $\bar{\mathcal{A}}$  car le choix de numérotation est différent ici, bien qu'on ait toujours la propriété énoncée dans la remarque ci-dessus.



On a dans ce cas :

$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$	$\bar{X}_5$	$\bar{X}_6$	$\bar{X}_7$	$\bar{X}_8$	$\bar{X}_9$
1	3	1	0	1	2	0	0	0

Et  $\bar{\mathcal{S}} = (0, 0, 2, 2, 1, 1, 2, 1, 0, -1)$  est différent de  $\mathcal{S}$ , mais on peut constater qu'il s'agit toujours d'une excursion, dont la représentation est :



## 2 Liens entre arbres et marches aléatoires

Dans cette partie, nous allons fixer une numérotation, et déduire ainsi des liens entre arbres de G-W et marches aléatoires sur  $\mathbb{Z}$ .

On choisit désormais de toujours numéroter les nœuds d'un arbre de la façon suivante :

- On attribue à la racine le numéro 1.
- Si on a déjà attribuer les numéros 1 à  $i$ , on considère l'individu qui a le plus grand numéro et qui a au moins un enfant non numéroté, et on attribue le numéro  $i + 1$  à son premier enfant non numéroté (le plus à gauche).

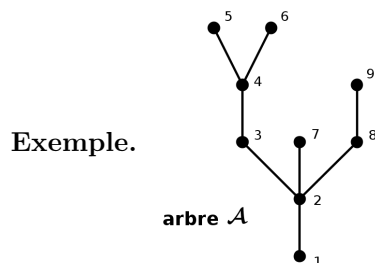
**Remarque.** Cette numérotation correspond au premier exemple donné dans la partie 1, et possède donc les propriétés attendues. En outre, tous les résultats de la partie 1 pourront s'appliquer.

### 2.1 Interprétation à numérotation fixée

Reprenons dans cette sous-partie le premier exemple donné dans la partie 1. Le but ici est d'illustrer le lien entre  $\mathcal{A}$  et l'excursion  $\mathcal{S}$ .

**Remarque.** Puisque la numérotation est fixée, une seule suite  $X_1, \dots, X_N$  peut correspondre à l'arbre  $\mathcal{A}$ , donc une seule excursion  $\mathcal{S}$  également.

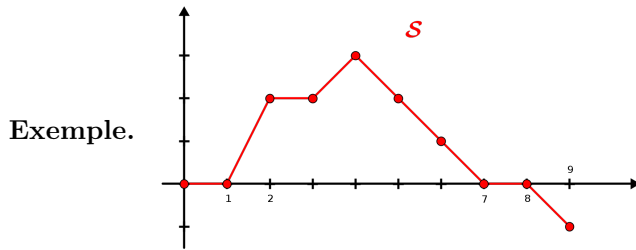
Depuis le début, on interprète  $\mathcal{A}$  comme "arbre généalogique" d'une famille d'individus. La numérotation adoptée suggère alors une chronologie particulière : on choisit ici de s'intéresser d'abord au premier enfant d'un individu, puis à toutes les générations qu'il engendre, avant de s'intéresser au deuxième enfant, etc... Cela revient à penser comme si le premier enfant naît, puis toutes ses descendances, et le deuxième enfant naît une fois que toutes les générations engendrées par le premier sont éteintes, ainsi de suite.



Dans cet exemple, le premier individu à naître après 2 est son premier enfant 3, mais le suivant n'est pas son deuxième enfant 7, qui ne naît qu'au moment où les générations engendrés par 3 ont disparu.

Dans cet ordre d'idée, l'excursion  $\mathcal{S}$  traduit alors l'évolution de la famille au cours du temps. Si l'individu  $i$  possède exactement 1 enfant, alors  $S_i = S_{i-1}$ , donc à la mort du parent  $i$ , la famille n'aura pas évolué en nombre d'individus.  $S_i - S_{i-1}$  représente en fait

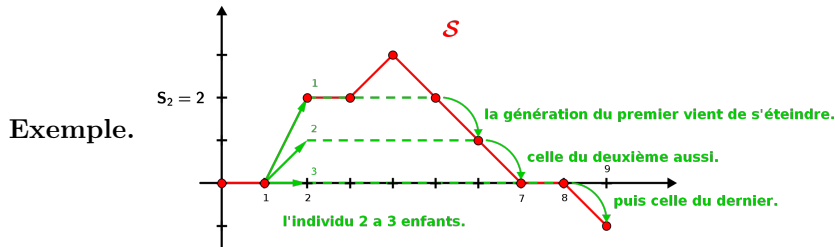
le bénéfice apporté par l'individu  $i$  en "descendance" pour la famille.



Ici, l'individu 3 permet à la famille de se stabiliser en nombre d'individus, en comptant exactement un enfant, tout comme l'individu 8 et la racine 1.

Plus précisément, il est possible de retracer l'histoire de la famille à partir de  $\mathcal{S}$  sans connaître a priori  $\mathcal{A}$ . En effet, il suffit de s'intéresser à la valeur de  $S_i - S_{i-1}$  (qui est  $\geq -1$  puisque  $X_i$  est supposé  $\geq 0$ , comme  $X$ ).

- Si  $S_i - S_{i-1} = -1$ , i.e. si le graphe de  $\mathcal{S}$  décroît, alors l'individu  $i$  n'a pas d'enfants, et il n'engendre pas d'autres générations. En outre, la génération dont il est issu s'éteint.
- En revanche, si  $S_i - S_{i-1} \geq 0$ , i.e. si le graphe de  $\mathcal{S}$  croît ou stationne, l'individu  $i$  a au moins un enfant. Or, dès que le graphe de  $\mathcal{S}$  retombe sous le niveau  $S_i$ , c'est le signe que la génération engendrée par le premier enfant de  $i$  vient de s'éteindre.



Avec ces différentes interprétations, on devine aisément les liens entre un arbre et l'unique excursion qui s'en déduit à numérotation fixée. En réalité, il n'est pas difficile de démontrer un tel lien, ce que la deuxième sous-partie a pour objectif d'exploiter afin de donner un lien entre arbres de G-W et marches aléatoires sur  $\mathbb{Z}$ .

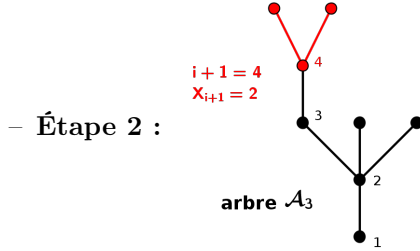
## 2.2 Arbre généré à partir d'une excursion

Établir un lien entre arbres de G-W et marches aléatoires nous permettra de comprendre comment générer un arbre de G-W à partir d'un vecteur de variables aléatoires simulant une marche sur  $\mathbb{Z}$ , avec unicité de la construction en fixant la numérotation.

**Proposition 2.1** À numérotation fixée, il est équivalent de se donner l'arbre (ordonné)  $\mathcal{A}$ , et l'unique excursion  $\mathcal{S}$  associée.

**Preuve 2.1.** On a déjà vu comment construire  $\mathcal{S}$  à partir de  $\mathcal{A}$  (par définition de  $\mathcal{S}$ ). Il reste alors à reconstruire  $\mathcal{A}$  à partir de  $\mathcal{S} = (S_0, \dots, S_N)$ . On exploite le fait que la numérotation est fixée (et que  $\mathcal{A}$  est ordonné) pour comprendre que l'arbre se déduit de la suite  $X_1, \dots, X_N$ , et alors on retrouve  $\mathcal{A}$  de la façon suivante :

- **Étape 1 :** On se donne la racine, que l'on numérote 1, et à laquelle on attache  $X_1 = S_1 - S_0 + 1$  nœuds non numérotés.



Si on a déjà reconstruit une partie  $\mathcal{A}_i$  de l'arbre, où les numéros 1 à  $i$  sont attribués, on considère le nœud le plus à gauche dans  $\mathcal{A}_i$  (en longeant les arêtes du côté gauche, partant de la racine) qui n'a pas encore de numéro. On lui attribue le numéro  $i + 1$ , et on lui attache  $X_{i+1} = S_{i+1} - S_i + 1$  nœuds non numérotés.

Par conséquent, si on voit toute excursion  $\mathcal{S}$  comme une marche aléatoire sur  $\mathbb{Z}$  allant de 0 à -1, dont le vecteur des variations serait  $(S_1 - S_0, \dots, S_N - S_{N-1})$ , on convient qu'à toute marche aléatoire de cette forme correspond un unique arbre de G-W.

En particulier, il est équivalent de se donner un tel vecteur et un arbre de G-W, et donc un vecteur de la forme  $(X_1 - 1, \dots, X_N - 1)$  et un arbre de G-W. Une méthode générale pour construire un arbre de G-W à partir d'une excursion  $\mathcal{S}$  serait alors :

- Déduire les valeurs des variables  $X_1, \dots, X_N$  à partir de  $(S_1 - S_0, \dots, S_N - S_{N-1})$ , donc de  $(X_1 - 1, \dots, X_N - 1)$ .
- Reconstruire  $\mathcal{A}$  à partir de  $X_1, \dots, X_N$  de façon évidente : puisque la numérotation est fixée, reprendre l'algorithme donné dans la preuve 2.1.

Par ailleurs, la probabilité de sortie d'un arbre de G-W par ce procédé dépend finalement de la distribution des variables  $X_1, \dots, X_N$ , dont la loi jointe est donnée dans la proposition suivante.

**Proposition 2.2.** Soit  $n \in \mathbb{N} - \{0\}$ . Soient  $x_1, \dots, x_n \in \mathbb{N}$ , notons  $p_x = \mathbb{P}(X = x)$  pour tout  $x \in \mathbb{N}$ . Alors :

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n, N = n) = \begin{cases} p_{x_1} \dots p_{x_n} & \text{si } \forall i \in \{1, \dots, n-1\}, x_1 + \dots + x_i \geq i \\ & \text{et si } x_1 + \dots + x_n = n-1. \\ 0 & \text{sinon.} \end{cases}$$

**Preuve 2.2.** Si  $\forall i \in \{1, \dots, n-1\}, x_1 + \dots + x_i \geq i$  et si  $x_1 + \dots + x_n = n-1$ , alors on a clairement l'inclusion  $\{X_1 = x_1, \dots, X_n = x_n\} \subset \{N = n\}$ . Donc conditionnellement à cet événement, on a :

$$\begin{aligned} \mathbb{P}(X_1 = x_1, \dots, X_n = x_n, N = n) &= \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \\ &= \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_n = x_n) \text{ par indépendance} \\ &= p_{x_1} \dots p_{x_n} \end{aligned}$$



**Remarque.** Ici encore on suppose toujours que  $N$  est fini, mais le lien avec les marches aléatoires suggère une généralisation au cas  $N = \infty$ . En particulier, en notant  $\mathbb{E}(X)$  l'espérance mathématique de  $X$  :

- Si  $\mathbb{E}(X) < 1$ , la taille  $N$  de l'arbre est finie p.s. (c'est notre hypothèse).
- Si  $\mathbb{E}(X) = 1$ , la marche aléatoire est récurrente et  $N$  est arbitrairement grand.
- Si  $\mathbb{E}(X) > 1$ ,  $N$  est infini p.s., donc l'arbre aussi.

Il est également possible de considérer des **arbres de G-W conditionnés de taille  $N$**  (abrégé par G-W( $N$ )), pour  $N \in \mathbb{N} - \{0\}$ , où la taille de l'arbre est fixée d'emblée égale à  $N$ . On se donne encore une suite  $X_1, \dots, X_N$  pour le générer, mais il est important de remarquer que dans ce cas les  $X_i$  ne sont plus i.i.d. puisque, par exemple si  $N \geq 2$ ,  $X_1$  est obligatoirement à valeurs dans  $\mathbb{N} - \{0\}$ .

Il ressort donc de cette discussion le théorème récapitulatif :

**Théorème 2.3** Avec la numérotation choisie, fixée à l'avance :

- $\mathcal{A}$  est un arbre de G-W  $\iff \mathcal{S}$  est une marche aléatoire  $(S_0, \dots, S_N)$  sur  $\mathbb{Z}$  avec  $S_0 = 0$  et les  $(S_i - S_{i-1})$  i.i.d., distribuées comme  $X - 1$ , et où  $N$  est le premier instant tel que la marche aléatoire touche  $-1$ .
- $\mathcal{A}$  est un arbre de G-W( $N$ )  $\iff \mathcal{S}$  est une marche aléatoire  $(S_0, \dots, S_N)$  sur  $\mathbb{Z}$  conditionnée de taille  $N$ .

### 2.3 Arbre généré à partir d'un pont

On a ainsi vu une première façon de générer un arbre de G-W à partir d'une marche aléatoire particulière, i.e. une excursion, caractérisée elle-même par le vecteur des variations  $(X_1 - 1, \dots, X_N - 1)$ , où  $X_1, \dots, X_N$  sont à valeurs dans  $\mathbb{N}$ .

On va voir dans cette partie comment construire un arbre de G-W à partir d'une marche aléatoire un peu plus générale, allant de 0 à  $-1$  toujours, mais non nécessairement positive ou nulle avant  $-1$ , que l'on nommera "pont". Le point clef est de constater qu'il est possible de déduire une excursion à partir de cette forme de marches aléatoires. Pour cela, commençons plutôt par comprendre comment former une telle marche aléatoire à partir d'une excursion.

On se donne un arbre  $\mathcal{A}$  de taille  $N$  comme avant, et  $M \in \{1, \dots, N\}$  le numéro d'un individu de  $\mathcal{A}$ , que l'on imagine marqué. Le couple  $(\mathcal{A}, M)$  forme alors ce que l'on appellera un **arbre marqué**. On associe à  $(\mathcal{A}, M)$  un **pont**  $\mathcal{S}' = (S'_0, \dots, S'_N)$  défini par :

$$\begin{cases} S'_0 = 0 \\ \forall i \in \{1, \dots, N\}, S'_i = S'_{i-1} + X'_i - 1 \end{cases}$$

où  $X'_i = X_{\sigma(i)}$  avec  $1 \leq \sigma(i) \leq N$  tel que  $\sigma(i) \equiv i + M - 1 [N]$ .

**Proposition 2.4.** Le pont  $\mathcal{S}'$  se déduit de l'excursion  $\mathcal{S}$ , si  $M$  marqué, par la transformation :

$$S'_i = \begin{cases} S_{i+M-1} - S_{M-1} & \text{pour } 0 \leq i \leq N - M + 1. \\ S_{i+M-N-1} - S_{M-1} - 1 & \text{pour } N - M + 1 < i \leq N. \end{cases}$$

**Preuve 2.4.**  $S'_0 = 0 = S_{0+M-1} - S_{M-1}$  si  $i = 0$ , et sinon :

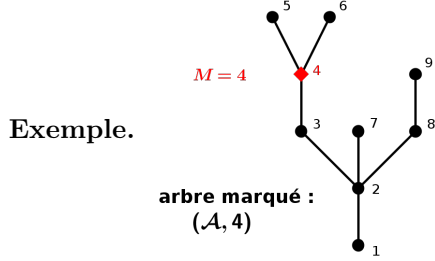
Si  $1 \leq i \leq N - M + 1$ ,  $X'_i = X_{i+M-1}$ , donc :

$$\begin{aligned} S'_i &= S'_{i-1} + X_{i+M-1} - 1 \\ &= (S'_{i-2} + X_{i+M-2} - 1) + X_{i+M-1} - 1 \\ &= \dots \\ &= S'_0 + X_M + \dots + X_{i+M-1} - i \\ &= 0 + (S_M - S_{M-1} + 1) + \dots + (S_{i+M-1} - S_{i+M-2} + 1) - i \\ &= S_{i+M-1} - S_{M-1} \end{aligned}$$

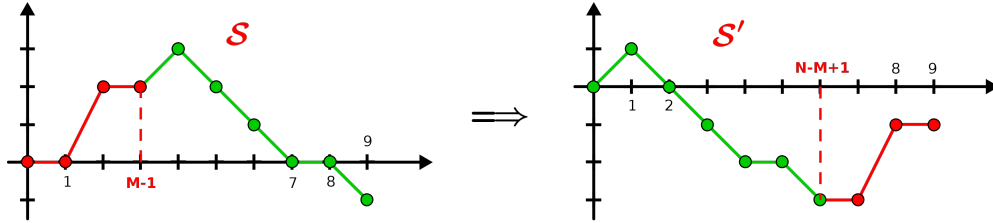
Enfin, si  $N - M + 1 < i \leq N$ , alors  $M > 1$  et  $X'_i = X_{i+M-N-1}$ , d'où :

$$\begin{aligned} S'_i &= S'_{i-1} + X_{i+M-N-1} - 1 \\ &= (S'_{i-2} + X_{i+M-N-2} - 1) + X_{i+M-N-1} - 1 \\ &= \dots \\ &= S'_{N-M+1} + X_1 + \dots + X_{i+M-N-1} - (i - N + M - 1) \\ &= (S_N - S_{M-1}) + X_1 + \dots + X_{i+M-N-1} - i + N - M + 1 \text{ par ce qui précède} \\ &= -1 - S_{M-1} + (S_1 - S_0 + 1) + \dots + (S_{i+M-N-1} - S_{i+M-N-2} + 1) - i + N - M + 1 \\ &= S_{i+M-N-1} - S_{M-1} - 1 \end{aligned}$$

**Remarque.** Cette transformation effectuée sur  $\mathcal{S}$  revient à décaler les variations  $S_i - S_{i-1}$ , qui valaient  $X_i - 1$  pour  $\mathcal{S}$  et désormais  $X_{\sigma(i)} - 1$  pour  $\mathcal{S}'$ , où  $\sigma$  est une translation modulo  $N$  de vecteur  $M - 1$ . Cela revient à couper  $\mathcal{S}$  au point  $M - 1$  en deux morceaux et à les intervertir. Remarquons de plus que si  $M = 1$ ,  $\mathcal{S}' = \mathcal{S}$ .



On considère ici l'arbre  $\mathcal{A}$  du premier exemple, que l'on suppose ici marqué en  $M = 4$ . On redonne le chemin  $\mathcal{S}$  associé, et le pont  $\mathcal{S}'$  obtenu par la définition, mettant ainsi en évidence la transformation que la définition revient à effectuer :



**Proposition 2.5.**  $\mathcal{S}'$  vérifie aussi  $S'_0 = 0$  et  $S'_N = -1$ . En revanche,  $\mathcal{S}'$  n'est plus nécessairement une excursion. En particulier,  $\mathcal{S}'$  atteint son premier minimum au point  $N - M + 1$ , qui vaut  $-1 - S_{M-1}$ .

**Preuve 2.5.**  $S'_0 = 0$  par définition. Et d'après la proposition 2.4.,

$$S'_N = \begin{cases} S_{N+1-1} - S_{1-1} = -1 & \text{si } M = 1, \\ S_{N+M-N-1} - S_{M-1} - 1 = -1 & \text{sinon.} \end{cases}$$

Comme  $S_N = -1$  est le minimum de  $(S_0, \dots, S_N)$ , donc de  $(S_{M-1}, \dots, S_N)$ , et comme  $S_0 = 0$  est un minimum de  $(S_0, \dots, S_{M-1})$  car  $M - 1 \neq N$ , il est clair que  $S'_{N-(M-1)}$  est le premier minimum de  $\mathcal{S}'$  d'après la transformation vue dans l'exemple précédent, mais pas nécessairement le seul. Or,  $S'_{N-(M-1)} = S_N - S_{M-1} = -1 - S_{M-1}$  d'après la proposition 2.4.

**Proposition 2.6.** À numérotation fixée, il est équivalent de se donner l'arbre marqué  $(\mathcal{A}, M)$ , et le pont  $\mathcal{S}'$  associé.

**Preuve 2.6.** Il suffit de montrer comment  $(\mathcal{A}, M)$  se retrouve à partir du pont  $\mathcal{S}'$ . Il est facile de déduire  $M$  en regardant le premier minimum de  $\mathcal{S}'$ , que l'on sait être  $S'_{N-M+1}$ . On retrouve alors aisément  $\mathcal{S}$  par la transformation inverse, qui consiste à couper  $\mathcal{S}'$  par analogie au point  $N - M + 1$ , et d'intervertir les deux morceaux.  $\mathcal{A}$  se retrouve ensuite à partir de  $\mathcal{S}$ , par la proposition 2.1.

Finalement, un pont  $\mathcal{S}'$  se déduit simplement de  $\mathcal{S}$  par la transformation vue précédemment, où il suffit de choisir un point  $M \in \{1, \dots, N\}$ . Il est clair (par la proposition 2.6.) que tout autre choix de  $M$  fournit un pont  $\mathcal{S}'$  différent pour un même arbre  $\mathcal{A}$ . En particulier, il existe  $N$  ponts  $\mathcal{S}'$  distincts qui fournissent un même arbre de G-W  $\mathcal{A}$  par la transformation étudiée.

**Remarque.**  $\mathcal{S}'$  vue comme une marche aléatoire est plus générale que  $\mathcal{S}$ , mais on impose encore la condition d'avoir des variations dans  $\{-1\} \cup \mathbb{N}$ . En particulier, une marche aléatoire quelconque de 0 à  $-1$  ne fournit pas forcément un arbre de G-W par une telle transformation (il faut éviter d'avoir un nombre d'enfants strictement négatif).

## 2.4 Probabilité de sortie d'un arbre

Les ponts  $\mathcal{S}'$  sont donc les marches aléatoires les plus générales que l'on considère pour générer un arbre de G-W dans le cadre qui nous occupe. Comme à chaque excursion  $\mathcal{S}$  de taille  $N$  on a associé  $N$  ponts  $\mathcal{S}'$  différents mais d'égale probabilité, il n'est pas étonnant d'obtenir le résultat suivant sur la probabilité qu'un arbre de G-W soit de taille  $N$ .

**Proposition 2.8.** Pour  $n \in \mathbb{N} - \{0\}$ ,  $\mathbb{P}(N = n) = \frac{1}{n} \mathbb{P}(S_n = -1)$ .

**Preuve 2.8.** Soit  $n \in \mathbb{N} - \{0\}$ . Pour  $1 \leq M \leq n$ , on pose :

$$\forall i \in \{1, \dots, n\}, S_i^{(M)} = \begin{cases} S_{i+M-1} - S_{M-1} & \text{pour } 0 \leq i \leq N - M + 1. \\ S_{i+M-N-1} - S_{M-1} - 1 & \text{pour } N - M + 1 < i \leq N. \end{cases}$$

$\mathcal{S}^{(M)}$  décrit quand  $M$  varie tous les ponts possibles  $\mathcal{S}'$  obtenus à partir de  $\mathcal{S}$  comme on l'a vu à  $M$  fixé. Comme  $N = \inf\{k/S_k = -1\}$ , on pose aussi  $N^{(M)} = \inf\{k/S_k^{(M)} = -1\}$ .

$$\begin{aligned}
\mathbb{P}(N = n) &= \mathbb{P}(S_n = -1, N = n) \\
&= \mathbb{P}(S_n^{(1)} = -1, N^{(1)} = n) && \text{car } \mathcal{S}' = \mathcal{S} \text{ si } M = 1 \\
&= \frac{1}{n} \sum_{M=1}^n \mathbb{P}(S_n^{(M)} = -1, N^{(M)} = n) && \text{car } M \text{ est uniforme sur } \{1, \dots, n\}. \\
&= \frac{1}{n} \mathbb{E} \left[ \sum_{M=1}^n 1_{\{S_n^{(M)} = -1, N^{(M)} = n\}} \right] \\
&= \frac{1}{n} \mathbb{E}[1_{\{S_n = -1\}}] \\
&= \frac{1}{n} \mathbb{P}(S_n = -1).
\end{aligned}$$

Par ailleurs, par analogie avec la probabilité de sortie de marches aléatoires sur  $\mathbb{Z}$  telles que les ponts  $\mathcal{S}'$ , on peut déduire la probabilité de sortie d'un arbre de G-W de taille  $N$ .

**Proposition 2.9.** Soit  $n \in \mathbb{N} - \{0\}$ . Soient  $y_1, \dots, y_n \in \mathbb{N}$  tels que  $y_1 + \dots + y_n = n - 1$ , on note ici aussi  $p_x = \mathbb{P}(X = x)$  pour tout  $x \in \mathbb{N}$ . Alors :

$$\mathbb{P}(S'_i - S'_{i-1} = y_i - 1 \text{ pour tout } i \in \{1, \dots, n\}) = \frac{1}{n} p_{y_1} \dots p_{y_n}.$$

**Preuve 2.9.** D'après ce qu'on a vu, comme  $y_1 + \dots + y_n = n - 1$  et que  $y_i \in \mathbb{N}$ , les  $y_1, \dots, y_n$  définissent un pont, à partir duquel on déduit une unique excursion marquée en un point  $m$ , tel que  $N - m + 1 = \min\{i/\forall k, S'_k \geq S'_i\}$ . En particulier, il existe une permutation  $\sigma$  telle que  $y_{\sigma(1)} + \dots + y_{\sigma(i)} \geq i$  pour tout  $i < n$ , et telle qu'on ait :

$$\begin{aligned}
\mathbb{P}(\forall i \in \{1, \dots, n\}, S'_i - S'_{i-1} = y_i - 1) &= \mathbb{P}(X_1 = y_{\sigma(1)}, \dots, X_n = y_{\sigma(n)}, N = n, M = m) \\
&= \frac{1}{n} \mathbb{P}(X_1 = y_{\sigma(1)}, \dots, X_n = y_{\sigma(n)}, N = n) \\
&= \frac{1}{n} p_{y_1} \dots p_{y_n}.
\end{aligned}$$

Dans cette partie, nous avons fait le lien entre arbres de G-W de taille  $N$  et certaine formes de marches aléatoires, à partir desquelles on peut déduire de nombreux résultats sur la taille d'un arbre de G-W, la représentation par une excursion à numérotation fixée, ou par un pont à  $M$  fixé, ou même la probabilité de sortie à taille fixée.

On peut donner le dernier théorème récapitulatif concernant cette fois-ci les ponts et les arbres de G-W conditionnés de taille  $N$  :

**Théorème 2.10.** Avec la numérotation choisie, fixée à l'avance, les assertions suivantes sont équivalentes :

- $(\mathcal{A}, M)$  est un arbre de G-W( $N$ ), et  $M$  est uniforme sur les nœuds de  $\mathcal{A}$ .
- $\mathcal{S}'$  est une marche aléatoire  $(S'_0, \dots, S'_N)$  sur  $\mathbb{Z}$  avec  $S'_0 = 0$  et les  $(S'_i - S'_{i-1})$  i.i.d., distribuées comme  $X - 1$ , conditionnée par l'évènement  $S'_N = -1$ .

### 3 Exemples d'arbres de Galton-Watson

On étudie ici des exemples d'arbres de G-W conditionnés de taille  $N = n \in \mathbb{N} - \{0\}$ , où  $X - 1$ , donc chaque composantes du vecteur des variations de  $\mathcal{S}'$ , suit une loi de probabilité bien connue.

#### 3.1 Le cas d'une distribution binaire

On suppose dans cet exemple que  $X$  est à valeurs dans  $\{0, 2\}$  avec :

$$\begin{cases} \mathbb{P}(X = 2) = p \in ]0, 1[, \\ \mathbb{P}(X = 0) = 1 - p = q. \end{cases}$$

$\mathcal{S}'$  s'apparente alors à une marche aléatoire simple (i.e. de pas 1) sur  $\mathbb{Z}$ .

**Proposition 3.1.** Soit  $n \in \mathbb{N} - \{0\}$  impair. Alors, dans le cas d'une distribution binaire, pour tout  $x_1, \dots, x_n \in \mathbb{N}$  :

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n, N = n) = \begin{cases} p^{\frac{n-1}{2}} q^{\frac{n+1}{2}} & \text{si } \forall i \in \{1, \dots, n-1\}, x_1 + \dots + x_i \geq i \\ & \text{et si } x_1 + \dots + x_n = n-1. \\ 0 & \text{sinon.} \end{cases}$$

**Preuve 3.1.**  $\mathcal{S}'$  est un pont tel que  $S'_0 = 0$ ,  $S'_n = -1$  et :

$$\begin{cases} \mathbb{P}(S'_i - S'_{i-1} = 1) = p \\ \mathbb{P}(S'_i - S'_{i-1} = -1) = q, \text{ pour } i \in \{1, \dots, n\}. \end{cases}$$

On remarque que pour aller de 0 à  $-1$  en  $n$  étapes par pas de 1, il est nécessaire que  $n$  soit impair, car il faut descendre exactement une fois de plus qu'on est monté, donc monter  $\frac{n-1}{2}$  fois (avec probabilité  $p$ ) et descendre  $\frac{n+1}{2}$  fois. La proposition 2.2., avec  $\frac{n-1}{2}$  des  $p_{x_i}$  qui valent  $p$  et  $\frac{n+1}{2}$  qui valent  $q$ , donne alors le résultat.

Par conséquent, puisque le nombre d'enfants d'un individu ne peut être que 0 ou 2, il est clair d'après la partie 2 que **n'importe quel arbre binaire de taille  $n$  (ordonné avec racine) est un arbre de G-W( $n$ )** associé à cette distribution.

Pour générer uniformément un arbre de G-W conditionné de taille  $n$ , on peut donc employer la première méthode suggérée par la partie 2, ou bien la deuxième méthode inspirée de la proposition 3.1. :

- **méthode 1** : Générer un pont  $\mathcal{S}'$  conditionné de taille  $n$ , i.e. une marche aléatoire de pas 1, allant de 0 à  $-1$  en  $n$  étapes, et en déduire l'unique excursion  $\mathcal{S}$  associée, puis l'unique arbre  $\mathcal{A}$  associé à cette excursion.
- **méthode 2** : Distribuer  $\frac{n-1}{2}$  balles aléatoirement dans  $n$  boîtes avec au plus une balle par boîte. Poser alors  $S'_i = 2\text{card}\{\text{balle parmi les boîtes } 1 \text{ à } i\} - i$  pour tout  $i \in \{1, \dots, n\}$ , et  $S'_0 = 0$ . Soit  $m = \min\{k/\forall i \in \{1, \dots, n\}, S'_i \geq S'_k\}$ , on pose alors  $X_i = S'_j - S'_{j-1} + 1$ , avec  $i \in \{1, \dots, n\}$ , et  $1 \leq j \leq n$  tel que  $j \equiv m + i[n]$ .

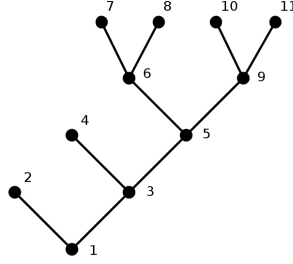
**Exemple.** Appliquons la méthode 2 à l'exemple suivant :



Dans cet exemple,  $\mathcal{S}' = (0, -1, 0, 1, 0, -1, 0, -1, -2, -1, -2, -1)$ . Alors  $m = 8$ , d'où :

$$\begin{aligned} X_1 &= S'_9 - S'_8 + 1 = 2, & X_5 &= 2, & X_9 &= 2, \\ X_2 &= S'_{10} - S'_9 + 1 = 0, & X_6 &= 2, & X_{10} &= 0, \\ X_3 &= S'_{11} - S'_{10} + 1 = 2, & X_7 &= 0, & X_{11} &= 0, \\ X_4 &= S'_1 - S'_0 + 1 = 0, & X_8 &= 0, & & \end{aligned}$$

Et l'arbre associé à cette répartition de balles est bien un arbre binaire :



**Proposition 3.2.** Il y a  $\binom{n}{\frac{n-1}{2}}$  ponts de pas 1 allant de 0 à  $-1$  en  $n$  étapes. Et il y a donc  $\frac{1}{n} \binom{n}{\frac{n-1}{2}}$  arbres binaires de G-W conditionnés de taille  $n$ .

**Preuve 3.2.** Par la méthode 2, il y a  $\binom{n}{\frac{n-1}{2}}$  manières de placer  $\frac{n-1}{2}$  balles dans  $n$  cases, chacune correspond à un unique pont  $\mathcal{S}'$ , et exactement  $n$  ponts distincts fournissent le même arbre  $\mathcal{A}$ .

Dans cet exemple, il est aussi possible de voir une excursion comme une chaîne de Markov inhomogène de taille  $n$ , dont la matrice de transition  $Q$  n'est pas très compliqué à déterminer ici puisque qu'on ne peut se retrouver que dans un état voisin pour une distribution binaire; disons avec probabilité  $p = \frac{1}{2}$  pour simplifier encore.

On considère alors la configuration un peu plus générale qui consiste à partir d'un état  $j$  au lieu de 0 ( $S_0 = j$ ). Soit  $q_m(j) = \mathbb{P}_j(S_m = -1; S_1, \dots, S_{m-1} \geq 0)$ .  $q_m(j)$  est donc la probabilité partant de  $j$ , d'arriver en  $-1$  pour la première fois au temps  $m$ .

**Remarque.**  $q_m(j) = 0$  si  $m$  et  $j$  ont la même parité.

**Proposition 3.3.** Pour  $0 \leq m \leq n-1$ ,

$$Q^{(m)}(i, j) = \begin{cases} \frac{q_{n-m-1}(j)}{2q_{n-m}(i)} & \text{si } j = i \pm 1 \\ 0 & \text{si } |j - i| > 1. \end{cases}$$

**Preuve 3.3.** Si  $|j - i| > 1$ , il n'est jamais possible de passer de  $i$  à  $j$  en 1 étape dans le cas binaire, donc  $Q^{(m)}(i, j) = 0$  pour tout  $0 \leq m \leq n - 1$ . Pour  $j, j_1, \dots, j_m \geq 0$  tel que  $|j_k - j_{k-1}| = 1$  si  $k \in \{1, \dots, m - 1\}$ ,

$$\begin{aligned} \mathbb{P}_j(S_1 = j_1, \dots, S_m = j_m | N = n) &= \frac{\mathbb{P}_j(S_1=j_1, \dots, S_m=j_m, N=n)}{\mathbb{P}_j(N=n)} \\ &= \frac{\mathbb{P}_j(S_1=j_1, \dots, S_m=j_m, S_n=-1, S_{m+1}, \dots, S_{n-1} \geq 0)}{\mathbb{P}_j(N=n)} \\ &= \frac{\mathbb{P}_j(S_1=j_1, \dots, S_m=j_m) q_{n-m}(j_m)}{\mathbb{P}_j(N=n)} \\ &= \frac{\mathbb{P}_j(S_1=j_1, \dots, S_{m-1}=j_{m-1}) q_{n-m}(j_m)}{2^{\mathbb{P}_j(N=n)}} \\ &= \frac{\mathbb{P}_j(S_1=j_1, \dots, S_m=j_m, S_n=-1, S_{m+1}, \dots, S_{n-1} \geq 0) q_{n-m}(j_m)}{2^{\mathbb{P}_j(N=n)} \mathbb{P}_j(S_n=-1, S_{m+1}, \dots, S_{n-1} \geq 0)} \\ &= \mathbb{P}_j(S_1 = j_1, \dots, S_{m-1} = j_{m-1} | N = n) \frac{q_{n-m}(j_m)}{2^{q_{n-m+1}(j_{m-1})}} \end{aligned}$$

Et comme on a par définition de  $Q$  :

$$\mathbb{P}_j(S_1 = j_1, \dots, S_m = j_m | N = n) = \mathbb{P}_j(S_1 = j_1, \dots, S_{m-1} = j_{m-1} | N = n) Q^{(m-1)}(j_{m-1}, j_m),$$

on obtient finalement :  $Q^{(m-1)}(j_{m-1}, j_m) = \frac{q_{n-m}(j_m)}{2^{q_{n-m+1}(j_{m-1})}}$ .

**Lemme 3.4.** Pour  $m < n - 1$  et  $1 \leq j < n - m$ ,  $q_m(j - 1) = \frac{j}{m} \binom{m}{\frac{m+j}{2}} \frac{1}{2^m}$ .

**Preuve 3.4.**  $q_m(j - 1) = \mathbb{P}_{j-1}(N = m) = \frac{j}{m} \mathbb{P}_{j-1}(S_m = -1)$  d'après la formule de Kemperman (admise ici). Or,  $\mathbb{P}_{j-1}(S_m = -1) = \mathbb{P}_0(S_m = j) = \mathbb{P}_0(\frac{m+S_m}{2} = \frac{m+j}{2})$ . Et comme, sous  $\mathbb{P}_0$ ,  $\frac{m+S_m}{2}$  suit une loi binomiale  $\mathcal{B}(m, \frac{1}{2})$ , on a pour finir :

$$\mathbb{P}_0(\frac{m+S_m}{2} = \frac{m+j}{2}) = \binom{m}{\frac{m+j}{2}} \frac{1}{2^m}, \text{ d'où } q_m(j - 1) = \frac{j}{m} \binom{m}{\frac{m+j}{2}} \frac{1}{2^m}.$$

**Proposition 3.5.** Pour une distribution binaire, la matrice de transition de la chaîne de Markov  $Q$  associée à l'excursion  $\mathcal{S}$  est donnée pour tout  $m < n - 1$ , et  $1 \geq j < n - m$ , par :  $Q^{(m)}(j - 1, j) = \frac{(j+1)(n-m-j)}{2j(n-m-1)}$  et  $Q^{(m)}(j - 1, j - 2) = \frac{(j-1)(n-m+j)}{2j(n-m-1)}$ .

**Preuve 3.5.** Cela se déduit facilement de la proposition 3.3. et du lemme 3.4. :

$$\begin{aligned} Q^{(m)}(j - 1, j) &= \frac{q_{n-m-1}(j)}{2^{q_{n-m}(j-1)}} = \frac{\frac{j+1}{n-m-1} \binom{n-m-1}{\frac{n-m+j}{2}} \frac{1}{2^{n-m-1}}}{\frac{2j}{n-m} \binom{n-m}{\frac{n-m+j}{2}} \frac{1}{2^{n-m}}} \\ &= \frac{(j+1)(n-m)(n-m-\frac{n-m+j}{2})}{j(n-m-1)(n-m)} = \frac{(j+1)(n-m-j)}{2j(n-m-1)}. \end{aligned}$$

$$\text{Et de la même façon, } Q^{(m)}(j - 1, j - 2) = \frac{q_{n-m-1}(j-2)}{2^{q_{n-m}(j-1)}} = \dots = \frac{(j-1)(n-m+j)}{2j(n-m-1)}.$$

Avec la proposition 3.5., on obtient une dernière méthode très effective pour générer uniformément des arbre de G-W(n) associés à la distribution binaire :

- Générer une excursion  $\mathcal{S}$  de taille  $n$  grâce aux matrices de transition  $Q^{(m)}$ .
- Déduire l'arbre  $\mathcal{A}$  de l'excursion  $\mathcal{S}$ .


### 3.2 Le cas d'une distribution géométrique

On suppose maintenant que  $X$  suit une loi géométrique :  $\mathbb{P}(X = k) = pq^k$  pour tout  $k \in \mathbb{N}$ , avec  $1 - q = p \in ]0, 1[$ . Alors, pour tout  $x_1, \dots, x_n \in \mathbb{N}$  :

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n, N = n) = \begin{cases} p^n q^{n-1} & \text{si } \forall i \in \{1, \dots, n-1\}, x_1 + \dots + x_i \geq i \\ & \text{et si } x_1 + \dots + x_n = n-1. \\ 0 & \text{sinon.} \end{cases}$$

Comme  $X$  est à valeurs dans  $\mathbb{N}$  dans ce cas, **n'importe quel arbre (ordonné avec racine) de taille  $n$  est également un arbre de G-W( $n$ )** associé à cette distribution. Cela suggère une méthode analogue à celle de l'exemple précédent pour générer uniformément n'importe quel arbre de taille  $n$  :

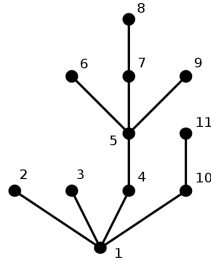
- Distribuer  $n - 1$  balles aléatoirement dans  $n$  boîtes (avec éventuellement plusieurs dans la même cette fois-ci). Poser  $S'_i = \text{card}\{\text{balles dans les boîtes } 1 \text{ à } i\} - i$  pour tout  $i \in \{1, \dots, n\}$ , et  $S'_0 = 0$ .
- Soit  $m = \inf\{k/\forall i \in \{1, \dots, n\}, S'_k \leq S'_i\}$ , poser pour tout  $i \in \{1, \dots, n\}$  :  
 $X_i = S'_{m+j} - S'_{m+j-1} + 1$ , où  $1 \leq j \leq n$  tel que  $j \equiv i[n]$ .

**Exemple.**  $n = 11$ ,   **$n - 1$  balles.**

Ici,  $\mathcal{S}' = (0, -1, -2, -2, -3, 0, -1, -2, -2, 0, -1, -1)$  et  $m = 4$ , donc on a :

$$\begin{aligned} X_1 &= S'_5 - S'_4 + 1 = 4, & X_5 &= 3, & X_9 &= 0, \\ X_2 &= 0, & X_6 &= 0, & X_{10} &= 1, \\ X_3 &= 0, & X_7 &= S'_{11} - S'_{10} + 1 = 1, & X_{11} &= 0. \\ X_4 &= 1, & X_8 &= S'_1 - S'_0 + 1 = 0, \end{aligned}$$

L'arbre associé à cette répartition de balles est dans ce cas :



Finalement, l'étude des liens entre arbres et marches aléatoires permet, comme on peut le voir à travers ces deux exemples, d'imaginer de nouvelles façons de générer des arbres de G-W (en particulier conditionnés de taille  $n$ ), et de déduire par analogie des propriétés importantes, telles que par exemple leur nombre, à l'aide d'arguments combinatoires simples, ou leur forme.



## Annexe C

# Mémoire de M2

### C.1 Présentation du domaine de recherche : « Les algorithmes génétiques simples »

Rédaction de fin de magistère :

# Les algorithmes génétiques simples

Jimmy PAROT,  
étudiant à l'université Paris-Sud XI (Orsay).

30 octobre 2015

---

## Table des matières

<b>1</b>	<b>Présentation</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Les algorithmes génétiques simples . . . . .	3
1.3	L'évaluation Route Royale . . . . .	3
<b>2</b>	<b>Des idées de chercheurs</b>	<b>4</b>
2.1	Les travaux de VAN NIMWEGEN et al . . . . .	4
2.2	L'article de Raphaël CERF . . . . .	5
<b>3</b>	<b>Mon mémoire de M2</b>	<b>7</b>
3.1	Quelques résultats . . . . .	7
3.2	Les grandes perspectives . . . . .	8

---

# 1 Présentation

## 1.1 Introduction

Les algorithmes génétiques sont des algorithmes aléatoires célèbres qui s'inspirent du mode de reproduction naturelle des chromosomes au sein d'une population. Ils font particulièrement leurs preuves là où les algorithmes déterministes standards se heurtent à une trop grande quantité de calculs pour résoudre un problème d'optimisation. Citons l'exemple célèbre du voyageur de commerce, dont le parcours optimal est particulièrement difficile à obtenir par le calcul lorsque le nombre de villes à visiter devient relativement grand (100 à 200...). C'est donc par exemple dans ce genre de situation qu'un algorithme aléatoire est le seul choix possible, car on peut alors espérer que celui-ci trouve la (ou une) solution optimale « par hasard » et la retourne en un temps raisonnable.

Les algorithmes génétiques sont donc parmi les algorithmes aléatoires les plus utilisés en simulation aléatoire depuis longtemps, bien que personne n'ait pu démontrer leur efficacité par le calcul. Les résultats satisfaisants obtenus à l'aide de paramètres empiriques dans des situations concrètes ont semblé suffire aux praticiens, mais d'un point de vue mathématique une preuve algébrique était grandement attendue. En 1999, Erik VAN NIMWEGEN et deux autres chercheurs se sont penchés avec habileté sur cette question, dans l'article [NCM99] qui fait référence depuis cette date (avec ceux qui ont suivi). Leur objectif était double : d'une part prouver l'existence de paramètres pour lesquels la performance des algorithmes génétiques est assurée, du moins dans un cas particulier simple, et d'autre part exhiber des jeux précis de paramètres garantissant cette performance. Cependant, leur travail n'a pas abouti sur une démonstration complète à cause de la méthode employée, mais beaucoup de résultats présents dans l'article obtenus par simulation restent aujourd'hui encore très intéressants.

Quoi qu'il en soit, dans un article récent de février 2015, à savoir [Cer15], le chercheur (au DMA de l'ENS Ulm) Raphaël CERF propose cette fois une démonstration complète, en utilisant des méthodes de grandes déviations, qui garantit l'existence de jeux de paramètres pour lesquels les algorithmes génétiques simples, dans un cadre très général cette fois, sont performants en différents sens que nous détaillerons plus loin.

En revanche, bien que cette démonstration soit encourageante du point de vue théorique, son caractère général ne permet en fait pas d'exhiber des paramètres utilisables en pratique. Dans le cadre de mon mémoire de M2, et sous la direction de Raphaël CERF, mon rôle fut d'adapter la démonstration présente dans l'article à un cas particulier simple qui est le même que celui considéré par VAN NIMWEGEN, dans l'espoir d'obtenir des résultats plus explicites, ainsi que des paramètres utilisables en simulation. Cependant, je n'ai pas réussi à terminer les calculs dans le temps qui m'était imparti.

Dans la fin de cette partie, nous parlerons de généralités sur les algorithmes considérés pour la recherche. Nous discuterons ensuite des idées présentes dans les articles célèbres cités plus haut, puis nous aborderons pour finir dans la partie 3 l'apport de mon mémoire de M2 dans ce domaine.

## 1.2 Les algorithmes génétiques simples

Les algorithmes génétiques simples imitent les étapes d'évolution d'une population, en simulant les processus de sélection naturelle, de reproduction et même celui des mutations génétiques. Concrètement, un tel algorithme démarre avec la donnée d'une population de  $m$  chromosomes (nous pouvons supposer  $m$  pair), où chaque chromosome est en fait constitué de  $l$  sites pouvant prendre exclusivement les valeurs 0 ou 1. Ensuite, pour passer d'une génération à l'autre, l'algorithme répète les étapes successives suivantes :

**Sélection :** L'algorithme tire aléatoirement (avec remplacement)  $m$  nouveaux chromosomes parmi les  $m$  de départ, selon une loi à préciser qui sélectionne les « meilleurs » chromosomes de façon préférentielle.

**Croisements :** Les chromosomes ainsi choisis sont rangés par paires, et chaque paire procède à un échange d'informations aléatoire, que l'on appelle croisement, avec probabilité  $p_C \in [0, 1]$ .

**Mutations :** Les  $m$  chromosomes résultant des deux opérations précédentes subissent pour finir des mutations indépendantes sur chaque site avec probabilité  $p_M \in [0, 1]$ .

Bien que l'idée générale soit relativement claire et simple à comprendre, il est important de noter que les trois étapes de sélection, croisements et mutations doivent être détaillées pour comprendre le fonctionnement réel de l'algorithme.

L'étape la plus importante à décrire est celle de la sélection. Pour mieux la comprendre, on peut considérer qu'elle s'opère en deux temps : le premier est une phase d'*évaluation*, qui consiste à déterminer l'efficacité de chaque chromosome dans une population, puis le deuxième temps correspond à la phase de sélection à proprement parler, en fonction des efficacités trouvées juste avant. Nous parlerons de l'exemple célèbre d'évaluation Route Royale à la fin de cette partie, et des modes de sélection utilisés par chaque chercheur en temps voulu. Les étapes de croisements et mutations ne varient pas d'un article à l'autre, et nous les laissons donc de côté pour cette rédaction.

## 1.3 L'évaluation Route Royale

Pour parler de l'évaluation que nous appelons Route Royale, il est préférable de commencer par dire un mot sur le modèle mathématique employé dans ce domaine de recherche. D'après les hypothèses faites, les chromosomes sont des séquences de 0 et de 1 de longueur  $l$ . Ils sont donc modélisés simplement par un vecteur-ligne, ou encore par un élément de  $\{0, 1\}^l$ , et une population de  $m$  chromosomes est ainsi modélisée par une matrice  $m \times l$  où chaque ligne représente un chromosome, ou aussi bien par un élément de  $(\{0, 1\}^l)^m$ . Si  $x$  est une population, on note alors agréablement :

$$x = \begin{pmatrix} x(1) \\ \vdots \\ x(m) \end{pmatrix} = \begin{pmatrix} x(1, 1) & \dots & x(1, l) \\ \vdots & \ddots & \vdots \\ x(m, 1) & \dots & x(m, l) \end{pmatrix}$$

Ainsi, une *fonction d'évaluation* est la donnée d'une application de  $\{0, 1\}^l$  dans  $\mathbb{R}^+$  qui à tout chromosome associe une efficacité. De plus, l'algorithme est performant s'il permet de converger vers un chromosome qui réalise le maximum de cette fonction, que l'on appelle alors *chromosome dominant*<sup>1</sup>. Dans le modèle considéré, on supposera pour simplifier, et sans perte de généralité, que la séquence  $u^* = 111\dots 1$  est un chromosome dominant (non nécessairement unique).

Ceci étant dit, il existe bien des façons de choisir une telle fonction d'évaluation, et nous présentons ici l'une des plus célèbres, qui est celle utilisée par VAN NIMWEGEN et al dans l'article [NCM99]. L'évaluation Route Royale a l'avantage d'être à la fois simple et suffisamment générale pour fournir des résultats satisfaisants, comme le prouvent les simulations présentées dans cet article. L'idée est de découper un chromosome en  $N$  blocs de tailles  $K$ , c'est-à-dire de supposer que  $l = NK$ , avec  $N$  et  $K$  des entiers naturels non nuls.

$$\underbrace{0001011101}_K \underbrace{1011110\dots\dots\dots}_K \underbrace{0001011011}_K$$

En notant alors  $u_i$  le  $i^{\text{ème}}$  bloc de longueur  $K$  d'un chromosome  $u$ , pour tout  $i \in \{1, \dots, N\}$ , on peut définir la fonction d'évaluation Route Royale  $f : \{0, 1\}^l \rightarrow \{0, 1, \dots, N\}$  par

$$f(u) = \sum_{i=1}^N \mathbb{1}_{u_i = u_i^*} = \sum_{i=1}^N \mathbb{1}_{u_i = 1\dots 1}.$$

Lorsqu'un bloc  $u_i$  de longueur  $K$  d'un chromosome ne contient que des 1, on dit que ce bloc est *aligné*, sous-entendu avec les blocs du chromosome dominant  $u^*$ . La fonction d'évaluation Route Royale compte donc simplement le nombre de blocs d'un chromosome qui sont alignés. Si l'algorithme est performant, on peut donc déjà imaginer que celui-ci doit converger (en un certain sens) vers le chromosome dominant  $u^*$ .

## 2 Des idées de chercheurs

Nous commentons ici les idées générales présentes dans les deux articles importants pour cette théorie.

### 2.1 Les travaux de VAN NIMWEGEN et al

Mathématiquement, il est naturel de modéliser les actions d'un tel algorithme par un processus  $(X_n)_{n \in \mathbb{N}}$ , qui partant d'une population initiale  $x_0 \in (\{0, 1\}^l)^m$  subit de manière répétée les étapes de sélection, croisements et mutations. Ces trois étapes définissent de plus à chaque génération  $n$  un opérateur aléatoire  $\Phi_n$  agissant sur le processus, c'est-à-dire tel que  $\forall n \in \mathbb{N}, X_{n+1} = \Phi_n(X_n)$ . Un

---

1. Un tel maximum est garanti tant que  $l$  reste fini, car alors l'espace de départ est fini.

premier constat est de voir que le processus  $(X_n)_{n \in \mathbb{N}}$  précédemment défini est en fait une chaîne de Markov, même homogène, puisqu'à chaque génération, l'algorithme ignore les modifications réalisées antérieurement et transforme les populations en répétant le même procédé aléatoire à chaque tour.

L'idée de Erik VAN NIMWEGEN, James P. CRUTCHFIELD et Melanie MITCHELL fut alors d'étudier la matrice de passage du processus, et plus précisément ses itérations afin de comprendre la dynamique d'un tel algorithme. Même dans un cadre général, la matrice de transition est d'ailleurs relativement facile à déterminer. Il suffit pour cela d'étudier quelque peu en détail chacune des trois étapes de l'algorithme. Cette étude permet rapidement de déterminer les probabilités de transitions  $P_S(x, y)$ ,  $P_C(y, z)$  et  $P_M(z, x')$ , qui sont respectivement les probabilités de passer de la population  $x$  à la population  $y$  par l'étape de sélection, celle de passer de  $y$  à  $z$  par l'étape de croisements, et enfin celle de passer de  $z$  à  $x'$  par l'étape de mutation<sup>2</sup>. Ces trois transitions déterminent alors la probabilité de passer de la population  $x$  à la population  $x'$  en un tour d'algorithme, via la formule

$$P(x, x') = \sum_{y, z} P_S(x, y) P_C(y, z) P_M(z, x').$$

Pour un maximum de simplicité, VAN NIMWEGEN limita cependant son étude à un algorithme génétique plus léger, en ne gardant que les étapes de sélection et mutations (ce qui revient à choisir  $p_C = 0$ ). Il considéra de plus la fonction d'évaluation particulière Route Royale, et choisit enfin la sélection proportionnelle à l'évaluation, qui est la plus simple que l'on puisse imaginer<sup>3</sup>. Malgré cela, l'étude de la matrice de transition s'est révélée être trop délicate pour aboutir, et seules des simulations ont été possibles pour tenter d'expliquer la dynamique de l'algorithme simplifié.

## 2.2 L'article de Raphaël CERF

Dans l'article de Raphaël CERF, la démarche est tout autre, bien que les objectifs soient communs.

Premièrement, l'étude est menée dans une généralité la plus totale, en ne spécifiant aucune fonction d'évaluation, et en considérant l'algorithme complet, avec notamment l'étape intéressante de croisements. En outre, l'auteur considère le cas de la sélection par rang, qui s'opère en triant les chromosomes par ordre d'efficacité avant de les choisir uniquement en fonction de leur position (sélection linéaire), ou bien la sélection par tournois, qui rentre également dans le cadre de son étude. Notons seulement que la sélection par rang est une des plus simples que l'on puisse choisir, plus simple encore que la sélection proportionnelle, ce qui compense le fait de travailler avec des fonctions d'évaluation générales  $f : \{0, 1\}^l \rightarrow \mathbb{R}^+$ , mais possède aussi l'inconvénient de perdre de l'information sur

2. Ces calculs sont détaillés dans l'article de Raphaël CERF, nous ne les répétons pas ici.

3. Chaque chromosome est sélectionné avec une probabilité proportionnelle à son nombre de blocs alignés.

l'efficacité des chromosomes pour les choisir, et donc de permettre à l'algorithme de converger peut-être moins rapidement.

Deuxièmement, l'approche mathématique est également différente. Il ne s'agit plus de s'attaquer directement aux itérations de la matrice de transition de la chaîne de Markov, mais plutôt de s'intéresser à des processus dérivés du processus initial  $(X_n)_{n \in \mathbb{N}}$ , définis à  $\lambda \in \mathbb{R}$  fixé par

$$\forall n \in \mathbb{N}, \quad N(X_n, \lambda) = \text{card}\{i \in \{1, \dots, m\} \mid f(X_n(i)) \geq \lambda\}.$$

Ces processus dénombrent finalement les chromosomes présents à chaque génération qui possèdent une efficacité plus grande que  $\lambda$ , et sont donc à valeurs dans  $\{0, \dots, m\}$ . L'idée ici est en fait d'oublier la composition détaillée de la population, et de s'intéresser uniquement à ces quantités, qui suffisent a priori pour décider de la performance de l'algorithme. Cependant, si ces quantités sont en apparence plus simples, ce ne sont plus du tout des chaînes de Markov, car la probabilité de passer d'un nombre donné de chromosomes d'efficacité plus grande que  $\lambda$  à un autre ne dépend pas uniquement du nombre de départ, mais aussi de la composition précise de la population.

La méthode de démonstration utilisée est donc de comparer à l'aide d'un couplage ces processus à des processus dérivés d'un algorithme modifié de sorte à obtenir des chaînes de Markov. Il s'agit ensuite d'étudier les chaînes de Markov majorantes et minorantes, et les inégalités obtenus par des méthodes de grandes déviations sur ces chaînes sont ainsi répercutées sur les processus initiaux.

Grâce à cette méthode, les résultats obtenus par Raphaël CERF sont de différentes sortes. Nous présentons ici les principaux.

**Le régime désordonné :** Le premier résultat fut d'établir, par l'exemple, qu'il existe des jeux de paramètres (i.e. des choix de  $f$ ,  $x_0$ ,  $p_C$ ,  $p_M$ ,  $l$ ,  $m$ , etc...) tels que les algorithmes génétiques simples ne sont pas performants au sens où, asymptotiquement, un chromosome dominant est détruit au bout d'un certain temps avec une probabilité strictement positive (et même tendant vers 1).

**Le régime de quasi-espèce :** Les deux grands résultats suivants sont du même ordre d'idée, et prouvent la performance asymptotique des algorithmes génétiques simples en deux sens. Nous traduisons directement les énoncés de l'article [Cer15] ici, où  $\sigma$  désigne un paramètre lié au mode de sélection, en laissant volontairement la formulation peu rigoureuse de l'auteur dans la première formule<sup>4</sup>.

**Théorème 1** (Cerf [2015]) Soit  $f_0^* = \max_{1 \leq i \leq m} f(x_0(i))$  l'efficacité maximale présente dans la population initiale, et soit  $\pi = \sigma(1 - p_C)(1 - p_M)^l$ . Nous supposons que  $\pi > 1$ , alors il existe  $V^* > 0$  et  $p^* > 0$ , dépendant uniquement de  $\pi$ , tels que pour tous  $l \geq 1$  et  $m \geq 1$

---

4. Que quiconque peut traduire mathématiquement sans confusion.

$$\mathbb{P}\left(\begin{array}{c} \text{la population contient toujours un chromosome} \\ \text{d'efficacité plus grande ou égale à } f_0^* \\ \text{jusqu'à la génération } \exp(V^*m) \end{array}\right) \geq p^*.$$

**Théorème 2** (Cerf [2015]) Soit  $\Lambda(X_n, r) = \max\{\lambda \in \mathbb{R} \mid N(X_n, \lambda) \geq r\}$  la plus grande efficacité  $\lambda$  représentée au moins  $r$  fois dans la population à la génération  $n$ , et soit  $\pi = \sigma(1 - p_C)(1 - p_M)^l$ . Nous supposons que  $\pi > 1$ , alors il existe  $\rho^* > 0$ ,  $c^* > 0$  et  $m_0 > 0$ , dépendant uniquement de  $\pi$ , tels que pour tout  $m \geq m_0$

$$\mathbb{P}\left(\forall n \leq \exp(c^*m), \max_{1 \leq i \leq m} f(X_n(i)) \geq \max_{0 \leq s \leq n} \Lambda(X_s, \lfloor \rho^*m \rfloor)\right) \geq 1 - \exp(-c^*m).$$

Ce dernier théorème prouve la performance des algorithmes génétiques en un sens plus fort que le précédent, car il affirme que la probabilité pour que, au moins jusqu'à un temps  $\exp(c^*m)$ , l'efficacité maximale de la population soit toujours au moins égale à la plus grande efficacité relevée jusqu'à présent, possédée par les  $\lfloor \rho^*m \rfloor$  meilleurs chromosomes simultanément au sein d'une population, tend vers 1 quand  $m$  tend vers l'infini.

Grâce à ces résultats, nous avons donc maintenant une démonstration que les algorithmes génétiques simples peuvent permettre de résoudre des problèmes d'optimisation sous certaines conditions. Concernant ces conditions, le chercheur Raphaël CERF s'est appliqué à mettre en évidence un paramètre critique  $\pi$ , qui détermine apparemment le régime de l'algorithme. Cela permet notamment de se rendre compte que pour espérer un algorithme performant, il est nécessaire de choisir les paramètres de sorte que  $\pi > 1$ . Cependant, ces théorèmes sont asymptotiques, et il est difficile de quantifier un rang à partir duquel on peut obtenir de bons résultats en pratique.

### 3 Mon mémoire de M2

Ce mémoire fut intitulé *Évaluation Route Royale et sélection par rang pour les algorithmes génétiques simples, sans croisements*, et rédigé sous la direction de Raphaël CERF dans le cadre de la validation de mon M2 Mathématiques Fondamentales et Appliquées, spécialité Probabilité et Statistique. Nous présentons ici un résumé succinct du travail effectué, ainsi que des pistes pour les recherches à venir dans ce domaine.

#### 3.1 Quelques résultats

Comme son nom l'indique, le cadre de mon mémoire est analogue à celui proposé par Erik VAN NIMWEGEN en 1999, à savoir les algorithmes génétiques sans croisements, et avec la fonction d'évaluation Route Royale. Le but est bien sûr de considérer les hypothèses les plus simplificatrices possible afin d'obtenir un premier résultat convaincant. La seule différence est d'avoir considéré ici la sélection par rang, encore plus simple que la sélection proportionnelle.



L'objectif de ce mémoire était de réinvestir les méthodes de démonstration présentes dans l'article [Cer15] dans ce cadre très particulier, dans l'espoir que la donnée d'une fonction d'évaluation particulière fournisse un théorème du même type que ceux énoncés plus haut mais davantage explicite, c'est-à-dire notamment avec des constantes connues (qui s'expriment avec les paramètres de l'algorithme).

Après un premier temps pour assimiler les grandes idées de la démonstrations, j'ai déroulé les calculs jusqu'à aboutir à une première inégalité de grandes déviations, particulièrement importante. Pour comparaison, les processus dérivés à considérer furent ceux particulièrement simples définis pour tout

$$\lambda \in \{0, \dots, N\} \text{ par } \forall n \in \mathbb{N}, \quad N(X_n, \lambda) = \sum_{i=1}^m \mathbb{1}_{\{f(X_n(i))=\lambda\}}.$$

Les calculs effectués ont principalement permis de montrer que, pour tout  $s \in ]0, 1[$  et  $\Gamma \subset [0, 1]$ ,

$$\mathbb{P}(N_{n+1} \in m\Gamma \mid N_n = \lfloor sm \rfloor) \underset{m \rightarrow \infty}{\sim} \exp(-m \inf\{J(s, t, u) : t \in \Gamma, u \in [t, 1]\}),$$

où  $J$  est une fonction assez compliquée à décrire, mais qui s'exprime en fonction des paramètres de l'algorithme. C'est à partir de cette équivalence que l'on peut ensuite déduire dans une démonstration complète un théorème comme le théorème 1. La seule remarque que l'on peut vraiment faire à ce stade des calculs est de noter que la borne inférieure trouvée ici s'exprime comme attendu de façon plus précise vis à vis des paramètres de l'algorithme qu'il ne l'était dans le cadre général de l'article de référence.

### 3.2 Les grandes perspectives

Pour terminer la présentation de ce domaine de recherche, disons simplement que les perspectives à venir sont très probablement dans l'esprit de mon mémoire, où le premier travail serait bien évidemment de terminer les calculs commencés dans celui-ci, et de voir si nous obtenons bel et bien des constantes explicites.

Par la suite, il s'agirait également de quantifier si possible dans quelles mesures certains jeux de paramètres sont du fait utilisables en simulation pour cet algorithme. Puis enfin d'élargir le calcul à d'autres algorithmes particuliers, comme par exemple en autorisant l'étape de croisements, ou encore en considérant plutôt la sélection proportionnelle, a priori plus prometteuse que la sélection par rang.

## Références

- [Cer15] Raphaël Cerf. *The quasispecies regime for the genetic algorithm with ranking selection*. 2015.
- [NCM99] Erik Van Nimwegen, James P. Crutchfield, and Melanie Mitchel. *Statistical Dynamics of the Royal Road Genetic Algorithm*. 1999.

## C.2 Mémoire : « Évaluation Route Royale et sélection par rang pour les algorithmes génétiques simples, sans croisements »

Mémoire de M2 recherche :

# Évaluation Route Royale et sélection par rang pour les algorithmes génétiques simples, sans croisements

Jimmy PAROT,  
étudiant en M2 MFA,  
Paris-Sud XI (Orsay).

6 septembre 2015

---

Ce mémoire est encadré par Raphaël CERF, enseignant-chercheur, DMA de l'ENS Ulm (Paris).

## Résumé

Le sujet de ce mémoire est l'étude des algorithmes génétiques simples dans le cas de la sélection Route Royale. Nous reprenons dans ce cas particulier les méthodes de grandes déviations utilisées par Raphaël CERF dans un article de 2015<sup>1</sup> avec l'espoir d'établir des résultats plus explicites, et d'explicitier par ailleurs des jeux de paramètres utilisables en simulation. Dans ce deuxième ordre d'idée, nous chercherons également à comparer nos résultats avec ceux obtenus dans les travaux très intéressants de VAN NIMWEGEN, P. CRUTCHFIELD et MITCHELL à ce sujet, publiés dans un article de 1999<sup>2</sup>, afin de mieux juger de la validité des méthodes utilisées pour cette étude des algorithmes génétiques.

---

1. *The quasispecies regime for the simple genetic ranking selection* (non publié).  
2. *Statistical Dynamics of the Royal Road Genetic Algorithm*.

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Présentation du modèle mathématique</b>	<b>3</b>
2.1	L'algorithme sans croisements . . . . .	4
2.2	Évaluation Route Royale et sélection par rang . . . . .	4
2.3	Définition des processus . . . . .	5
<b>3</b>	<b>Domination stochastique inférieure</b>	<b>6</b>
3.1	Minoration du processus . . . . .	6
3.2	Construction du couplage . . . . .	8
3.3	Minoration stochastique . . . . .	9
<b>4</b>	<b>Estimation du temps de survie</b>	<b>10</b>
4.1	Première inégalité de grande déviation . . . . .	10
4.2	Étude de la borne . . . . .	13
4.3	Dynamique du processus . . . . .	14
4.4	Temps moyen de survie . . . . .	16

## 1 Introduction

Cet document est écrit dans le cadre d'un mémoire de M2 recherche, dont le but est l'étude des algorithmes génétiques dans le cas de la sélection que nous appellerons Route Royale. Il est encadré par l'enseignant-chercheur Raphaël CERF, du département des mathématiques et applications (DMA) de l'ENS Ulm à Paris, qui a déjà publié de nombreux articles sur le sujet à ce jour, et dont nous nous inspirons naturellement.

Les algorithmes génétiques sont des algorithmes d'optimisation très célèbres, qui s'inspirent du mode de reproduction naturel des chromosomes (ou des gènes) dans une population, en réinvestissant les concepts de sélection naturelle, croisements et mutations aléatoires. Leur célébrité vient principalement de la diversité des domaines d'application dans lesquels ils sont utilisés, où les algorithmes déterministes ne peuvent être utilisés efficacement, comme par exemple dans la résolution du fameux problème du voyageur de commerce et bien d'autres. La clé de leur efficacité vient donc de l'étape de sélection naturelle, comme dans la nature, qui permet de converger vers une population (au bout d'un grand nombre de générations) qui ne contient plus que les chromosomes *dominants* parmi ceux présents au départ, c'est-à-dire ceux qui offrent la meilleure adaptation à l'environnement, etc... Pour le reste, les croisements et les mutations font de même partie du processus naturel d'évolution d'une population, et servent quant à eux à améliorer cette convergence, au sens où ils rendent sans cesse possible l'apparition d'un chromosome encore jamais observé et donc potentiellement plus performant que ses prédécesseurs.

En simulation, nous avons donc besoin de critères pour pouvoir dire si un chromosome est plus performant qu'un autre, afin de refléter la sélection naturelle. Cela repose mathématiquement sur la donnée d'une fonction d'évaluation, capable de mesurer l'*efficacité* d'un chromosome donné. Le cas des fonctions dites Route Royale fit l'objet d'un article important de 1999, publié par VON NIMWEGEN, P. CRUTCHFIELD et MITCHELL, dans lequel les auteurs ont tenté d'explicitier par une approche qualitative les régimes (tailles de population, taux de mutation, etc...) dans lesquels les algorithmes génétiques sans croisements convergent comme voulu vers une population idéale. Ces travaux ont fourni des régimes de convergence empiriques, à l'aide de nombreuses simulations, mais n'ont pas pu aboutir à une démonstration rigoureuse des résultats observés par simulation malgré de nombreuses hypothèses simplificatrices du modèle. Et pour cause, la méthode, cependant intéressante et naturelle, qui consistait à modéliser l'algorithme par une chaîne de Markov puis étudier les itérations de la matrice de transition est une tâche qui s'avère finalement trop difficile pour permettre d'obtenir des résultats explicites et rigoureux.

Plus récemment, dans un article de 2015, Raphaël CERF propose grâce à des méthodes de grandes déviations la démonstration de plusieurs résultats de convergence asymptotique très généraux, dans le cadre des algorithmes génétiques avec croisements, sélection par rang ou par tournois, et avec des fonctions d'évaluation arbitraires. La convergence asymptotique se caractérise alors par la présence du meilleur chromosome possible à l'infini, ou bien par la présence d'une grande proportion de chromosomes dominants. Ces résultats, certes rassurants et encourageants, restent cependant trop théoriques et généraux pour permettre en pratique de fixer des paramètres (forcément finis) de simulation, avec des fonctions d'évaluation données, pertinants et fidèles au cadre d'application de ces théorèmes.

Dans ce mémoire, notre but est de s'intéresser aux fonctions d'évaluation Route Royale, à l'instar de VON NIMWEGEN, P. CRUTCHFIELD et MITCHELL, mais en reprenant les méthodes de grandes déviations utilisées par Raphaël CERF dans l'espoir d'obtenir dans ce contexte des résultats à la fois plus simples et plus précis. Par souci de simplification, nous ne considérons ici que le cas le plus simple de la sélection par rang, et d'algorithmes sans croisements. D'autres examens pourraient cependant fournir des résultats fort intéressants par la suite, comme notamment le cas de la sélection proportionnelle à l'évaluation, ou l'étude complète avec croisements.

## 2 Présentation du modèle mathématique

Le modèle mathématique présenté ici est le même que celui considéré par Raphaël CERF dans son article. Nous le décrivons ici principalement pour rappeler ou redéfinir les notations.

## 2.1 L'algorithme sans croisements

Les algorithmes génétiques modélisent l'évolution d'une population de chromosomes au cours du temps. Chaque chromosome est constitué de plusieurs sites dont l'état est variable, et on modélise chaque état possible par une lettre ou un chiffre, comme A, C, G, T pour les séquences d'ADN. Ici, nous étudions le cas le plus simple à deux états possibles, 0 ou 1.

Nous fixons un entier  $m \geq 1$  représentant la taille de la population, et un paramètre  $p_M \in ]0, 1[$  de mutations. Le fonctionnement d'un algorithme génétique simple sans croisements peut alors être décrit de la façon suivante.

### Algorithme :

---

À la génération  $n = 0$ , on dispose d'une population initiale  $x_0$  de  $m$  chromosomes.

Puis, pour tout  $n \geq 0$ , la population à l'étape  $n + 1$ , notée  $x_{n+1}$ , s'obtient suite aux deux mécanismes suivants.

**Sélection :** On tire aléatoirement avec remplacement  $m$  chromosomes de la population  $x_n$  pour former une population intermédiaire  $y_n$ .

**Mutations :** Chaque site de chaque chromosome de la population  $y_n$  mute indépendamment avec probabilité  $p_M$  pour former la population  $x_{n+1}$ .

---

Pour les mutations, l'état d'un site change donc simplement de 1 à 0, ou de 0 à 1, le cas échéant. En revanche, le mécanisme de sélection est plus subtil, et passe par l'évaluation quantifiée des chromosomes pour être choisis ou non. Il est donc détaillé dans la section suivante.

## 2.2 Évaluation Route Royale et sélection par rang

L'évaluation des chromosomes permet de choisir préférentiellement les plus *efficaces* au moment de la sélection. Il existe de nombreuses façons d'évaluer les chromosomes, et de nombreuses façons de les sélectionner aussi. Le modèle que nous considérons est basé sur une évaluation célèbre, dite Route Royale, et la sélection non moins célèbre par rang, qui présente l'avantage d'être plus simple que la sélection proportionnelle, a priori.

**Évaluation Route Royale :** Nous notons  $l \geq 1$  le nombre de sites d'un chromosome. L'évaluation d'un chromosome revient alors à se donner une fonction d'évaluation<sup>3</sup>  $f : \{0, 1\}^l \rightarrow \mathbb{R}^+$ . Il est important de noter qu'une telle fonction possède un maximum, atteint en au moins un chromosome que l'on notera  $u^*$ , et que nous appelons la **Séquence Maître**. Nous pouvons supposer sans perte de généralité que  $u^* = 1 \dots 1$ .

Le principe de l'évaluation Route Royale est d'évaluer indépendamment chaque morceau d'un chromosome, et de faire une somme pondérée des résultats. Nous pouvons supposer ici, dans l'idée d'obtenir des résultats asymptotiques,

---

3. Majoritairement désignée par *fitness*, venant de l'anglais.

que l'entier  $l$  peut s'écrire  $l = NK$ , avec  $N \geq 1$  et  $K \geq 1$ , c'est-à-dire que l'on peut découper chaque chromosome en  $N$  blocs de même longueur  $K$  :

$$\underbrace{0001011101}_{K} \underbrace{1011110}_{K} \dots \dots \dots \underbrace{0001011101}_{K}$$

En désignant par  $u$  un chromosome quelconque, et  $u_i$  le  $i^{\text{ième}}$  bloc de longueur  $K$  du chromosome  $u$ , de sorte que  $u = u_1 \dots u_N$ , nous pouvons définir la fonction Route Royale  $f : \{0, 1\}^l \rightarrow \{0, 1, \dots, N\}$  la plus simple possible (c'est-à-dire en considérant une pondération identique sur chaque bloc) par :

$$f(u) = \sum_{i=1}^N 1_{\{u_i = u_i^*\}} = \sum_{i=1}^N 1_{\{u_i = 1 \dots 1\}}.$$

Concrètement, chaque terme intervient dans la somme si et seulement si le  $i^{\text{ième}}$  bloc du chromosome  $u$  correspond à celui de la Séquence Maître  $u^*$ , à savoir une séquence entière de 1 pour nous. Quand cela arrive, nous disons que le bloc  $u_i$  est donc **aligné** (sous-entendu sur le bloc  $u_i^* = 1 \dots 1$ ), et la fonction d'évaluation Route Royale compte simplement le nombre de blocs alignés.

**Sélection par rang** : Une fois la fonction d'évaluation choisie, la sélection par rang s'effectue en deux temps.

**Classement** : On ordonne les  $m$  chromosomes dans un ordre croissant d'efficacité, par rapport à la fonction  $f$ . Pour cela, on choisit uniformément au hasard une permutation des chromosomes parmi toutes celles qui conviennent, multiples en cas d'efficacité identiques.

**Tirage** : On réalise ensuite les  $m$  tirages avec remplacement en fonction du rang de chaque chromosome dans le classement précédent, en fixant pour cela deux paramètres  $\eta^+ \in \mathbb{R}$  et  $\eta^- \in \mathbb{R}$  tels que  $0 \leq \eta^- < \eta^+$  et  $\eta^+ + \eta^- = 2$ , et en définissant la probabilité  $b_m^i$  de tirer le  $i^{\text{ième}}$  (plus mauvais) chromosome par

$$b_m^i = \frac{1}{m} \left( \eta^- + (\eta^+ - \eta^-) \frac{i-1}{m-1} \right). \quad (1)$$

### 2.3 Définition des processus

Mathématiquement, nous modélisons l'algorithme en considérant l'espace  $(\{0, 1\}^l)^m$  de toutes les populations de taille  $m$  de chromosomes à  $l$  sites. Plus précisément, un chromosome  $u \in \{0, 1\}^l$  est modélisé par un vecteur ligne  $(u(1), \dots, u(l))$  de sites, et l'on note plus simplement  $u = u(1) \dots u(l)$ , comme nous l'avons déjà fait. Puis nous modélisons une population  $x \in (\{0, 1\}^l)^m$  par un vecteur colonne de chromosomes, soit par une matrice de  $\mathcal{M}_{m,l}(\{0, 1\})$ , dont les lignes sont les chromosomes.

$$x = \begin{pmatrix} x(1) \\ \vdots \\ x(m) \end{pmatrix} = \begin{pmatrix} x(1, 1) & \dots & x(1, l) \\ \vdots & \ddots & \vdots \\ x(m, 1) & \dots & x(m, l) \end{pmatrix}$$

Nous notons enfin  $x_0$  la population initiale dans l'algorithme. Ce dernier définit un processus  $X = (X_n)_{n \in \mathbb{N}}$  à valeurs dans  $(\{0, 1\}^l)^m$ , partant de  $X_0 = x_0$ . Ce processus est clairement une chaîne de Markov homogène. Il est par ailleurs aisé d'expliciter la matrice de transitions, mais celle-ci n'est pas utile pour notre étude.

Pour comprendre le comportement de l'algorithme, nous nous intéressons à la proportion de Séquences Maîtres à chaque instant. Et plus généralement, nous nous intéressons à la façon dont évolue le nombre de chromosomes qui possèdent la meilleure évaluation à chaque instant. Pour cela, nous définissons de nouveaux processus, pour tout  $\lambda \in \{0, \dots, N\}$ , par

$$\forall n \in \mathbb{N}, \quad M(X_n, \lambda) = \sum_{i=1}^m \mathbf{1}_{\{f(X_n(i))=\lambda\}}$$

Le processus  $M(X_n, \lambda)$  compte donc le nombre de chromosomes de  $X_n$  d'efficacité  $\lambda$ . Nous avons en particulier,  $\forall n \in \mathbb{N}$ ,  $\sum_{\lambda=0}^N M(X_n, \lambda) = m$ .

Or, les nouveaux processus  $(M(X_n, \lambda))_{n \in \mathbb{N}}$ , pour  $\lambda \in \{0, \dots, N\}$ , ne sont plus des chaînes de Markov car le nombre de chromosomes d'efficacité donnée à l'instant  $n + 1$  ne dépend pas seulement de leur nombre à l'instant  $n$ . En effet, leur nombre à l'instant  $n$  n'intervient que lors de la phase de sélection, et la phase de mutation peut créer de nouveaux chromosomes d'efficacité  $\lambda$ , avec des probabilités qui ne dépendent pas de seulement de  $n$  (comme dans le cas d'une chaîne de Markov inhomogène), mais qui dépendent de la configuration précise de la population à l'instant  $n$ . Dans la partie suivante, notre but sera donc de comparer ces processus avec des chaînes de Markov, sur lesquelles nous pourrions dire plus facilement des choses, afin de déduire des informations sur les processus d'origine.

### 3 Domination stochastique inférieure

Fixons  $\lambda \in \{0, \dots, N\}$  dans toute cette partie. Nous supposons de plus que, pour  $n \in \mathbb{N}$  fixé,  $\lambda = \max\{\varepsilon : M(X_n, \varepsilon) \neq 0\}$ . Nous cherchons alors à estimer les transitions  $\mathbb{P}(M(X_{n+1}, \lambda) = j \mid M(X_n, \lambda) = i)$  pour tout  $j \in \{0, \dots, m\}$  et tout  $i \in \{1, \dots, m\}$ , car l'hypothèse sur  $\lambda$  implique  $i \geq 1$ . Comme la population initiale  $x_0$  est quelconque, il suffit d'estimer les transitions précédentes pour  $n = 0$ .

#### 3.1 Minoration du processus

Dans ce mémoire, nous nous contenterons de minorer ces probabilités<sup>4</sup>. Comme nous avons supposé que  $\lambda$  correspond à la meilleure efficacité dans

---

4. Une étude plus complète consisterait à majorer également les probabilités de transition étudiées, pour justifier par la suite la convergence du processus d'origine vers la limite de cette borne inférieure, comme cela est fait dans les articles de référence.



la population de départ, il est impossible d'obtenir des chromosomes d'efficacité  $\lambda$  par mutation de chromosomes d'efficacité plus élevée. Notre idée est donc de négliger les chances d'obtenir des chromosomes d'efficacité  $\lambda$  par mutation de chromosomes d'efficacité moins élevée, et de ne considérer que l'étape de sélection pour produire des chromosomes d'efficacité  $\lambda$  à la génération suivante. En particulier, nous pouvons écrire, pour tout  $i \in \{1, \dots, m\}$  et tout  $j \in \{0, \dots, m\}$ ,

$$\begin{aligned} & \mathbb{P}(M(X_1, \lambda) = j \mid M(X_0, \lambda) = i) \\ & \geq \sum_{k=j}^m \mathbb{P}(\text{sélectionner exactement } k \text{ chromosomes d'efficacité } \lambda) \\ & \quad \times \mathbb{P}(\text{exactement } k - j \text{ des sélectionnés mutent sur un bloc aligné}). \end{aligned}$$

Pour la première probabilité dans la somme, d'après l'hypothèse faite sur  $\lambda$ , nous avons directement que :

$$\begin{aligned} & \mathbb{P}(\text{sélectionner exactement } k \text{ chromosomes d'efficacité } \lambda) \\ & = \mathbb{P}(\text{tirer exactement } k \text{ chromosomes dans les } i \text{ meilleurs}) \end{aligned}$$

Or, tirer exactement  $k$  chromosomes dans les  $i$  meilleurs (c'est-à-dire ceux qui ont un rang plus grand que  $i$  après classement) suit une loi binomiale, dont les deux paramètres sont  $m$ , et

$$b_m^{\geq i} = \sum_{r=m-i+1}^m b_m^r = \frac{i}{m} \left( \eta^- + \frac{(\eta^+ - \eta^-)}{m-1} \times \frac{(2m-i-1)}{2} \right)$$

d'après l'équation (1). Par commodité pour la suite, nous définissons tout de suite  $B_m(x) = x \left( \eta^- + \frac{(\eta^+ - \eta^-)(2-x-\frac{1}{m})}{2(1-\frac{1}{m})} \right)$ , de sorte que nous ayons  $b_m^{\geq i} = B_m\left(\frac{i}{m}\right)$ .

Par ailleurs, la deuxième probabilité suit de manière plus évidente une loi binomiale, dont les deux paramètres sont  $k$ , et  $c_\lambda$  dont l'expression est donnée par

$$\begin{aligned} c_\lambda & = \mathbb{P}(\text{muter sur au moins un bloc aligné}) \\ & = 1 - \mathbb{P}(\text{ne muter sur aucun des } \lambda \text{ blocs alignés}) \\ & = 1 - (1 - p_M)^{K\lambda} \end{aligned}$$

D'où finalement, en remplaçant dans l'inégalité de départ, nous obtenons

$$\begin{aligned} & \mathbb{P}(M(X_1, \lambda) = j \mid M(X_0, \lambda) = i) \\ & \geq \sum_{k=j}^m \binom{m}{k} B_m\left(\frac{i}{m}\right)^k \left(1 - B_m\left(\frac{i}{m}\right)\right)^{m-k} \times \binom{k}{k-j} c_\lambda^{k-j} (1 - c_\lambda)^j. \quad (2) \end{aligned}$$

Ainsi, nous avons ici une borne inférieure pour les probabilités de transition de notre processus  $M(X_0, \lambda)$ . De plus, cette borne est intéressante car nous pouvons l'interpréter comme la probabilité de transition d'un processus plus simple, que nous construisons dans la partie suivante.

## 3.2 Construction du couplage

### Couplage du processus initial

Dans cette partie, le but est de réaliser un couplage entre le processus de départ  $X$  et tous les paramètres de l'algorithme, et notamment pour la population initiale  $x_0$ . Nous ne supposons donc plus  $n = 0$  ici, car nous avons besoin de recréer tout le processus en entier. Ceci nous permettra ensuite d'obtenir une domination stochastique sur  $X$ , et d'en déduire une domination stochastique pour les processus dérivés  $M(X_n, \lambda)$ . Réexaminons pour cela toutes les étapes de l'algorithme une par une.

**Sélection :** Pour la sélection, nous avons besoin d'abord d'ordonner la population. Soient  $x$  une population et  $U$  un v.a. uniforme sur  $[0, 1]$ , nous définissons donc l'application  $\mathcal{O}$  qui au couple  $(x, U)$  associe une population ordonnée de la façon suivante :

1.  $\mathcal{O}$  énumère toutes les permutations de  $\{1, \dots, m\}$  qui ordonnent (dans l'ordre croissant relativement à  $f$ ) la population  $x$ , supposons qu'il y en ait  $I$ .
2.  $\mathcal{O}$  choisit la permutation numéro  $[UI]$ , et renvoie la population ordonnée. Soient maintenant  $y$  une population ordonnée (au sens précédent), et  $V$  une v.a. uniforme sur  $[0, 1]$ , le choix d'un nouveau chromosome peut s'obtenir en définissant l'application  $\mathcal{S}$ , qui pour un couple  $(y, V)$  sélectionne le chromosome  $y(\kappa)$ , avec  $\kappa = \min \left\{ t \in \{1, \dots, m\} : V < \sum_{i=1}^t b_m^i \right\}$ .

**Mutation :** Enfin, si  $u$  est un chromosome et  $W_1, \dots, W_l$  des v.a. uniformes sur  $[0, 1]$ , nous définissons l'application  $\mathcal{M}$  qui à un uplet  $(u, W_1, \dots, W_l)$  associe le chromosome  $v$  tel que  $v(i) = u(i)$  si  $W_i > p_M$ , et  $v(i) \neq u(i)$  sinon.

Autrement dit, comme l'algorithme doit sélectionner  $m$  chromosomes puis muter ces  $m$  chromosomes, et ce pour tout instant  $n \geq 1$ , nous nous donnons, pour tout  $n \geq 1$ , les familles de v.a. uniformes sur  $[0, 1]$  suivantes

- $U^n$ ,
- $V_1^n, \dots, V_m^n$ ,
- $(W_{k,1}^n)_{1 \leq k \leq l}, \dots, (W_{k,m}^n)_{1 \leq k \leq l}$ ,

puis définissons les matrices du couplage pour toute population  $x$ , par

$$\Phi_n(x) = \begin{pmatrix} \mathcal{M}(\mathcal{S}(\mathcal{O}(x, U^n), V_1^n), W_{1,1}^n, \dots, W_{l,1}^n) \\ \vdots \\ \mathcal{M}(\mathcal{S}(\mathcal{O}(x, U^n), V_m^n), W_{1,m}^n, \dots, W_{l,m}^n) \end{pmatrix}$$

Ceci permet donc bien de reconstruire tout le processus  $X$ , via les relations

$$\begin{cases} X_0 = x_0, \\ \forall n \geq 1, \quad X_n = \Phi_n(X_{n-1}). \end{cases}$$

### Couplage des processus de comptage

Maintenant que le processus initial est couplé avec les paramètres de l'algorithme, nous souhaitons définir un second processus  $Y$ , dont la loi du nombre de chromosomes d'efficacité  $\lambda$  est donnée par le processus  $M(Y_n, \lambda)$ , de sorte que les probabilités de transition coïncident avec l'expression obtenue dans l'inégalité (2). Il suffit alors que le processus  $Y$  soit conditionné à ne pas pouvoir créer de nouveaux blocs alignés par mutation, et donc de n'autoriser les mutations que sur les sites des blocs déjà alignés !

Pour le couplage, avec les mêmes familles de v.a. utilisées pour le processus  $X$ , nous définissons alors, pour tous  $n \geq 1$ ,  $i \in \{1, \dots, l\}$  et  $j \in \{1, \dots, m\}$ ,

$$\tilde{W}_{i,j}^n = \begin{cases} W_{i,j} & \text{si le bloc contenant le site } i \text{ (du chromosome } j) \text{ est aligné} \\ 1 & \text{sinon} \end{cases}$$

ainsi que les matrices de couplage pour le processus  $Y$  par

$$\Psi_n(x) = \begin{pmatrix} \mathcal{M}(\mathcal{S}(\mathcal{O}(x, U^n), V_1^n), \tilde{W}_{1,1}^n, \dots, \tilde{W}_{l,1}^n) \\ \vdots \\ \mathcal{M}(\mathcal{S}(\mathcal{O}(x, U^n), V_m^n), \tilde{W}_{1,m}^n, \dots, \tilde{W}_{l,m}^n) \end{pmatrix}.$$

### 3.3 Minoration stochastique

Le couplage réalisé dans la partie précédente permet donc de définir les processus  $X$  et  $Y$  sur un même espace de probabilité, et ce dans le but d'obtenir la domination stochastique  $X \geq Y$ . En particulier, cette domination permet d'obtenir la domination suivante pour les processus de comptage, à savoir :  $\forall n \in \mathbb{N}, M(X_n, \lambda) \geq M(Y_n, \lambda)$ . Cependant, cette déduction n'est pas si évidente qu'il n'y paraît, et repose sur une propriété de monotonie des chaînes considérées, celle-ci provenant elle-même de la construction du couplage précédent. Nous admettons cette domination dans ce mémoire; voir les articles de référence pour une justification analogue.

Dorénavant, nous nous intéresserons donc au processus  $M(Y_n, \lambda)$ , et simplifions la notation en posant  $M_n = M(Y_n, \lambda)$ . Il s'agit maintenant de comprendre comment le nombre  $M_n$  de chromosomes ayant la meilleure efficacité évolue pour le processus  $Y$ , et si nous obtenons une convergence vers une proportion non nulle de tels chromosomes, nous pourrions en déduire qu'une proportion au moins égale sera présente dans le processus initial  $X$ .

La construction du processus  $Y$  dans la section précédente a justement été

pensée pour que

$$\begin{aligned} & \mathbb{P}(M_{n+1} = j \mid M_n = i) \\ &= \sum_{k=j}^m \binom{m}{k} B_m \left( \frac{i}{m} \right)^k \left( 1 - B_m \left( \frac{i}{m} \right) \right)^{m-k} \times \binom{k}{k-j} c_\lambda^{k-j} (1 - c_\lambda)^j. \end{aligned} \quad (3)$$

C'est-à-dire pour que la domination stochastique admise dans cette section corresponde à l'inégalité (2).

## 4 Estimation du temps de survie

Comme le processus  $Y$  ne peut pas avoir de bloc aligné autrement que par sélection, le processus  $(M_n)_{n \in \mathbb{N}}$  est une chaîne de Markov, homogène, et de matrice de transition de la forme

$$Q = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \star_{1,0} & \dots & \dots & \star_{1,N} \\ \vdots & \ddots & \ddots & \vdots \\ \star_{N,0} & \dots & \dots & \star_{N,N} \end{pmatrix}$$

avec les  $\star_{i,j} > 0$ , indépendants de  $n$ , et dont les expressions sont données par la formule (3). Notons au passage que 0 est le seul état absorbant, car tous les blocs alignés peuvent être détruits par mutations, mais les mutations ne créent jamais de bloc aligné. Pour commencer, nous nous intéressons donc au temps que le processus  $(M_n)_{n \in \mathbb{N}}$  va mettre avant de tomber dans son état absorbant. Nous définissons pour cela  $\tau_\lambda = \min\{n \in \mathbb{N} : M_n = 0\}$ , et nous cherchons donc à estimer  $\mathbb{E}[\tau_\lambda]$ . Les méthodes standards de grandes déviations sont alors d'un grand secours.

### 4.1 Première inégalité de grande déviation

Fixons un ensemble  $\Gamma \subset [0, 1]$  quelconque, la première quantité que nous évaluerons est, pour tout  $i \in \{0, \dots, m\}$ ,  $\mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = i)$ . Notons alors

$$p_\lambda(i, j, k) = \binom{m}{k} B_m \left( \frac{i}{m} \right)^k \left( 1 - B_m \left( \frac{i}{m} \right) \right)^{m-k} \times \binom{k}{k-j} c_\lambda^{k-j} (1 - c_\lambda)^j.$$

Pour tout  $i \in \{0, \dots, m\}$ , nous avons d'après l'équation (3)

$$\begin{aligned} \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = i) &= \sum_{j \in m\Gamma \cap \mathbb{N}} \mathbb{P}(M_{n+1} = j \mid M_n = i) \\ &\leq (m+1)^2 \max\{p_\lambda(i, j, k) : k \in \{0, \dots, m\}, j \in m\Gamma \cap \mathbb{N}, j \leq k\}. \end{aligned}$$

Or, la connaissance des méthodes classiques de grandes déviations permet de voir qu'il est possible d'évaluer ce maximum en estimant la quantité

$$\begin{aligned}
\ln p_\lambda(i, j, k) &= \ln \frac{m!}{k!(m-k)!} + k \ln B_m\left(\frac{i}{m}\right) + (m-k) \ln \left(1 - B_m\left(\frac{i}{m}\right)\right) \\
&\quad + \ln \frac{k!}{(k-j)!j!} + (k-j) \ln c_\lambda + j \ln(1 - c_\lambda) \\
&= -m \left( \frac{k}{m} \ln \frac{1}{B_m\left(\frac{i}{m}\right)} + \frac{m-k}{m} \ln \frac{1}{1 - B_m\left(\frac{i}{m}\right)} \right) + \ln \frac{m!}{k!(m-k)!} \\
&\quad - k \left( \frac{k-j}{k} \ln \frac{1}{c_\lambda} + \frac{j}{k} \ln \frac{1}{1 - c_\lambda} \right) + \ln \frac{k!}{(k-j)!j!} \\
&= -m \left( \frac{k}{m} \ln \frac{m}{kB_m\left(\frac{i}{m}\right)} + \frac{m-k}{m} \ln \frac{m}{(1 - B_m\left(\frac{i}{m}\right))(m-k)} \right) \\
&\quad + \left( \ln \frac{m!}{k!(m-k)!} + k \ln \frac{k}{m} + (m-k) \ln \frac{m-k}{m} \right) \\
&\quad - k \left( \frac{k-j}{k} \ln \frac{k}{(k-j)c_\lambda} + \frac{j}{k} \ln \frac{k}{j(1 - c_\lambda)} \right) \\
&\quad + \left( \ln \frac{k!}{(k-j)!j!} + (k-j) \ln \frac{k-j}{k} + j \ln \frac{j}{k} \right) \\
&= -mI\left(B_m\left(\frac{i}{m}\right), \frac{k}{m}\right) + G(k, m) - kI\left(c_\lambda, \frac{k-j}{k}\right) + G(j, k)
\end{aligned}$$

en reconnaissant la fonction  $I$  de grande déviation pour une loi binomiale de paramètres  $n$  et  $p$ , définie pour tous  $p \in [0, 1]$  et  $t \in \mathbb{R}^+$  par

$$I(p, t) = \begin{cases} 0 & \text{si } t = p, \\ t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{1-p} & \text{si } 0 < p < 1 \text{ et } 0 \leq t \leq 1 \text{ avec } t \neq p, \\ +\infty & \text{si } (p \in \{0, 1\} \text{ et } t \neq p) \text{ ou } t > 1. \end{cases}$$

et en posant, pour tous  $n \geq 1$  et  $k \leq n$ ,

$$G(k, n) = \ln \frac{n!}{k!(n-k)!} + (n-k) \ln \frac{n-k}{n} + k \ln \frac{k}{n}.$$

Un lemme déterministe<sup>5</sup> permet alors d'obtenir rapidement que, pour tous  $n \geq 1$  et  $k \leq n$ ,  $G(k, n) \leq 2 \ln n + 3$ . Et comme les inégalités  $k \leq m$  et  $j \leq k \leq m$  sont toujours vraies, nous déduisons

$$\ln p_\lambda(i, j, k) \leq -m \left( I\left(B_m\left(\frac{i}{m}\right), \frac{k}{m}\right) + \frac{k}{m} I\left(c_\lambda, 1 - \frac{j}{k}\right) \right) + 4 \ln m + 6.$$

---

5. Lemme classique de grandes déviations, dont la démonstration sera donnée en annexe.

Ce calcul nous permet de voir venir une fonction exploitable des variables discrètes  $\frac{i}{m}$ ,  $\frac{k}{m}$  et  $\frac{j}{m}$  pour déterminer un maximum, à la difficulté près que le terme  $B_m\left(\frac{i}{m}\right)$  dépend de  $m$  autrement qu'à travers la variable  $\frac{i}{m}$ . Dans ce calcul, la variable  $i$  étant choisie comme on le souhaite (avec  $i \neq 0$  puisque 0 est un état absorbant), nous pouvons fixer  $s \in ]0, 1]$  de sorte que pour tout  $m \geq 1$ , nous choisissons de mener le calcul avec  $i = \lfloor sm \rfloor$ . De cette façon, nous pouvons chercher un équivalent de  $B_m\left(\frac{i}{m}\right)$  quand  $m$  tend vers l'infini, toujours dans le but d'obtenir les résultats asymptotiques souhaités. Nous avons alors

$$\begin{aligned} B_m\left(\frac{i}{m}\right) &= s \left( \eta^- + \frac{\eta^+ - \eta^-}{2} \left( 2 - s - \frac{1}{m} \right) (1 + o(1)) \right) \\ &= s \left( \eta^- + \frac{\eta^+ - \eta^-}{2} (2 - s) \right) + o(1). \end{aligned}$$

On note  $f(s) = s \left( \eta^- + \frac{\eta^+ - \eta^-}{2} (2 - s) \right) = s \left( \eta^+ - s \frac{\eta^+ - \eta^-}{2} \right)$ . Comme  $f$  est croissante sur  $]0, 1]$ , nous avons bien  $0 < f(s) \leq \frac{\eta^+}{2} < 1$ , ce qui nous permet de réaliser l'estimation suivante

$$\begin{aligned} I\left(B_m\left(\frac{i}{m}\right), \frac{k}{m}\right) &= \frac{k}{m} \left( \ln\left(\frac{k}{m}\right) + \ln\left(\frac{1}{B_m\left(\frac{i}{m}\right)}\right) \right) \\ &\quad + \left(1 - \frac{k}{m}\right) \left( \ln\left(1 - \frac{k}{m}\right) + \ln\left(\frac{1}{1 - B_m\left(\frac{i}{m}\right)}\right) \right) \\ &= \frac{k}{m} \left( \ln\left(\frac{k}{m}\right) + \ln\left(\frac{1}{f(s)}\right) + \ln\left(\frac{1}{1 - o\left(\frac{1}{f(s)}\right)}\right) \right) \\ &\quad + \left(1 - \frac{k}{m}\right) \left( \ln\left(1 - \frac{k}{m}\right) + \ln\left(\frac{1}{1 - f(s)}\right) + \ln\left(\frac{1}{1 - o\left(\frac{1}{1 - f(s)}\right)}\right) \right) \\ &= I\left(f(s), \frac{k}{m}\right) + o(1). \end{aligned}$$

Conclusion, pour  $m$  assez grand,

$$\frac{1}{m} \ln p_\lambda(\lfloor sm \rfloor, j, k) \leq -I\left(f(s), \frac{k}{m}\right) - \frac{k}{m} I\left(c_\lambda, 1 - \frac{j}{m}\right) + o(1).$$

Ou bien asymptotiquement, pour tout  $s \in ]0, 1]$  et tout  $\Gamma \subset [0, 1]$ ,

$$\begin{aligned} & \frac{1}{m} \ln \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \\ & \leq -\min \left\{ J \left( \frac{\lfloor sm \rfloor}{m}, \frac{j}{m}, \frac{k}{m} \right) : j \in m\Gamma \cap \mathbb{N}, k \in \{j, \dots, m\} \right\} \quad (4) \end{aligned}$$

en posant  $J(s, t, u) = I(f(s), u) + uI \left( c_\lambda, 1 - \frac{t}{u} \right)$ .

## 4.2 Étude de la borne

Afin de mieux comprendre le comportement du minimum obtenu dans la borne, la technique standard de grandes déviations et de troquer ce dernier, qui porte sur des ensembles finis pour chaque  $m$  mais dont la convergence est difficile à cerner, contre un infimum sur un domaine plus large, dont l'existence reste à prouver mais qui devient indépendant de  $m$ . Par ailleurs, cet infimum minorera inévitablement chaque minimum obtenu à  $m$  fixé, donc le sens de l'inégalité sera toujours valable.

Pour cela, l'idée est de minorer sur les  $t$  dans  $\Gamma$  et les  $u$  dans  $[t, 1]$ , en imaginant donc que  $j = \lfloor tm \rfloor$  et  $k = \lfloor um \rfloor$ . On observe alors (par la propriété de semi-continuité inférieure de la fonction  $I$ ) que  $\overline{\lim} J \left( \frac{\lfloor sm \rfloor}{m}, \frac{j}{m}, \frac{k}{m} \right) \geq J(s, t, u)$ . Ce qui donne par l'inéquation (4) la vraie inégalité de grandes déviations recherchée, à savoir :

$$\begin{aligned} & \overline{\lim} \frac{1}{m} \ln \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \\ & \leq \overline{\lim} -\min \left\{ J \left( \frac{\lfloor sm \rfloor}{m}, \frac{k}{m}, \frac{j}{m} \right) : j \in m\Gamma \cap \mathbb{N}, k \in \{j, \dots, m\} \right\} \\ & \leq -\inf \{ J(s, t, u) : t \in \Gamma, u \in [t, 1] \}. \end{aligned}$$

Or, nous allons montrer que cette borne majorante est en fait la bonne, au sens où elle est optimale. Il nous faut pour cela reprendre le calcul au départ, mais en cherchant cette fois une minoration. Nous pouvons alors écrire, pour tout  $i \in \{1, \dots, m\}$  et tout  $\Gamma \subset [0, 1]$ , plus simplement :

$$\begin{aligned} \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = i) &= \sum_{j \in m\Gamma \cap \mathbb{N}} \mathbb{P}(M_{n+1} = j \mid M_n = i) \\ &\geq \max \{ p_\lambda(i, j, k) : j \in m\Gamma \cap \mathbb{N}, k \in \{j, \dots, m\} \}. \end{aligned}$$

En s'inspirant du travail précédent, nous fixons maintenant, pour chaque  $m$ , un  $s \in ]0, 1]$ , un  $t \in \overset{\circ}{\Gamma}$  (il suffit de prendre  $\Gamma$  de sorte que cela soit possible), et un  $u \in [t, 1]$  tels que  $i = \lfloor sm \rfloor$ ,  $j = \lfloor tm \rfloor$  et  $k = \lfloor um \rfloor$ , où  $i, j$  et  $k$  réalisent le maximum obtenu ci-dessus. Notons alors que le choix de  $t \in \overset{\circ}{\Gamma}$  garantit que pour  $m$  assez grand, chaque  $\frac{j}{m}$  appartient à  $\overset{\circ}{\Gamma}$ .

Par ailleurs, avec les mêmes notations que dans la partie précédente, nous avons directement que

$$\begin{aligned} \frac{1}{m} \ln p_\lambda(i, j, k) &\geq -I\left(B_m\left(\frac{i}{m}\right), \frac{k}{m}\right) - \frac{k}{m} I\left(c_\lambda, 1 - \frac{j}{k}\right) \\ &\geq -I\left(f\left(\frac{\lfloor sm \rfloor}{m}\right), \frac{k}{m}\right) - \frac{k}{m} I\left(c_\lambda, 1 - \frac{j}{k}\right) + o(1). \end{aligned}$$

Et alors

$$\begin{aligned} \underline{\lim} \frac{1}{m} \ln \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \\ \geq \underline{\lim} \max \left\{ J\left(\frac{\lfloor sm \rfloor}{m}, \frac{\lfloor tm \rfloor}{m}, \frac{\lfloor um \rfloor}{m}\right) : j \in m\Gamma \cap \mathbb{N}, k \in \{j, \dots, m\} \right\} \\ \geq -\inf\{J(s, t, u) : t \in \overset{\circ}{\Gamma}, u \in [t, 1]\}. \end{aligned}$$

Cette borne diffère légèrement de celle obtenue précédemment, mais l'infimum sur l'intérieur de  $\Gamma$  est toujours plus grand que celui sur  $\Gamma$ . Donc les deux inégalités de grandes déviations démontrent l'existence de la limite qui suit.

$$\lim_{m \rightarrow \infty} \frac{1}{m} \ln \mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = \lfloor sm \rfloor) = -\inf\{J(s, t, u) : t \in \Gamma, u \in [t, 1]\} \quad (5)$$

Nous noterons  $J_{inf}(s, \Gamma) = -\inf\{J(s, t, u) : t \in \Gamma, u \in [t, 1]\}$ , ce qui présente l'avantage de mieux voir la dépendance en  $s$  et  $\Gamma$  de cette borne, et de simplifier les écritures à venir.

### 4.3 Dynamique du processus

L'équation (5) peut se réécrire de manière plus explicite comme suit :

$$\mathbb{P}(M_{n+1} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \underset{m \rightarrow \infty}{\sim} \exp(-m \inf\{J(s, t, u) : t \in \Gamma, u \in [t, 1]\})$$

Cette formule nous permet donc d'identifier les événements qui ont une chance non nulle de se produire asymptotiquement, à savoir ceux pour lesquels  $\inf\{J(s, t, u) : t \in \Gamma, u \in [t, 1]\} = 0$ . Le choix de  $s$  et de  $\Gamma$  importe à ce niveau là. D'après la définition de  $I$ , la condition de nullité précédente équivaut à

$$u = f(s) \quad \text{et} \quad c_\lambda = 1 - \frac{t}{u},$$

c'est-à-dire, vu que la variable  $s$  (associée à  $i$ ) représente moralement le point de départ, et  $t$  (associée à  $j$ ) celui d'arrivée, en exprimant  $t$  en fonction de  $s$

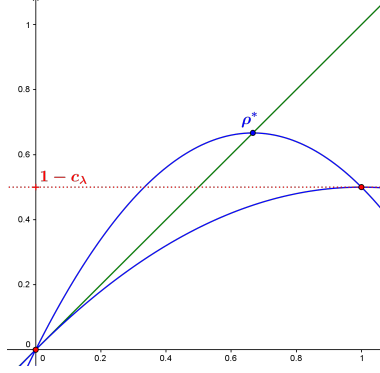
$$t = u(1 - c_\lambda) = f(s)(1 - c_\lambda).$$

Nous posons alors  $F(s) = s \left( \eta^+ - s \frac{\eta^+ - \eta^-}{2} \right) (1 - c_\lambda)$  pour  $s \in ]0, 1]$ . Il est ainsi possible d'interpréter la dynamique de la fonction  $F$  comme étant le



comportement « typique » d’une transition du processus, et les itérations de cette fonction donnent le comportement global (i.e. sur plusieurs itérations) du processus. Intéressons-nous donc aux points fixes de  $F$ .

$F$  est un polynôme de degré 2 en  $s$ . Comme  $F(0) = 0$  (par prolongement) et  $F(1) = 1 - c_\lambda < 1$ ,  $F$  possède un point fixe sur  $\rho^* \in ]0, 1]$  si et seulement si  $F'(0) > 1$ , i.e.  $\eta^+(1 - c_\lambda) > 1$ .



Notons que si  $F'(0) \leq 1$ , le processus converge exclusivement vers 0, et il n’y a donc aucune chance d’avoir un grand nombre de chromosomes dominants dans la population finale avec cet algorithme. Nous devons donc imposer  $\eta^+(1 - c_\lambda) > 1$  pour la suite.

Rappelons que  $c_\lambda = 1 - (1 - p_M)^{\lambda K}$ , et que dans notre étude, la longueur des chromosomes  $l = NK$  tend vers l’infini avec  $K$ . Si la probabilité de mutation reste constante, les blocs alignés à partir d’un certain temps n’auraient aucune chance de le rester. Il est donc bon de faire décroître  $p_M$  avec  $\frac{1}{K}$ . Dans le but de déterminer un régime (et pas forcément tous) où l’algorithme est capable de converger comme expliqué en introduction, nous choisissons de poser  $p_M = \frac{a}{K}$ , avec  $a > 0$  dont nous gardons la possibilité d’ajuster la valeur si besoin. Sous ces hypothèses, nous avons donc  $c_\lambda \xrightarrow{K \rightarrow \infty} 1 - \exp(-a\lambda)$ .

La condition  $F'(0) > 1$  peut se traduire asymptotiquement par  $a < \frac{\ln \eta^+}{\lambda}$ . Cette dernière inégalité laisse par ailleurs penser que  $p_M$  doit également varier au cours de l’algorithme en fonction de l’efficacité du meilleur chromosome présent dans la population : plus l’algorithme aura découvert de blocs alignés chez un même individu ( $\lambda$  grand), plus la probabilité de mutation doit diminuer pour les conserver (en  $\frac{1}{\lambda}$ )... Par ailleurs, comme  $\lambda \leq N$ , il est également possible de fixer la valeur de  $p_M$  dès le début de l’algorithme en choisissant par exemple  $a < \frac{\ln \eta^+}{N}$ , mais cette option ralentit certainement l’efficacité de l’algorithme au démarrage (avec trop peu de mutations pour découvrir rapidement de nouveaux blocs alignés).

#### 4.4 Temps moyen de survie

Rappelons que notre but dans cette section est de déterminer  $\mathbb{E}[\tau_\lambda | M_0 = i]$ , où  $\tau_\lambda = \min\{n \in \mathbb{N} : M_n = 0\}$  et  $i \neq 0$ . Là encore nous réinvestissons les méthodes de grandes déviations usuelles pour y parvenir.

La première étape, ici, va être de généraliser l'équation (5) pour connaître la probabilité de transition sur un nombre quelconque de générations.

**Lemme 1** *Soit  $d \geq 1$ , toujours avec les notations précédentes, nous avons*

$$\begin{aligned} \lim_{m \rightarrow \infty} \frac{1}{m} \ln \mathbb{P}(M_{n+d} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \\ = - \inf \left\{ \sum_{r=0}^{d-1} J_{inf}(s_r, \{s_{r+1}\}) : s_1, \dots, s_{d-1} \in ]0, 1], s_0 = s, s_d \in \Gamma \right\}. \end{aligned}$$

**Preuve 1**

$$\begin{aligned} & \mathbb{P}(M_{n+d} \in m\Gamma \mid M_n = \lfloor sm \rfloor) \\ &= \sum_{j_1, \dots, j_{d-1}} \mathbb{P}(M_{n+d} \in m\Gamma, M_{n+d-1} = j_{d-1}, \dots, M_{n+1} = j_1 \mid M_n = \lfloor sm \rfloor) \\ &= \sum_{j_1, \dots, j_{d-1}} \mathbb{P}(M_{n+d} \in m\Gamma \mid M_{n+d-1} = j_{d-1}) \times \dots \times \mathbb{P}(M_{n+1} = j_1 \mid M_n = \lfloor sm \rfloor) \\ &\leq (d-1) \max \left\{ \mathbb{P}(M_{n+d} \in m\Gamma \mid M_{n+d-1} = j_{d-1}) \times \dots \times \mathbb{P}(M_{n+1} = j_1 \mid M_n = \lfloor sm \rfloor) \right. \\ & \quad \left. : j_1, \dots, j_{d-1} \right\} \end{aligned}$$

*En passant au ln, divisant par m et en prenant la limite, nous obtenons bien la somme de  $d-1$  infimums donnés par l'équation (5) (la probabilité de transitions étant indépendante de n), ce qui est égal à l'infimum de la somme.*

*Réciproquement, plutôt que de le majorer par  $d-1$  fois le maximum, l'avant-dernier terme se minore plus simplement par 1 fois le maximum, et les mêmes opérations aboutissent à la même borne.*

Par souci de simplification des notations, nous notons plus généralement  $J_{inf}^d(s, \Gamma) = - \inf \left\{ \sum_{r=0}^{d-1} J_{inf}(s_r, \{s_{r+1}\}) : s_1, \dots, s_{d-1} \in ]0, 1], s_0 = s, s_d \in \Gamma \right\}$ . Nous avons en effet  $J_{inf}(s, \Gamma) = J_{inf}^1(s, \Gamma)$ .