Université Paris Nanterre
200 av. de la République
92001 Nanterre Cedex
www.parisnanterre.fr

Membre de l'université Paris Lumière

École doctorale 139 : Connaissance, Langage, Modélisation
Laboratoire Modal'X

Nathan Noiry

Matrices aléatoires et graphes aléatoires

Thèse présentée et soutenue publiquement le 08/10/2020
en vue de l'obtention du doctorat de **Mathématiques** de l'Université Paris Nanterre
sous la direction de **Nathanaël ENRIQUEZ**
et de **Laurent MÉNARD**
Jury :

Examinatrice	**Mme Anne-Laure Basdevant**	Maître de conférence, Université Paris Nanterre
Président de jury	**M. Francis Comets**	Professeur, Université Paris Diderot
Directeur de thèse	**M. Nathanaël Enriquez**	Professeur, Université Paris-Saclay
Examinateur	**M. Maxime Février**	Maître de conférence, Université Paris-Saclay
Rapporteure	**Mme Christina Goldschmidt**	Professor, University of Oxford
Directeur de thèse	**M. Laurent Ménard**	Maître de conférence, Université Paris Nanterre
Rapporteur	**M. Justin Salez**	Professeur, Université Paris Dauphine

À mes parents,

Remerciements

Mes premiers remerciements s'adressent bien sûr à mes deux directeurs de thèse, Nathanaël Enriquez et Laurent Ménard, dont les conseils avisés me guident depuis mon stage de première année à l'École Normale Supérieure de Lyon. Leur disponibilité, la profondeur de leurs idées et leur optimisme à toute épreuve ont créé les conditions idéales pour mes premiers pas dans le monde de la recherche. Nathanaël et Laurent partagent une vision tout en finesse des mathématiques qui n'a cessé de me séduire, et le souvenir de ces quatre années passées à leur côté restera impérissable. Je ne serai jamais assez reconnaissant de leur engagement de *tous* les instants.

Je voudrais ensuite remercier chaleureusement Christina Goldschmidt et Justin Salez, qui ont accepté la lourde tâche d'être les rapporteurs de cette thèse. Je leur adresse toute ma gratitude pour le temps qu'ils ont accordé à la relecture minutieuse du manuscrit, tâche délicate que le contexte de l'année deux mille vingt a certainement compliqué. Ces deux chercheurs sont un exemple pour moi, et je suis particulièrement honoré de l'intérêt qu'ils ont porté à mon travail.

J'aimerais également adresser toute ma reconnaissance à Anne-Laure Basdevant, Francis Comets et Maxime Février, qui me font l'honneur d'avoir accepté de participer à mon jury.

Maxime Février a été un acteur essentiel de cette thèse : je souhaite le remercier pour le temps qu'il m'a accordé et pour ses explications limpides, qui m'ont permis de me familiariser avec les matrices aléatoires. Je suis par ailleurs très heureux de l'intérêt qu'Alain Rouault a porté à mon travail, et je voudrais le remercier de m'avoir donné l'occasion de collaborer avec lui. J'aimerais aussi adresser un profond remerciement à Laure Dumaz, pour nos discussions mathématiques, bien sûr, mais aussi pour l'écoute amicale qu'elle m'a toujours offerte. J'ai notamment eu la chance d'organiser à ses côtés le séminaire mensuel du groupe de recherche MEGA – Matrices Et Graphes Aléatoires. Ceci m'amène naturellement à remercier Jamal Najim, pour sa confiance, ainsi que les autres organisateurs du séminaire : Guillaume Barraquand, Raphaël Butez et Camille Male. Je voudrais enfin remercier Justin Salez, dont le regard bienveillant m'accompagne depuis le début de la thèse.

Merci à tous les membres de Modal'X, grâce à qui règne une ambiance chaleureuse dans le laboratoire. Je pense particulièrement à Gabriel Faraud, avec qui j'ai eu la chance de collaborer, et à Patrice Bertail, pour nos nombreux échanges : j'espère que nos projets se concrétiseront dans l'avenir.

Au cours de ces quatre dernières années, j'ai eu le privilège de partager le bureau E08 avec de nombreux doctorants, post-doctorants et ATER : merci à El Mehdi, Taron, Charles, Paul, Julien, Boutheina, Nathalie, Samuel, Aurélie, Oussama, Christopher. Un merci tout particulier à Émilien : travailler à ses côtés fût un réel plaisir. Au-delà du laboratoire, j'ai eu la chance de croiser le chemin de nombreux doctorants. Je pense notamment à Clément, Simon, Assaf, Lucas, Armand, Thibault, Hugo, Yizhe, Slim, Delphin. J'ai également une pensée amicale pour tous mes camarades de Lyon, en particulier Mickaël, Robin, Ziad, Matthieu, et bien sûr Victor.

Au risque de rallonger ces remerciements, je voudrais profiter de cette occasion pour souligner l'influence déterminante qu'on eut certains enseignants sur mon parcours. C'est Olivier Coutarel qui, le premier, me détermina à réaliser des études de mathématiques. L'investissement d'Olivier Danthon m'a ensuite permis de découvrir l'exigence et l'esthétique de cette discipline ; je lui suis extrêmement reconnaissant de m'avoir accordé autant de temps. Sans Stéphane Attal et Jérôme Germoni, je ne serais probablement pas en train d'écrire ces lignes. Stéphane Attal m'a donné les moyens de réaliser mon projet, et ses conseils, si précieux à mes yeux, m'ont permis de surmonter l'angoisse du concours. Quant à Jérôme Germoni, il m'a offert l'opportunité de

réaliser un TIPE passionnant : j'ai beaucoup appris à ses côtés, et garderai un agréable souvenir de nos nombreuses discussions. Plus récemment, les travaux dirigés de Loïc Richier et Sébastien Martineau ont confirmé mon goût pour les probabilités. Je tiens à remercier Sébastien pour son enthousiasme permanent, si rare et si nécessaire.

Mes derniers remerciements s'adressent tout naturellement à mes proches. Mes amis, d'abord, pour leur soutien constant. En simplifiant : les *Villeurbannais* Sofiane, Sylvain, Thomas, Quentin, Louise, Déborah, Victoria, Romain, Éléonore, Joris, Marie, Clémence, Cyrielle ; et les *Montferrandais* Gaëtan, Antoine, Lisa, Laura, Marine, Caroline, Joséphine. Mes années de thèse n'auraient pas eu la même saveur sans eux ; ni sans les repas traditionnels du dimanche partagés avec Juliet et Antoine, d'ailleurs. Je voudrais ensuite remercier Anne-Marie et Jean-Pierre, dont les encouragements m'ont toujours touché ; certaines idées de cette thèse sont nées à Beaulieu, ce qui m'impose sans doute de remercier le Cantal ? Merci à ma mamie et à mon papi, pour leur présence rassurante et bienveillante. Hugo, papa, maman, merci de votre soutien inconditionnel qui m'accompagne depuis toutes ces années. Papa, maman, cette thèse vous est dédiée, merci de m'avoir donné le goût du travail et laissé libre de choisir ; votre fierté est aussi la mienne. Les derniers mots sont pour Inès, dont la présence quotidienne restera le moteur essentiel de cette thèse.

Résumé

Cette thèse se compose de plusieurs travaux ayant trait à la théorie des matrices aléatoires et à la théorie des graphes aléatoires.
Dans le contexte des matrices aléatoires, un premier travail porte sur l'étude spectrale des matrices de Wishart dont la taille tend vers l'infini et dont les moments des coefficients explosent. Dans ce cadre, nous calculons un développement asymptotique de la limite des mesures spectrales empiriques au voisinage de la loi de Marchenko-Pastur. Dans un second travail, nous nous sommes intéressés aux modèles matriciels déformés. Nous démontrons que l'étude des mesures spectrales dans la direction des vecteurs propres des matrices de perturbation apporte de nombreuses informations sur le spectre de ces modèles, notamment sur les coordonnées des vecteurs propres. Enfin, dans un troisième travail, nous exploitons un outil classique de la théorie des matrices aléatoires – la transformée de Stieltjes – afin d'identifier une classe soluble de processus de renouvellement.
Les deux autres contributions de cette thèse concernent la géométrie des modèles de configuration, (multi)-graphes aléatoires dont la suite des degrés est décidée à l'avance. Dans le régime sur-critique, nous nous sommes intéressés à l'analyse de l'algorithme de parcours en profondeur et à l'une de ses variantes, alternant entre parcours en profondeur et parcours en largeur. Nous démontrons qu'après une mise à l'échelle adéquate, les processus de contour associés à ces algorithmes convergent vers des profils déterministes, établissant en particulier l'existence de chemins simples de longueur linéaire, et l'existence de cycles de longueur linéaire ne possédant pas de raccourci à courte portée.

Mots clés. Matrices aléatoires, graphes aléatoires, modèle de configuration, processus de renouvellement, transformée de Stieltjes.

Abstract

This thesis consists of several works related to the theory of random matrices and the theory of random graphs.
In the context of random matrices, a first work concerns the spectrum of Wishart matrices whose size tends to infinity and whose entries have exploding moments. In this setting, we compute an asymptotic expansion of the limit of the empirical spectral measures in the vicinity of the Marchenko-Pastur law. In a second work, we were interested in deformed matrix models. We prove that the study of spectral measures in the direction of the eigenvectors of the perturbation matrices brings a lot of information on the spectrum of these models, in particular on the coordinates of the eigenvectors. Finally, in a third work, we exploit a classical tool of random matrix theory – the Stieltjes transform – in order to identify a solvable class of renewal processes.
The two other contributions of this thesis concern the geometry of configuration models, which are (multi)-random graphs whose sequence of degrees is fixed. In the supercritical regime, we study the depth first search algorithm and one of its variants, which alternates between depth first and breadth first search. We prove the convergence of the renormalized contour processes associated with these algorithms to deterministic profiles, establishing in particular the existence of simple paths of linear length, and the existence of cycles of linear length without shortcut at short range.

Keywords. Random matrices, random graphs, configuration model, renewal processes, Stieltjes transform.

Table des matières

Première partie

Introduction

Chapitre 1

Matrices aléatoires

Les matrices aléatoires apparaissent pour la première fois dans la littérature en 1929 avec les travaux de J. Wishart qui proposa, dans un contexte d'inférence statistique, l'étude de matrices de variance-covariance de vecteurs gaussiens indépendants et de même loi. On peut cependant assimiler la véritable naissance de la théorie des matrices aléatoires aux célèbres travaux du physicien E. Wigner, dans les années cinquante. L'idée fondatrice et particulièrement féconde de Wigner fût d'approximer les niveaux d'énergies de certains atomes lourds par les valeurs propres de grandes matrices aléatoires symétriques. Celles-ci sont alors envisagées comme des approximations de l'Hamiltonien du système physique étudié, qui est un opérateur de dimension infinie. Cette démarche a été confirmée par de nombreuses expérimentations et a révélé la structure remarquable des valeurs propres des matrices aléatoires, motivant ainsi leur étude approfondie par de nombreuses générations de mathématiciens.

Dans cette première partie de ce chapitre, nous présentons les résultats principaux concernant deux modèles de matrices aléatoires symétriques : le modèle historique de Wigner et celui de Wishart. Deux généralisations de ces modèles sont ensuite discutées. La première s'intéresse à des matrices aléatoires symétriques dont la loi des coefficients dépend de la dimension ; elle est en lien avec le Chapitre 3 de cette thèse, issu de l'article [Noi18]. La seconde concerne les modèles matriciels déformés, qui font l'objet du Chapitre 4 issu de l'article [Noi20]. Dans une dernière partie, nous présentons les résultats du Chapitre 5, issu d'un article réalisé en collaboration avec Nathanaël Enriquez [EN20], où nous avons identifié une classe soluble de processus de renouvellement et discuté de ses applications à un modèle de physique statistique. Si son objet d'étude diffère de celui des matrices aléatoires, ce travail se fonde sur un lien entre un modèle de polymères et un modèle de matrices aléatoires déformées, ce qui justifie à nos yeux sa présence dans cette partie de l'introduction.

1.1 Le modèle de Wigner

Comportement global du spectre. Pour tout $n \geq 1$, on considère une matrice aléatoire symétrique $X_n \in \mathbf{R}^{n\times n}$ de taille $n \times n$ dont les coefficients sont, à la symétrie près, des variables aléatoires indépendantes et identiquement distribuées. On supposera de plus que la variance de ces variables est égale à 1. Dans ce contexte, on appellera matrice de Wigner de taille n la matrice aléatoire

$$W_n := \frac{1}{\sqrt{n}} X_n.$$

Les valeurs propres de W_n seront notées $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$. Il est commode de résumer leur information à l'aide de la *mesure spectrale empirique*

$$\mu_{W_n} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i^{(n)}}.$$

C'est une variable aléatoire dans l'espace des mesures de probabilités dont le comportement asymptotique fait l'objet du Théorème de Wigner. Celui-ci établit la convergence des mesures spectrales empiriques vers la loi du demi-cercle, définie par

$$\mu_{sc}(\mathrm{d}x) := \frac{\sqrt{4-x^2}}{2\pi}\mathbf{1}_{|x|\leq 2}\mathrm{d}x.$$

Théorème 1. *Presque sûrement, la mesure μ_{W_n} converge étroitement vers μ_{sc}. Autrement dit, pour toute fonction $f : \mathbf{R} \mapsto \mathbf{R}$ continue et bornée, la convergence suivante a lieu presque sûrement :*

$$\frac{1}{n}\sum_{i=1}^{n} f\left(\lambda_i^{(n)}\right) \underset{n\to+\infty}{\longrightarrow} \int_{-2}^{2} f(x)\frac{\sqrt{4-x^2}}{2\pi}\mathrm{d}x.$$

Notons que la convergence des mesures spectrales empiriques ne dépend pas du choix de la loi des coefficients des matrices : on parle d'*universalité*. On consultera la Figure 1.1 pour une illustration du Théorème 1 dans le cas gaussien.

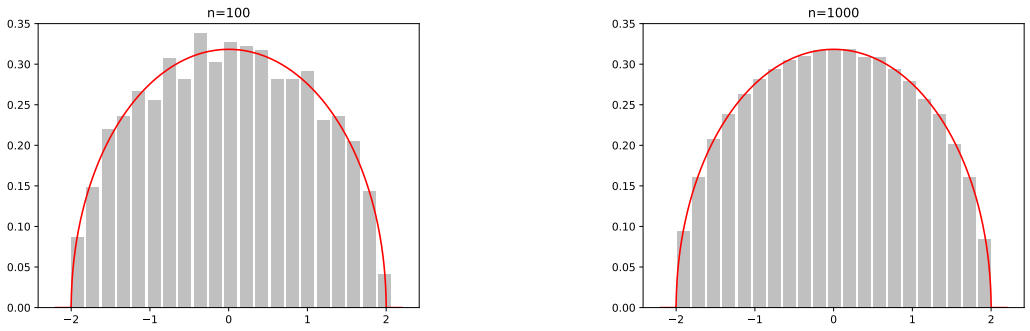


FIGURE 1.1 – Histogrammes des valeurs propres de 10 matrices de Wigner de taille $n \times n$ dont les coefficients sont des variables aléatoires gaussiennes centrées et réduites.

Wigner a d'abord démontré une version de ce résultat [Wig55] lorsque les coefficients de la matrices X_n sont égaux à ± 1 avec probabilité $1/2$, et a ensuite adapté les arguments de sa preuve au cas plus général de variables aléatoires symétriques dont tous les moments sont finis [Wig58].

La méthode de preuve de Wigner est appelée *méthode des moments* ; elle consiste à établir la convergence de l'espérance des moments de la mesure spectrale empirique vers ceux de la loi du demi-cercle. Le début du calcul est une réécriture du moment d'ordre k :

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{W_n}(x) = \mathrm{Tr}\left(W_n^k\right) = \frac{1}{n^{k/2}} \sum_{1\leq i_1,\ldots,i_k\leq n} (X_n)_{i_1 i_2}(X_n)_{i_2 i_3}\cdots(X_n)_{i_k i_1}.$$

Puisque les coefficients $(X_n)_{ij}$ sont centrés, tous les produits $(X_n)_{i_1 i_2}(X_N)_{i_2 i_3}\cdots(X_n)_{i_k i_1}$ tels qu'une variable $X_{i_p i_{p+1}}$ n'apparaît qu'une fois sont d'espérance nulle. La contribution asymptotique moyenne des autres termes peut être analysée en réécrivant la somme suivant une classe d'équivalence bien choisie sur les mots $i_1 i_2 \cdots i_k i_1$. Des considérations combinatoires permettent

alors d'établir que les moments moyens d'ordre impair sont asymptotiquement nuls et que les moments moyens d'ordre pair vérifient la convergence suivante

$$\lim_{n\to+\infty} \mathbf{E}\left[\int_{\mathbf{R}} x^{2k} \mathrm{d}\mu_{W_n}(x)\right] = \frac{1}{k+1}\binom{2k}{k}.$$

Notons que le terme de droite est le k-ème *nombre de Catalan* qui dénombre – entre autre – les arbres planaires enracinés possédant k arêtes. Ceci permet de conclure la démonstration puisqu'un calcul direct montre que les moments d'ordre pair de la loi du demi-cercle sont donnés par la suite des nombres de Catalan. Quant à ses moments d'ordre impair, ils sont tous nuls par un argument de symétrie.

Une analyse combinatoire analogue du terme de variance permet de convertir la convergence des moments moyens en une convergence en probabilité [Gre63]. Par ailleurs, des arguments de troncage ont permis à Arnold [Arn67, Arn71] d'affaiblir les hypothèses sur les coefficients de X_n et d'obtenir des résultats presque sûrs.

Il existe une autre approche à la preuve du Théorème de Wigner, reposant sur l'étude de la transformée de Stieltjes d'une mesure de probabilité.

Définition 1. *Soit μ une mesure de probabilité sur $\mathbf{R}$. La transformée de Stieltjes de μ est l'application analytique s_μ définie par*

$$\forall z \in \mathbf{C}_+, \quad s_\mu(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu(x)}{x-z},$$

où $\mathbf{C}_+ = \{z \in \mathbf{C}, \Im z > 0\}$.

Cette transformée a le bon goût de caractériser une mesure de probabilité et l'on dispose de formules d'inversion effectives. Par exemple, les atomes de μ correspondent aux pôles de s_μ et leurs poids sont donnés par les résidus de ces atomes. De plus, si μ possède une partie absolument continue par rapport à la mesure de Lebesgue en $x \in \mathbf{R}$, celle-ci est donnée par

$$\frac{\mathrm{d}\mu(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{t\to 0^+} \Im s_\mu(x+it) \tag{1.1}$$

Par ailleurs, la convergence étroite dans l'espace des mesures de probabilités est équivalente à la convergence ponctuelle des transformées de Stieltjes. Pour démontrer le Théorème de Wigner, il suffit donc d'établir la convergence des transformées de Stieltjes des mesures spectrales empiriques

$$s_{\mu_{W_n}}(z) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\lambda_i^{(n)} - z} = \mathrm{Tr}\,(W_n - zI_n)^{-1},$$

où I_n désigne la matrice identité de taille n. En tant que fonction de $z \in \mathbf{C}_+$, le membre de droite est communément appelé *résolvante de W_n*. Pour cette raison, la deuxième méthode de preuve du Théorème de Wigner que l'on va esquisser ici est parfois appelée *méthode de la résolvante*. Plus loin dans cette introduction, nous en présenterons une légère adaptation dans le contexte des matrices aléatoires déformées.

Par un argument de concentration, il est possible de réduire l'étude à l'espérance de la transformée de Stieltjes de μ_{W_n}, qui vaut $\mathbf{E}[s_{\mu_{W_n}}(z)] = \mathbf{E}[(W_n - zI_n)^{-1}_{11}]$ par invariance en loi des coefficients diagonaux de la matrice $(W_n - zI_n)^{-1}$. L'idée centrale de la preuve repose sur l'utilisation de la formule des compléments de Schur :

$$(W_n - zI_n)^{-1}_{11} = \frac{1}{(W_n - zI_n)_{11} - b^T C b}, \quad \text{où} \quad \begin{cases} b \text{ est le vecteur } ((W_n)_{i1})_{2\le i\le n}, \\ C \text{ est la matrice } ((W_n - zI_n)^{-1}_{ij})_{2\le i,j\le n}. \end{cases} \tag{1.2}$$

Encore une fois, un argument de concentration permet de remplacer la forme quadratique $b^T C b$ par son espérance, qui vaut $N^{-1}\sum_{2\leq i\leq N}\mathbf{E}[C_{ii}] = N^{-1}\mathrm{Tr}(C)$ puisque les coefficients de b sont centrés, de variance $n^{-1/2}$ et indépendants de ceux de C. Finalement, comme les valeurs propres de W_n s'entrelacent avec celles de son mineur principal $((W_n)_{ij})_{2\leq i,j\leq n}$, le terme $n^{-1}\mathrm{Tr}(C)$ peut être approximé par $n^{-1}\mathrm{Tr}(W_n - zI_n)^{-1}$ au prix d'une erreur d'ordre n^{-1}. Nous espérons ainsi avoir convaincu le lecteur que l'égalité suivante peut être obtenue de manière rigoureuse,

$$s_{\mu_{W_n}}(z) = \frac{1}{-z - s_{\mu_{W_n}}(z)} + o_{\mathbf{P}}(1). \tag{1.3}$$

Ainsi, tout point d'accumulation $s_\infty(z)$ de la suite $(s_{\mu_{W_n}}(z))_{n\geq 1}$ doit vérifier l'équation quadratique

$$s_\infty(z)^2 + z s_\infty(z) + 1 = 0.$$

et le choix de la branche peut être effectué en utilisant que $s_{\mu_{W_n}}(z) \sim -1/z$ lorsque $|z| \to +\infty$. Finalement, un argument de tension sur la suite $(s_{\mu_{W_n}})$ permet d'obtenir la convergence ponctuelle :

$$\forall z \in \mathbf{C}_+, \quad s_{\mu_{W_n}}(z) \xrightarrow[n\to+\infty]{\mathbf{P}} \frac{-z+\sqrt{z^2-4}}{2}.$$

Cela conclut la démonstration car le membre de droite est la transformée de Stieltjes de la loi du demi-cercle, comme le lecteur s'en convaincra en utilisant la formule d'inversion (1.1). Notons que la convergence en probabilité présentée ici peut être convertie en une convergence presque sûre par une étude plus approfondie du terme de reste dans l'Équation (1.3).

Plus grande valeur propre. Le Théorème de Wigner est une description macroscopique du spectre et il est naturel d'espérer pouvoir obtenir de plus fines informations, à commencer par le comportement asymptotique des valeurs propres extrêmes. Quitte à considérer $-W_n$, on se concentre ici sur la plus grande valeur propre. Celle-ci est asymptotiquement minorée par 2 car la loi du demi-cercle est supportée par l'intervalle $[-2,2]$. Füredi et Kómlos [FK81] furent les premiers à obtenir la convergence de la plus grande valeur propre vers le bord du support, en utilisant la *méthode des grandes traces* qui consiste à exploiter l'inégalité $\lambda_1^k \leq \mathrm{Tr}(W_n)^k$ pour des valeurs de k dépendant de n et tendant lentement vers l'infini. L'analyse du terme $\mathrm{Tr}(W_n)^k$ est réalisée avec des méthodes combinatoires analogues à la preuve du Théorème de Wigner par la méthode des moments. Les travaux de Füredi et Kómlos étaient restreints au cas où les moments des coefficients de X_n sont tous finis. L'hypothèse minimale nécessaire à la convergence de la plus grande valeur propre, à savoir l'existence d'un moment d'ordre quatre, a été identifiée par Baï et Yin [BY88].

Théorème 2. *En supposant que les coefficients de X_n sont des variables aléatoires centrées, réduites, et dont le quatrième moment est fini, la convergence suivante a lieu presque sûrement,*

$$\lambda_1^{(n)} \xrightarrow[n\to+\infty]{} 2.$$

Nous pouvons désormais nous interroger quant à la nature des fluctuations de la plus grande valeur propre autour de sa limite. De manière plus générale, existe-t-il une mise à l'échelle permettant d'étudier le processus ponctuel *microscopique* des valeurs propres autour d'un locus $x \in [-2,2]$ fixé ? Si oui, quelle est la limite de ce processus ? Ces questions s'avèrent ardues en général et ont d'abord été résolues dans le contexte de modèles *intégrables*, où des calculs explicites peuvent être menés à terme.

Modèles gaussiens. Au sein des matrices aléatoires, on peut distinguer trois modèles intégrables remarquables, chacun lié à une invariance en loi sous l'action respective des groupes unitaire, orthogonal et symplectique. On parle du GOE pour *Gaussian Orthogonal Ensemble*, du GUE pour *Gaussian Unitary Ensemble* et du GSE pour *Gaussian Symplectic Ensemble*. Le modèle le plus célèbre est sans doute le GUE et nous restreindrons notre présentation à ce cadre par soucis de concision. Pour tout $n \geq 1$, l'ensemble $\mathrm{GUE}(n)$ est défini comme une loi de probabilité sur les matrices *hermitiennes* de taille $n \times n$:

$$\mathbf{P}_n(\mathrm{d}X) := \frac{1}{Z_n} \exp\left(-\frac{1}{2}\mathrm{Tr}(XX^*)\right) \mathrm{d}X,$$

où Z_n est une constante de renormalisation. Un calcul astucieux, souvent présenté avec le niveau de rigueur des physiciens [Meh04], permet d'obtenir une formule explicite pour la densité des valeurs propres de ce modèle :

$$\mathbf{P}_n(\mathrm{d}\lambda) = \frac{1}{\widetilde{Z_n}} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \lambda_i^2\right) \mathrm{d}\lambda.$$

Une preuve rigoureuse pourra être consultée dans [Dei99] ou [Tao12]. À partir de cette formule, on peut vérifier que les valeurs propres du $\mathrm{GUE}(n)$ possèdent une structure déterminantale dont le noyau s'exprime en fonction des polynômes de Hermite. Rappelons ici qu'un processus déterminantal ayant n points et de noyau $K(\cdot,\cdot)$ est un processus ponctuel aléatoire $\{X_i\}_{1 \leq i \leq n}$ tel que, pour tout $1 \leq k \leq n$ et toute fonction mesurable $f : \mathbf{R}^k \to \mathbf{R}$:

$$\mathbf{E}\left[\sum_{i_1 \neq \cdots \neq i_k} f\left(X_{i_1}, \ldots, X_{i_k}\right)\right] = \int_{\mathbf{R}^k} f(x_1, \ldots, x_k) \det\left[\left(K(x_i, x_j)\right)_{1 \leq i,j \leq k}\right] \mathrm{d}x_1 \cdots \mathrm{d}x_k.$$

De manière informelle, en fixant $\varepsilon > 0$, il est possible de montrer que pour tout $x \in (-2+\varepsilon, 2-\varepsilon)$, le processus ponctuel formé par les valeurs propres du $\mathrm{GUE}(n)$ au voisinage de x converge, après une mise à l'échelle n^{-1}, vers le processus déterminantal de noyau

$$K_{\mathrm{Sine}}(x,y) := \frac{\sin(\pi(x-y))}{\pi(x-y)}.$$

Ce résultat a d'abord été démontré par Gaudin et Mehta [MG60]. En particulier, la loi de l'espacement entre deux valeurs propres consécutives à l'intérieur du spectre peut être calculée. Comme l'avait déjà remarqué Mehta [Meh60] par des calculs approchés, celle-ci diffère légèrement de la célèbre prédiction de Wigner, connue sous le nom de *Wigner's surmise* [Meh04].

Au bord du spectre, la densité de la loi du demi-cercle s'annule et les fluctuations des valeurs propres sont d'ordre $n^{-2/3}$. Leur comportement asymptotique est décrit par le processus déterminantal associé au noyau d'Airy, défini par

$$K_{\mathrm{Ai}}(x,y) := \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y},$$

où Ai est la fonction d'Airy, unique solution de l'équation différentielle $y'' = xy$ satisfaisant $y(x) \sim (4\pi\sqrt{x})^{1/2} \exp(-\frac{2}{3}x^{3/2})$ lorsque $x \to +\infty$. Définissons par ailleurs $q : \mathbf{R} \to \mathbf{R}$, l'unique solution de l'équation différentielle $q''(x) = xq(x) + 2q(x)^3$ satisfaisant $q(x) \sim \mathrm{Ai}(x)$ lorsque $x \to +\infty$. Alors les fluctuations de la plus grande valeur propre sont décrites par la célèbre loi éponyme de Tracy et Widom [TW94], dont la fonction de répartition F_{TW} est définie comme suit,

$$F_{\mathrm{TW}}(x) := \exp\left(-\int_x^{+\infty} (y-x)q(y)^2 \mathrm{d}y\right).$$

Théorème 3. *Pour tout $x \in \mathbf{R}$:*

$$\mathbf{P}\left(n^{2/3}\left(\frac{\lambda_1}{\sqrt{n}} - 2\right) \leq x\right) \underset{n\to+\infty}{\longrightarrow} F_{\mathrm{TW}}(x).$$

Mettons ici le lecteur en garde : lorsque l'on considère le modèle gaussien invariant par l'action du groupe orthogonal – le GOE – les fluctuations de la plus grande valeur propre sont décrites par une déformation de F_{TW}, pourtant toujours appelée loi de Tracy-Widom. Il existe en fait une famille à un paramètre de "lois de Tracy-Widom", chacune associée à un modèle intégrable que l'on va décrire dans le paragraphe suivant.

Digression : les β-ensembles de Hermite. Il est possible de généraliser les ensembles solubles précédents de la manière suivante. Pour tout $\beta > 0$ et tout $n \geq 1$, on considère $\mathbf{P}_{\beta,n}$ la mesure de probabilité sur l'ensemble des n-uplets de réels définie par

$$\mathbf{P}_{\beta,n}(\mathrm{d}\lambda) := \frac{1}{Z_{\beta,n}} \prod_{1\leq i<j\leq n} |\lambda_i - \lambda_j|^\beta \exp\left(-\frac{\beta}{4}\sum_{i=1}^n \lambda_i^2\right) \mathrm{d}\lambda.$$

L'ensemble gaussien du GUE correspond à $\beta = 2$. Mentionnons qu'en choisissant respectivement $\beta = 1$ et $\beta = 4$, on obtient une définition des ensembles $\mathrm{GOE}(n)$ et $\mathrm{GSE}(n)$. Dans le cas où β est un paramètre général, il n'y a *a priori* aucune raison pour que $\mathbf{P}_{\beta,n}$ décrive la loi des valeurs propres d'un modèle matriciel. De manière remarquable, Dumitriu et Edelman [DE02] ont proposé une représentation tridiagonale très élégante pour les β-ensembles, que nous rappelons ici. Définissons $(a_i)_{i\geq 1}$ et $(b_i(\beta))_{i\geq 1}$, deux familles indépendantes de variables aléatoires indépendantes, telles que, pour tout $i \geq 1$, a_i suit une loi normale centrée et de variance 2, et $b_i(\beta)$ une loi du χ de paramètre $i\beta$. Autrement dit :

$$\mathbf{P}(a_i \in \mathrm{d}x) = \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{x^2}{4}\right) \qquad \text{et} \qquad \mathbf{P}(b_i(\beta) \in \mathrm{d}x) = \frac{2^{1-\frac{i\beta}{2}}}{\Gamma\left(\frac{i\beta}{2}\right)} x^{i\beta-1} \exp\left(-\frac{x^2}{2}\right).$$

Pour tout $\beta > 0$, Dumitriu et Edelman considèrent la matrice tridiagonale suivante

$$H_\beta := \frac{1}{\sqrt{\beta}} \begin{bmatrix} a_n & b_{n-1}(\beta) & & & \\ b_{n-1}(\beta) & a_{n-1} & b_{n-2}(\beta) & & \\ & b_{n-2}(\beta) & a_{n-2} & \ddots & \\ & & \ddots & \ddots & b_1(\beta) \\ & & & b_1(\beta) & a_1 \end{bmatrix},$$

et montrent que la loi des valeurs propres de H_β est donnée par $\mathbf{P}_{\beta,n}$. Cette représentation des β-ensembles est loin d'être anecdotique : elle peut être exploitée pour obtenir une description asymptotique des processus ponctuels formés par les valeurs propres à l'intérieur et au bord du spectre limite. Une idée extrêmement féconde, d'abord proposée par Sutton et Edelman [ES07], consiste à interpréter les matrices tridiagonales comme des approximations d'opérateurs stochastiques de dimension infinie. À l'aide d'arguments heuristiques, ils conjecturent l'existence de deux opérateurs stochastiques limites : Sine_β pour "Sinus β" à l'intérieur du spectre, et Airy_β pour "Airy β" au bord du spectre. La preuve de la convergence vers Sine_β a été établie par Virág et Valkó [VV09] ; celle de la convergence vers Airy_β par Ramírez, Rider et Virág [RRV11].

Universalité et loi locale. Les matrices aléatoires présentent un phénomène remarquable d'universalité, déjà présent dans le Théorème de Wigner où la loi limite du spectre ne dépend pas de la loi des coefficients. Ce phénomène persiste pour le comportement microscopique des valeurs propres, où deux classes d'universalités existent selon que les coefficients de X_n sont des variables aléatoires complexes ou réelles. Dans le cas complexe, les statistiques locales du spectre coïncident avec celles du GUE. Dans le cas réel, ce sont celles du GOE, que le lecteur pourra consulter dans le livre de Mehta [Meh04].

Le premier résultat d'universalité a été obtenu par Soshnikov [Sos99], qui montre l'ubiquité de la loi de Tracy-Widom pour la plus grande valeur propre lorsque les coefficients de X_n sont centrés et sous-gaussiens. La preuve de ce résultat repose sur la méthode des grandes traces qui, malgré son apparente simplicité, permet d'obtenir des informations très précises sur le bord du spectre, au prix d'une étude combinatoire détaillée de $\mathrm{Tr} W_n^k$ pour des valeurs de k de l'ordre de $n^{2/3}$. Concernant l'universalité pour le comportement des valeurs propres à l'intérieur du support de la loi du demi-cercle, le premier résultat est sans doute dû à Johansson [Joh01a] pour des matrices de Wigner dont les coefficients sont obtenus par convolution avec des lois gaussiennes. Quelques années plus tard, Tao et Vu [TV10] montrent l'universalité des comportements locaux à l'intérieur et au bord du spectre de deux matrices de Wigner dont les quatre premiers moments des coefficients coïncident. Leur preuve est basée sur une comparaison avec le modèle Gaussien, rendu possible par l'intermédiaire d'une méthode de Lindeberg, d'abord introduite dans le contexte des matrices aléatoires par Chatterjee [Cha06].

Les résultats d'universalité les plus aboutis utilisent la machinerie des désormais célèbres lois locales. Celles-ci consistent à obtenir des estimées très précises sur les coefficients de la résolvante $(W_n - zI_n)^{-1}$, lorsque $z \in \mathbf{C}_+$ est autorisé à s'approcher de l'axe réel en fonction de la dimension des matrices. La première loi locale a été obtenue par Erdős, Schlein et Yau [ESY09]. Nous reproduisons ici l'énoncé de l'article de synthèse [BGK17] et renvoyons à ce même article pour une bibliographie détaillée. Introduisons d'abord, pour tout $z = E + i\eta$ et $\tau > 0$, le domaine du demi-plan complexe

$$\mathcal{D}_n^{(\tau)} := \left\{ z \in \mathbf{C};\, |E| \leq \tau^{-1},\, n^{-1+\tau} \leq \eta \leq \tau^{-1} \right\}, \tag{1.4}$$

et la fonction d'erreur

$$\psi_n(z) := \sqrt{\frac{\Im\left(s_{\mu_{sc}}(z)\right)}{n\eta}} + \frac{1}{n\eta}. \tag{1.5}$$

Théorème 4. *Soit $\tau > 0$. Alors, pour tout $\varepsilon > 0$, il existe $D > 0$ tel que les estimées suivantes sont vérifiées uniformément en $z \in \mathcal{D}_n^{(\tau)}$ et $i, j \in \{1, \ldots, N\}$,*

$$\mathbf{P}\left(\left| s_{\mu_{W_n}}(z) - s_{\mu_{sc}}(z) \right| \geq n^{\varepsilon} (n\eta)^{-1} \right) \leq n^{-D} \tag{1.6}$$

et

$$\mathbf{P}\left(\left| (W_n - z)^{-1}_{ij} - s_{\mu_{sc}}(z)\delta_{ij} \right| \geq n^{\varepsilon} \psi_n(z) \right) \leq n^{-D}. \tag{1.7}$$

Le Théorème 4 a de nombreuses conséquences ; il permet par exemple de démontrer que les valeurs propres de W_n sont proches des quantiles de la loi du demi-cercle et que les coefficients des vecteurs propres sont d'ordre $n^{-1/2}$. Avec plus de travail, il permet également d'établir, en un certain sens, l'universalité du comportement des valeurs propres à l'intérieur et au bord du spectre. Pour notre propos, une application mérite d'être soulignée.

Théorème 5. *Soit $x \in \mathbf{R}$ et une suite $(\varepsilon_n)_{n\geq 1}$ telle que $\varepsilon_n \gg n^{-1+\delta}$ pour un $\delta \in (0,1)$ fixé. Notons $I_x(\varepsilon_n) := (-x + \varepsilon_n, x - \varepsilon_n)$. Alors, pour toute suite $(\omega_n)_{n\geq 1}$ telle que $n^{-1+\delta} \ll \omega_n \ll \varepsilon_n$, il existe $D > 0$ tel que :*

$$\mathbf{P}\left(\left| \mu_{W_n}(I_x(\varepsilon_n)) - \mu_{sc}(I_x(\varepsilon_n)) \right| \geq \omega_n \right) \leq n^{-D}.$$

Le Théorème 5 est une confirmation locale du Théorème de Wigner : il garantit que la loi du demi-cercle est encore vérifiée pour des intervalles de longueurs microscopiques, avec grande probabilité. Sa démonstration repose sur l'utilisation de la transformée de Helffer-Sjöstrand, qui assure que toute fonction analytique $f : \mathbf{R} \to \mathbf{R}$[1] peut être représentée de la manière suivante

$$f(x) = \frac{1}{\pi} \int_{\mathbf{C}} \frac{\overline{\partial} \left(\widetilde{f}(z) \chi(z) \right)}{x - z} \mathrm{d}^2 z, \tag{1.8}$$

où :

- $\widetilde{f}(z) = \widetilde{f}(x+iy) = f(x) + iyf'(y)$ est le prolongement quasi-analytique de f ;
- χ est une fonction de cutoff analytique et telle que $\chi(x+iy) = 1$ pour tout $|y| \leq 1$ et $\chi(x+iy) = 0$ pour tout $|y| \geq 2$;
- $\overline{\partial} = \frac{1}{2}\left(\partial_x + i\partial_y\right)$ est la dérivée anti-holomorphe.

Ainsi, on peut réécrire :

$$\int_{\mathbf{R}} f(\lambda) \left(\mathrm{d}\mu_{W_n}(\lambda) - \mathrm{d}\mu_{sc}(\lambda) \right) = \int_{\mathbf{C}} \overline{\partial} \left(\widetilde{f}(z) \chi(z) \right) \left(s_{\mu_{W_n}}(z) - s_{\mu_{sc}}(z) \right) \mathrm{d}^2 z$$

et exploiter les lois locales du Théorème 4.

1.2 Le modèle de Wishart

Nous l'avons déjà mentionné : la première apparition des matrices aléatoires dans la littérature est l'article de J. Wishart [Wis28], où l'auteur identifie la loi d'une matrice de covariance pour des vecteurs gaussiens indépendants et de même loi. Nous présentons ici les résultats principaux concernant le spectre de matrices de covariance aléatoires, également appelées matrices de Wishart. Ces résultats étant analogues à ceux présentés dans le contexte des matrices de Wigner, nous nous contenterons d'une exposition succincte.

Comportement global du spectre. Pour tout $n \geq 1$, soit $m = m(n) \geq 1$ tel quel $m/n \to \alpha \in (0, \infty)$ lorsque $n \to +\infty$. Considérons également $X_n \in \mathbf{R}^{n \times m}$ une matrice rectangulaire de taille $n \times m$, dont on supposera les coefficients indépendants et de même loi, centrée et réduite. On appellera matrice de Wishart de taille n la matrice de covariance aléatoire

$$S_n := \frac{1}{n} X_n X_n^T,$$

dont les valeurs propres seront notées $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$. Comme dans le cadre des matrices de Wigner, on forme

$$\mu_{S_n} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i^{(n)}}$$

la mesure spectrale empirique associée à S_n. Celle-ci converge vers une mesure de probabilité déterministe, dénommée loi de Marchenko-Pastur de paramètre α,

$$\mu_{\mathrm{MP},\alpha}(\mathrm{d}x) := \frac{\sqrt{(b-x)(x-a)}}{2\pi x} \mathbf{1}_{x \in (a,b)} \mathrm{d}x + 1_{\alpha < 1}(1-\alpha)\delta_0(\mathrm{d}x), \tag{1.9}$$

où $a, b = (1 \pm \sqrt{\alpha})^2$.

[1] afin de démontrer le Théorème 5, on choisit une approximation de l'intervalle $I_x(\varepsilon_n)$.

Théorème 6. *Presque sûrement, la mesure μ_{S_n} converge étroitement vers $\mu_{\mathrm{MP},\alpha}$. Autrement dit, pour toute fonction $f : \mathbf{R} \mapsto \mathbf{R}$ continue et bornée, la convergence suivante a lieu presque sûrement,*

$$\frac{1}{n}\sum_{i=1}^{n} f\left(\lambda_i^{(n)}\right) \underset{n\to+\infty}{\longrightarrow} \int f(x)\mathrm{d}\mu_{\mathrm{MP},\alpha}(x).$$

On consultera la Figure 1.2 pour une illustration de ce Théorème dans le cas gaussien.

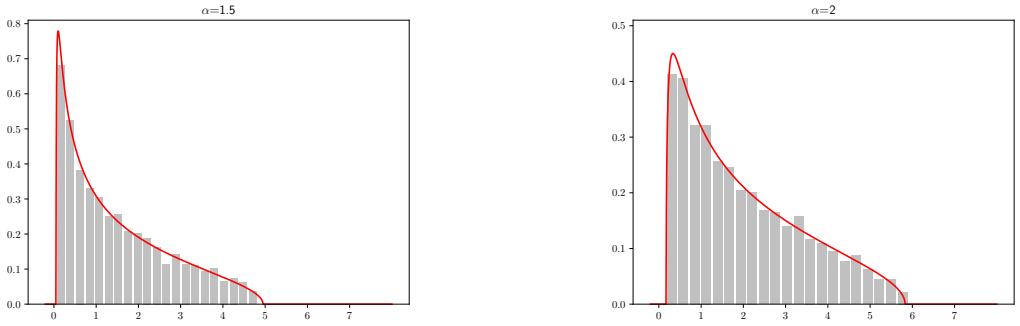


FIGURE 1.2 – Histogrammes des valeurs propres de 10 matrices de Wishart indépendantes telles que X_n est une matrice de taille $100 \times (\alpha \times 100)$ dont les coefficients sont des variables aléatoires gaussiennes centrées et réduites.

Ce résultat a été établi par Marchenko et Pastur [MP67]. Leur preuve inaugure la méthode de la résolvante pour l'étude des matrices aléatoires que nous avons présentée dans le contexte des matrices de Wigner. Quelques années plus tard, une preuve combinatoire utilisant la méthode des moments a été proposée par Jonsson [Jon82].

Remarque 1 (À propos des conventions de mise à l'échelle)**.** *Soulignons ici qu'il existe plusieurs manières de définir le modèle de Wishart. De nombreux auteurs considèrent $B_N = N^{-1}Y_N Y_N^T$, où Y_N est une matrice aléatoire de taille $p \times N$, de sorte que $p/N \to y \in (0, +\infty)$ lorsque $N \to +\infty$. Les changements de paramètres avec notre modèle sont donc $n = p, m = N, \alpha = 1/y$. De plus, en écrivant $B_N = (p/N)p^{-1}Y_N Y_N^T$, il est aisé de vérifier l'égalité en loi suivante :*

$$\mu_{B_N} = \Lambda_{n/m}(\mu_{S_n}),$$

où $\Lambda_\xi(\cdot)$ désigne le poussé en avant d'une mesure par la dilatation $x \mapsto \xi x$.

Plus grande valeur propre. La convergence vers la loi de Marchenko-Pastur garantit que la plus grande valeur propre est asymptotiquement minorée par le bord du spectre $b = (1 + \sqrt{\alpha})^2$. Comme pour les matrices de Wigner, la méthode des grandes traces permet en fait d'établir la convergence vers b. Sa première occurrence dans le contexte des matrices de covariance aléatoires est due à Geman [Gem80], dans le cas où les coefficients de X_n sont sous-gaussiens. Cette hypothèse a ensuite été affaiblie par Bai, Yin et Krishnaiah [YBK88] qui supposent uniquement l'existence d'un quatrième moment. Cette condition s'avère être nécessaire pour obtenir la convergence presque-sûre [BSY88].

Théorème 7. *En supposant que les coefficients de X_n sont des variables aléatoires centrées, réduites, et dont le quatrième moment est fini, la convergence suivante a lieu presque sûrement,*

$$\lambda_1^{(n)} \underset{n\to+\infty}{\longrightarrow} b = (1 + \sqrt{\alpha})^2.$$

Modèles gaussiens. En choisissant des coefficients gaussiens pour X_n, on obtient des modèles intégrables pour les matrices aléatoires symétriques et positives. Les ensembles qui en résultent sont nommés LOE, LUE et LSE pour *Laguerre Orthogonal, Unitary* et *Symplectic Ensemble*. Explicitons la loi de l'ensemble du LOE, obtenue en choisissant des gaussiennes réelles centrées et réduites, sans doute la plus utilisée du point de vue des applications. On se restreint sans perte de généralité au cas $m \geq n$, puisque $X_n X_n^T$ et $X_n^T X_n$ ont les mêmes valeurs propres non nulles. Pour tout $n \geq 1$, l'ensemble du LOE(n) correspond alors à la mesure de probabilité $\mathbf{P}_{n,m}$ induite par l'application $X \in \mathbf{R}^{n\times m} \mapsto XX^T$ sur l'espace des matrices symétriques et positives de taille $n \times n$, définie par

$$\mathbf{P}_{n,m}(\mathrm{dS}) := \frac{1}{Z_{n,m}} |\det(S)|^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2}\mathrm{Tr}(S)\right) \mathrm{d}S.$$

La loi des valeurs propres de ce modèle est donnée par la formule suivante :

$$\mathbf{P}_{n,m}(\mathrm{d}\lambda) = \frac{1}{\widetilde{Z_{n,m}}} \prod_{1\leq i<j\leq n} |\lambda_i - \lambda_j| \prod_{p=1}^{n} \lambda_i^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\lambda_i\right) \mathrm{d}\lambda.$$

Ces valeurs propres possèdent une structure déterminantale dont le noyau est relié aux polynômes de *Laguerre*, expliquant au passage l'appellation LOE. Des calculs asymptotiques sur ces polynômes permettent d'identifier les processus limites à l'intérieur et au bord du support de la loi de Marchenko-Pastur [PS11]. De manière remarquable, ceux-ci sont exactement les mêmes que les processus déterminantaux limites obtenus dans le cas symétrique du GOE. En particulier, les fluctuations de la plus grande valeur propre sont régies par la loi de Tracy-Widom correspondante, comme l'a d'abord démontré Johnstone [Joh01b].

Digression : les β-ensembles de Laguerre. Pour tout $m \geq n \geq 1$, et $\beta > 0$, le β-ensemble de Laguerre de paramètres n et m est décrit par une loi de probabilité $\mathbf{P}_{\beta,n,m}$ sur les n-uplets de réels positifs, définie de la manière suivante :

$$\mathbf{P}_{\beta,n,m}(\mathrm{d}\lambda) := \frac{1}{Z_{\beta,n,m}} \prod_{1\leq i<j\leq n} |\lambda_i - \lambda_j|^{\beta} \prod_{p=1}^{n} \lambda_i^{\frac{\beta}{2}(m-n+1)-1} \exp\left(-\frac{\beta}{2}\sum_{i=1}^{n}\lambda_i\right) \mathrm{d}\lambda.$$

Les valeurs particulières de $\beta = 1, 2, 4$ permettent de retrouver les ensembles LOE, LUE et LSE. Dumitriu et Edelman [DE02] ont proposé une représentation tridiagonale de cet ensemble, que nous rappelons ici. Pour tout $n \geq 1$, on considère la matrice bidiagonale suivante :

$$B_\beta := \begin{bmatrix} a_m & & & & \\ b_{n-1}(\beta) & a_{m-1} & & & \\ & b_{n-2}(\beta) & a_{m-2} & & \\ & & \ddots & \ddots & \\ & & & b_1(\beta) & a_{m-n+1} \end{bmatrix},$$

où $(a_{m-i})_{0\leq i\leq n-1}$ et $(b_i(\beta))_{1\leq i\leq n-1}$ sont deux familles indépendantes de variables aléatoires indépendantes telles que a_{m-i} suit une loi du χ de paramètre $\beta(m-i)$ et $b_i(\beta)$ une loi du χ de paramètre βi. Alors les valeurs propres de la matrice de covariance aléatoire $L_\beta := \beta^{-1} B_\beta B_\beta^T$ sont distribuées selon la loi $\mathbf{P}_{\beta,n,m}$. Dans ce contexte, l'approche de Sutton et Edelman, qui consiste à interpréter L_β comme la version discrète d'un opérateur de dimension infinie, fonctionne encore. De manière intéressante, les processus limites des valeurs propres à l'intérieur et au bord du spectre sont *les mêmes*, à savoir Sine_β et Ai_β, que dans le cas symétrique que l'on a déjà

abordé. On pourra trouver les résultats correspondants dans les articles [VV09] et [RRV11]. En particulier, dans le cas du LUE – où $\beta = 2$, le processus ponctuel formé par les valeurs propres au voisinage d'un point à l'intérieur du support converge vers le processus déterminantal de noyau K_{Sine}, et le processus ponctuel formé par les plus grandes valeurs propres vers le processus déterminantal de noyau K_{Ai}.

Universalité et loi locale. Comme dans le contexte des matrices de Wigner, le comportement local des valeurs propres des matrices de Wishart ne dépend que du corps de base choisi pour les coefficients.

Lorsque les coefficients de X_n sont symétriques et sous-gaussiens, la méthode des grandes traces a permis à Soshnikov [Sos02] de démontrer que les fluctuations de la plus grande valeur propre d'une matrice de Wishart sont asymptotiquement décrites par la loi de Tracy-Widom. L'universalité du comportement des valeurs propres à l'intérieur a ensuite été obtenue par Tao et Vu [TV12], à l'aide d'une adaptation de leur méthode des quatre moments.

La machinerie des lois locales s'applique aussi aux matrices de Wishart et permet d'obtenir les résultats d'universalité les plus fins. Dans ce contexte, l'analogue du Théorème 4 a été obtenu par Bloemendal et ses co-auteurs [BEK+14]. Afin d'énoncer leur résultat, rappelons au lecteur que le domaine du demi-plan supérieur $\mathcal{D}_n^{(\tau)}$ a été introduit dans l'Équation (1.4), et définissons la fonction d'erreur

$$\psi_n(z) := \sqrt{\frac{\Im\left(s_{\mu_{\text{MP},\alpha}}(z)\right)}{n\eta}} + \frac{1}{n\eta}, \tag{1.10}$$

où $s_{\mu_{\text{MP},\alpha}}$ désigne la transformée de Stieltjes de la loi de Marchenko-Pastur. Nous espérons qu'aucune confusion ne sera possible entre les deux définitions (1.5) et (1.10) de la fonction d'erreur ψ_n, chacune étant propre au contexte des matrices de Wigner ou à celui des matrices de Wishart.

Théorème 8. *Soit $\tau > 0$. Alors, pour tout $\varepsilon > 0$, il existe $D > 0$ tel que les estimées suivantes sont vérifiées uniformément en $z \in \mathcal{D}_n^{(\tau)}$ et $i, j \in \{1, \dots, N\}$:*

$$\mathbf{P}\left(\left|s_{\mu_{S_n}}(z) - s_{\mu_{\text{MP},\alpha}}(z)\right| \geq n^{\varepsilon}(n\eta)^{-1}\right) \leq n^{-D} \tag{1.11}$$

et

$$\mathbf{P}\left(\left|(S_n - z)^{-1}_{ij} - s_{\mu_{\text{MP},\alpha}}(z)\delta_{ij}\right| \geq n^{\varepsilon}\psi_n(z)\right) \leq n^{-D}. \tag{1.12}$$

Ici encore, la méthode de Hellfer-Sjöstrand permet d'obtenir une version locale du Théorème de Marchenko et Pastur.

Théorème 9. *Soit $x \in \mathbf{R}_+$ et $(\varepsilon_n)_{n\geq 1}$ telle que $\varepsilon_n \gg n^{-1+\delta}$ pour un $\delta \in (0,1)$ fixé. Notons $I_x(\varepsilon_n) := (-x + \varepsilon_n, x - \varepsilon_n)$. Alors, pour toute suite $(\omega_n)_{n\geq 1}$ telle que $n^{-1+\delta} \ll \omega_n \ll \varepsilon_n$, il existe $D > 0$ tel que :*

$$\mathbf{P}\left(\left|\mu_{S_n}(I_x(\varepsilon_n)) - \mu_{\text{MP},\alpha}(I_x(\varepsilon_n))\right| \geq \omega_n\right) \leq n^{-D}.$$

Avant de nous tourner vers de possibles généralisations des modèles de Wigner et de Wishart, mentionnons ici que de manière générique, le comportement local des valeurs propres des matrices aléatoires ne dépend pas des symétries globales du modèle considéré, mais uniquement du corps de base choisi pour les coefficients. Par exemple, les statistiques locales de l'ensemble gaussien GOE (resp. GUE) coïncident avec celles du LOE (resp. LUE). Ce phénomène remarquable a été démontré pour des modèles de plus en plus généraux ces dernières années. La technique désormais éprouvée, baptisée *the three steps strategy*, est due à Erdős et ses collaborateurs. Nous n'en discuterons pas dans cette thèse, et indiquons la référence [EY17] au lecteur curieux.

1.3 Première généralisation : les coefficients

Nous présentons ici deux généralisations des modèles de Wigner et de Wishart correspondant à deux affaiblissements possibles des hypothèses réalisées dans les parties précédentes au sujet des coefficients des matrices aléatoires X_n. La première généralisation concerne le cas de variables aléatoires appartenant au bassin d'attraction de lois stables. La seconde, qui fait l'objet du Chapitre 3 de cette thèse, issu de l'article [Noi18], concerne le cas de variables aléatoires dont la loi dépend de la dimension de la matrice.

Matrices de Lévy. Commençons par rappeler qu'une fonction $L : \mathbf{R}_+ \to \mathbf{R}_+^*$ est dite à variation lente lorsque, pour tout $x > 0$,

$$\lim_{t\to+\infty} \frac{L(tx)}{L(t)} = 1.$$

Soit $\beta \in (0,2)$[2]. Pour tout $n \geq 1$, soit X_n une matrice aléatoire symétrique, dont les coefficients sont i.i.d. de loi commune appartenant au bassin d'attraction d'une loi β-stable. Autrement dit, il existe une fonction à variation lente L telle que $\mathbf{P}(|X_n(1,1)| \geq x) = L(x)/x^\beta$. On appellera matrice de Lévy de paramètre β la matrice aléatoire

$$W_n^{(\beta)} := \frac{1}{n^{1/\beta}} X_n.$$

L'appellation est due à l'article fondateur de Bouchaud et Cizeau [CB94]. Dans la littérature mathématique, Ben Arous et Guionnnet [BAG08] ont démontré la convergence de la mesure spectrale empirique de ce modèle. De manière informelle, la limite est une mesure de probabilité dont la densité est de la forme $C_\beta/|x|^{1+\beta}$ lorsque $x \to +\infty$.

Dans le contexte des matrices de covariance aléatoires, on définit la matrice de Lévy de taille n par

$$S_n^{(\beta)} := \frac{1}{n^{2/\beta}} X_n X_n^T, \tag{1.13}$$

où $X_n \in \mathbf{R}^{n\times m}$ est une matrice rectangulaire de taille $n \times m$ dont les coefficients sont i.i.d. de loi commune appartenant au bassin d'attraction d'une loi β-stable. Ici, la convergence de la mesure spectrale empirique a été obtenue par Belinschi, Dembo et Guionnet [BDG09].

Avant de nous tourner vers une autre généralisation possible pour la loi des coefficients de X_n, il nous faut introduire quelques notations. Celles-ci concernent les graphes et les chemins tracés sur les graphes, et nous permettront d'énoncer les résultats à venir.

Quelques notations. Un graphe $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ est constitué d'un ensemble de sommets V et d'un ensemble d'arêtes $\mathrm{E} \subset \{\{u,v\};\, u,v \in \mathrm{V}\}$. Un graphe étiqueté est la donnée d'un graphe G et d'une application bijective entre V et $\{1,\dots,|V|\}$. Dans ce contexte, on appelle *ré-étiquetage* du graphe tout nouveau choix d'application bijective (il y en a $|V|!$). On appellera *mot de longueur k sur* G toute suite d'étiquettes $i_1,\dots,i_k$ telle que, pour tout $1 \leq j \leq k-1$, $\{i_j, i_{j+1}\}$ est une paire d'étiquettes dont les sommets associés forment une arête dans le graphe. Un mot est dit *fermé* lorsque $i_1 = i_k$. Deux mots fermés de longueur k, $\mathbf{i} = i_1 \dots i_k$ et $\mathbf{i}' = i'_1 \dots i'_k$, sont dits *équivalents* lorsqu'il existe une bijection σ de $\{1,\dots,|V|\}$ telle que $\sigma(i_j) = i'_j$ pour tout $1 \leq j \leq k$. Autrement dit, deux mots sont équivalents lorsqu'il existe un ré-étiquetage de G échangeant $\mathbf{i}$ et $\mathbf{i}'$. Cela définit une relation d'équivalence sur les mots fermés de longueur k.

[2]Nous nous excusons ici auprès du lecteur habitué à la notation consacrée "α". Ce paramètre est malheureusement déjà utilisé dans notre définition des matrices de Wishart, où il correspond au rapport entre le nombre de colonnes et le nombre de lignes de X_n.

Nous allons restreindre notre propos aux arbres planaires enracinés, dont on donne ici une définition formelle. Par convention, on notera $(\mathbf{N}^*)^0 = \emptyset$.

Définition 2. *Un arbre planaire enraciné* T *est un sous ensemble fini de* $\cup_{i\geq 0}(\mathbf{N}^*)^i$ *contenant* $\emptyset$ *et vérifiant les règles d'hérédités suivantes :*

- *si* $(u_1, \ldots, u_n) \in \mathrm{T}$, *alors* $(u_1, \ldots, u_{n-1}) \in \mathrm{T}$;
- *si* $(u_1, \ldots, u_n) \in \mathrm{T}$, *alors* $(u_1, \ldots, i) \in \mathrm{T}$ *pour tout* $1 \leq i \leq u_n - 1$.

L'élément $\emptyset$ *est appelé la racine de* T.

On se représentera un arbre planaire enraciné comme un arbre généalogique dessiné dans le plan de sorte que les enfants d'un individu soient ordonnés par ordre de naissance de gauche à droite. Dans ce dessin, l'ancêtre commun est la racine $\emptyset$. Étant donnée une arête $e = \{u, v\}$ d'un arbre planaire enraciné telle que u est le parent de v, on dira que e est une arête impaire (resp. paire) lorsque la distance de v à la racine est impaire (resp. paire).

Soit $k \geq 1$, $a \geq 1$, $1 \leq l \leq a$ et $\mathbf{b} = (b_1, \ldots, b_a)$ un a-uplet d'entiers tel que $b_1 \geq \cdots \geq b_a \geq 2$ et $b_1 + \cdots + b_a = 2k$. On notera :

- $\mathcal{W}_k(a, \mathbf{b})$ un ensemble de représentants des mots fermés de longueurs $2k$ sur les arbres planaires enracinés ayant a arêtes, tels que i_1 est la racine de l'arbre et tels que, pour tout $1 \leq i \leq a$, une arête est parcourue b_i fois par le mot ;
- $\mathcal{W}_k(a, l, \mathbf{b})$ un ensemble de représentants des mots fermés de longueurs $2k$ sur les arbres planaires enracinés ayant a arêtes dont l sont impaires, tels que i_1 est la racine de l'arbre et tels que, pour tout $1 \leq i \leq a$, une arête est parcourue b_i fois par le mot.

Avant de présenter une seconde généralisation possible concernant la loi des coefficients des matrices aléatoires de Wigner et de Wishart, mentionnons ici que les deux ensembles introduits ci-dessus permettent d'exprimer les moments de la loi du demi-cercle et de la loi de Marchenko-Pastur : pour tout $k \geq 0$,

$$\left\{ \begin{array}{lcl} \int_{\mathbf{R}} x^{2k} \mathrm{d}\mu_{sc}(x) & = & |\mathcal{W}_k(k, (2, \ldots, 2))| \,, \\ \int_{\mathbf{R}} x^k \mathrm{d}\mu_{\mathrm{MP},\alpha}(x) & = & \sum_{l=1}^{k} \alpha^l \, |\mathcal{W}_k(k, l, (2, \ldots, 2))| \,. \end{array} \right. \tag{1.14}$$

La première égalité provient du fait que $\mathcal{W}_k(k, (2, \ldots, 2))$ est en bijection avec l'ensemble des arbres planaires enracinés possédant k arêtes (via l'application "contour"), et possède donc $\mathrm{Cat}(k)$ éléments. Le lecteur pourra consulter une preuve de la seconde égalité dans la partie 3.4.1 du Chapitre 3.

Matrices dont les coefficients explosent. Considérons une suite de mesures de probabilités centrées $(P_n)_{n\geq 1}$. Pour tout $k \geq 1$, on notera $M_k(P_n)$ le moment d'ordre k de P_n, que l'on supposera fini,

$$M_k(P_n) := \int_{\mathbf{R}} x^k \mathrm{d}P_n(x) < +\infty.$$

De plus, nous supposerons qu'il existe une suite de réels positifs $(A_k)_{k\geq 2}$ telle que

$$\forall k \geq 2, \quad \lim_{n\to+\infty} \frac{M_k(P_n)}{n^{\frac{k}{2}-1} M_2(P_n)^{\frac{k}{2}}} = A_k. \tag{1.15}$$

Notons que l'on a toujours $A_2 = 1$.

Pour tout $n \geq 1$, soit $X_n \in \mathbf{R}^{n\times n}$ une matrice aléatoire symétrique dont les coefficients sont i.i.d. de loi commune P_n. On s'intéresse à la matrice de Wigner généralisée

$$W_n := \frac{1}{\sqrt{M_2(P_n)n}} X_n.$$

Lorsque la suite $(P_n)_{n\geq 1}$ est constante, on retrouve le modèle classique introduit dans la partie 1.1. La mesure spectrale empirique de W_n se prête à une étude par la méthode des moments, la combinatoire apparaissant à la limite étant sensiblement plus compliquée que dans le cas classique des matrices de Wigner. Dans la littérature, la première occurrence d'une telle méthode est due à Ryan [Rya98] qui obtient la convergence de la mesure spectrale empirique vers une mesure de probabilité déterministe dont les moments s'expriment uniquement en fonction de la suite $(A_k)_{k\geq 2}$. Les formules qu'il propose utilisent le formalisme des probabilités libres et notamment des *partitions non-croisées* dont on ne parlera pas au cours de cette thèse. Quelques années plus tard, Zakharevich [Zak06] a obtenu une description combinatoire plus directe, que nous reproduisons ici.

Théorème 10. *Pour tout $k \geq 1$,*

$$\lim_{n\to+\infty} \mathbf{E}\left[\int_{\mathbf{R}} x^{2k} \mathrm{d}\mu_{W_n}\right] = \sum_{a=1}^{k} \sum_{\substack{\mathbf{b}=(b_1,\dots,b_a) \\ b_1\geq\cdots\geq b_a\geq 2 \\ b_1+\cdots+b_a=2k}} |\mathcal{W}_k(a, a+1, \mathbf{b})| \prod_{i=1}^{a} A_{b_i}.$$

Plaçons nous maintenant dans le contexte des matrices de Wishart, ce qui revient à considérer la matrice de covariance aléatoire

$$S_n := \frac{1}{M_2(P_n)n} X_n X_n^T,$$

où $X_n \in \mathbf{R}^{n\times m}$ est une matrice rectangulaire de taille $n \times m$ dont les coefficients sont i.i.d. de loi commune P_n. On suppose bien sûr que $m/n \to \alpha \in (0, +\infty)$. Pour ce modèle, l'analogue des travaux de Ryan est dû à Benaych-Georges et Cabanal-Duvillard [BGCD12]. Le premier résultat de cette thèse propose une approche *à la Zakharevich* pour l'étude de la mesure spectrale empirique.

Théorème A : Chapitre 3, Proposition 1

Notons μ_{S_n} la mesure spectrale empirique de S_n. Il existe une mesure de probabilité $\mu_{\mathscr{A}} = \mu_{(A_k)_{k\geq 2}}$ telle que, en probabilité, μ_{S_n} converge étroitement vers $\mu_{\mathscr{A}}$. De plus, pour tout $k \geq 1$,

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\mathscr{A}}(x) = \sum_{a=1}^{k} \sum_{l=1}^{a} \alpha^l \sum_{\substack{\mathbf{b}=(b_1,\dots,b_a) \\ b_1\geq\cdots\geq b_a\geq 2 \\ b_1+\cdots+b_a=2k}} |\mathcal{W}_k(a, a+1, l, \mathbf{b})| \prod_{i=1}^{a} A_{b_i}. \tag{1.16}$$

Remarque 2. *Comme le nombre de passages d'un mot sur une arête d'un arbre est nécessairement pair, la formule (1.16) entraîne que $\mu_{\mathscr{A}}$ ne dépend que de $(A_{2k})_{k\geq 1}$. Dans le cas particulier où $A_{2i} = 0$ pour tout $i > 1$, la formule (1.16) se réduit à la deuxième ligne de l'Équation (1.14) et on retrouve la loi de Marchenko-Pastur : $\mu_{\mathscr{A}} = \mu_{\mathrm{MP},\alpha}$.*

On pourra trouver une preuve de ce Théorème dans la partie 3.3 du Chapitre 3. Ce résultat possède au moins deux applications. La première concerne les spectres de grands graphes aléatoires bipartis et dilués et sera présentée dans la partie 2.4 de cette Introduction. La seconde s'intéresse aux matrices de Lévy tronquées.

Matrices de covariance de Lévy tronquées. Rappelons au lecteur que la définition d'une matrice de covariance à queue lourde de paramètre $\beta \in (0,2)$ est donnée dans l'Équation (1.13). Notons P la loi des coefficients de X_n. Par hypothèse, celle-ci appartient au bassin d'attraction d'une loi β-stable et ne possède donc pas de second moment. Une idée naturelle pour pallier à ce déficit est de considérer une suite de versions tronquées de la loi P, que l'on introduit maintenant. On notera F la fonction de répartition de P. Pour tout $n \geq 1$, définissons les quantiles q_n^- et q_n^+ par les formules :

$$\begin{cases} F(q_n^-) &= 1/n, \\ 1 - F(q_n^+) &= 1/n. \end{cases}$$

Pour tout $B > 0$, on considère la version tronquée de X_n, notée $X_n^{(B)}$, dont les coefficients sont

$$X_n^{(B)}(i,j) = X_n(i,j)\mathbf{1}_{Bq_n^- \leq X_n(i,j) \leq Bq_n^+} + Bq_n^- \mathbf{1}_{X_n(i,j) < Bq_n^-} + Bq_n^+ \mathbf{1}_{X_n(i,j) > Bq_n^+}.$$

On notera P_n la loi de $X_n^{(B)}(i,j)$. La suite formée par les mesures de probabilité $(P_n)_{n\geq 1}$ satisfait l'hypothèse (1.15) et les coefficients associés s'écrivent sous la forme

$$A_{2k} = B^{\beta(k-1)} c_{k,\beta,F},$$

où $c_{k,\beta,F}$ est une constante explicite dépendant de k, β et de la fonction de répartition de P – un calcul explicite de cette constante fait l'objet de la partie 3.4.2. Ainsi, le Théorème A s'applique et la suite des mesures spectrales empiriques associées aux matrices $(M_2(P_n)n)^{-1/2} X_n^{(B)}$ converge vers une mesure de probabilité déterministe qu'on note $\mu_{\alpha,\beta,B}$. Lorsque B tend vers l'infini, les moments de cette mesure explosent et l'on s'attend à retrouver le régime des matrices de Lévy. À l'inverse, lorsque B tend vers 0, on s'attend à retrouver le régime classique des matrices de Wishart. Le Théorème A permet non seulement de confirmer cette prédiction, mais aussi d'identifier un terme d'erreur. Plus précisément, lorsque $B \to 0^+$, le moment d'ordre k de la mesure $\mu_{\alpha,\beta,B}$ s'écrit

$$\sum_{l=1}^{k} \alpha^l |\mathcal{W}_k(k, k+1, l, (2, \ldots, 2))| + A_4 \sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k(k, k+1, l, (4, 2, \ldots, 2))| + \mathcal{O}(B^\beta). \tag{1.17}$$

Nous avons déjà mentionné au cours de l'Équation (1.14) que le premier terme de ce développement correspond au moment d'ordre k de la loi de Marchenko-Pastur. De manière remarquable, le second terme correspond au moment d'ordre k de la mesure signée de masse nulle suivante

$$\mu_{\mathrm{MP},\alpha}^{(1)}(\mathrm{d}x) := \frac{x^2 - 2(\alpha+1)x + (\alpha^2+1)}{2\alpha\pi\sqrt{(b-x)(x-a)}} \mathbf{1}_{x\in(a,b)} \mathrm{d}x. \tag{1.18}$$

L'identification de $\mu_{\mathrm{MP},\alpha}^{(1)}$, obtenue dans la partie 3.4.1, est réalisée en deux étapes :

1. obtenir une équation fonctionnelle sur la série génératrice

$$G(z) := \sum_{k\geq 0} z^k \sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k(k, k+1, l, (4, 2, \ldots, 2))|,$$

2. identifier la densité (1.18) en utilisant la formule d'inversion (1.1) de la transformée de Stieltjes $S(z) = -z^{-1}G(z^{-1})$.

On obtient alors un développement asymptotique au premier ordre des moments de la mesure spectrale empirique de matrices de covariance à queues lourdes tronquées.

Théorème B : Chapitre 3, Théorème 2

Soit $\mu_{\mathrm{MP},\alpha}$ Pour tout $k \geq 1$, lorsque $B \to 0$,

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\alpha,\beta,B}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_{\mathrm{MP},\alpha}(x) + B^\beta c_{4,\beta,F} \int_{\mathbf{R}} x^k \mathrm{d}\mu^{(1)}_{\mathrm{MP},\alpha}(x) + \mathcal{O}(B^\beta).$$

Dans la partie 2.4 de cette Introduction, nous énoncerons un développement asymptotique similaire au sujet des spectres de graphes d'Erös-Rényi bipartis.

1.4 Deuxième généralisation : modèles matriciels perturbés

Nous présentons maintenant deux modèles de matrices aléatoires déformées correspondant respectivement à des perturbations additives de matrices de Wigner et à des perturbations multiplicatives de matrices de Wishart. Dans ce nouveau contexte, on s'intéresse à l'influence de la matrice de perturbation sur la matrice perturbée. À un niveau macroscopique, il est possible d'obtenir des déformations des lois de Wigner et de Marchenko-Pastur pour les limites des mesures empiriques. À l'échelle microscopique, nous verrons qu'il existe des transitions de phase concernant la possible existence de valeurs propres n'appartenant pas au support de la mesure empirique limite. L'étude de ce phénomène a été initiée par Baïk, Ben Arous et Péché en 2006 et a engendré un nombre impressionnant de travaux. Bien que leur article fondateur concerne les matrices de covariance aléatoires, nous ne respecterons pas la chronologie dans notre présentation et énoncerons d'abord les résultats principaux de la littérature dans le contexte des matrices de Wigner, puis dans celui des matrices de Wishart. Nous présentons ensuite notre contribution qui concerne l'étude de la *mesure spectrale dans la direction de la perturbation*. L'article correspondant fait l'objet du Chapitre 4, issu de l'article [Noi20]. Enfin, nous discuterons d'un lien entre le modèle des matrices de covariance déformées et un modèle de physique statistique, obtenu en collaboration avec Nathanaël Enriquez.

Nous utiliserons les mêmes notations que dans les parties 1.1 et 1.2, et supposerons que tous les moments des coefficients des matrices X_n sont finis.

Perturbation additive de matrices de Wigner. Pour tout $n \geq 1$, soit $A_n \in \mathbf{R}^{n\times n}$ une matrice symétrique de taille $n \times n$ dont on notera $\gamma_1^{(n)}, \dots, \gamma_n^{(n)}$ les valeurs propres et $v_1^{(n)}, \dots, v_n^{(n)}$ les vecteurs propres normalisés associés. On suppose que ces valeurs propres sont bornées et que leur distribution empirique converge vers une mesure de probabilité μ_A :

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \xrightarrow[n\to+\infty]{} \mu_A,$$

au sens de la convergence étroite. On s'intéresse alors à la matrice de Wigner déformée

$$W_n := \frac{1}{\sqrt{n}} X_n + A_n.$$

Soient $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ les valeurs propres de W_n et $\phi_1^{(n)}, \dots, \phi_n^{(n)}$ les vecteurs propres normalisés associés. En exploitant la méthode de la résolvante, Pastur [Pas72] a identifié la limite des mesures spectrales empiriques de ce modèle.

Théorème 11. *Presque-sûrement, la mesure spectrale empirique μ_{W_n} converge vers une mesure de probabilité déterministe $\mu_{sc} \boxplus \mu_A$, caractérisée par l'équation*

$$\forall z \in \mathbf{C}_+, \quad s_{\mu_{sc}\boxplus\mu_A}(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu_A(\lambda)}{\lambda - s_{\mu_{sc}\boxplus\mu_A}(z) - z}.$$

La notation $\mu_{sc} \boxplus \mu_A$ se lit "convolution libre des mesures μ_{sc} et μ_A", et n'est pas sans lien avec les probabilités libres... Nous renvoyons le lecteur au livre de Nica et Speicher [NS06] pour plus de détails sur ce lien profond entre la théorie des matrices aléatoires et celle des probabilités libres, développée par Voiculescu dans les années quatre-vingt.

Mentionnons que des lois locales analogues au Théorème 4 ont été récemment obtenues par Knowles et Yin [KY17] pour ce modèle ; le lecteur pourra consulter l'Équation (4.7) du Chapitre 4 pour une formulation précise.

On suppose maintenant qu'il existe $\theta \in \mathbf{R}$ tel que $\gamma_1^{(n)} \equiv \theta$. Supposons de plus que pour tout $n \geq 1$, aucune des autres valeurs propres de A_n n'est *atypique*, en ce sens que

$$\sup_{2\leq i\leq n} \operatorname{dist}\left(\gamma_i^{(n)}, \operatorname{Supp}(\mu_A)\right) \xrightarrow[n\to+\infty]{\mathbf{P}} 0.$$

Du point de vue de l'inférence statistique, on aimerait estimer le paramètre θ à partir de l'observation de W_n. Selon sa valeur, il peut être responsable de l'apparition d'une valeur propre atypique, un *outlier*, dans le spectre de W_n. Notons $w(z) = s_{\mu_{sc}\boxplus\mu_A}(z) + z$.

Théorème 12. *Supposons qu'il existe $x_\theta \notin \operatorname{Supp}(\mu_{sc} \boxplus \mu_A)$ tel que $w(x_\theta) = \theta$. Soit $\delta > 0$ tel que $[x_\theta - \delta, x_\theta + \delta] \cap \operatorname{Supp}(\mu_{sc} \boxplus \mu_A) = \emptyset$. Alors pour tout n suffisamment grand, il existe une unique valeur propre de W_n dans l'intervalle $[x_\theta - \delta, x_\theta + \delta]$. En notant i_n l'indice de cette valeur propre, les deux convergences suivantes ont lieu en probabilité,*

$$\lambda_{i_n}^{(n)} \xrightarrow[n\to+\infty]{} x_\theta \qquad \text{et} \qquad |\langle \phi_{i_n}^{(n)}, v_1^{(n)} \rangle|^2 \xrightarrow[n\to+\infty]{} \frac{1}{w'(x_\theta)}.$$

À l'inverse, si $w(x) = \theta$ n'admet pas de solution en dehors du support de $\mu_{sc} \boxplus \mu_A$, le spectre de W_n n'admet pas de valeur propre atypique, en ce sens que

$$\sup_{1\leq i\leq n} \operatorname{dist}\left(\lambda_i^{(n)}, \operatorname{Supp}(\mu_{sc} \boxplus \mu_A)\right) \xrightarrow[n\to+\infty]{\mathbf{P}} 0.$$

La convergence de la valeur propre atypique que l'on vient d'énoncer est une version simplifée d'un résultat dû à Capitaine, Donati-Martin, Féral et Février [CDMFF11], qui ont étudié le cas d'une matrice A_n possédant *plusieurs* valeurs propres atypiques. Le comportement asymptotique de la projection du vecteur propre associé à la valeur propre atypique a été obtenu plus tard par Capitaine [Cap13].

Lorsque $A_n = \theta e_1 e_1^T$, où e_1 désigne le premier vecteur de la base canonique, la mesure empirique limite est la loi du demi-cercle puisque $\mu_{sc} \boxplus \delta_0 = \mu_{sc}$. Dans ce contexte, des calculs explicites sont possibles et le Théorème 12 se traduit de la manière suivante.

Théorème 13. *Supposons que $A_n = \theta e_1 e_1^T$, $\theta \geq 0$. Alors les convergences suivantes ont lieu en probabilité :*

$$\lambda_1^{(n)} \xrightarrow[n\to+\infty]{} \begin{cases} 2 & si\ \theta \leq 1; \\ \theta + \frac{1}{\theta} > 2 & si\ \theta > 1. \end{cases} \tag{1.19}$$

et

$$|\langle \phi_1^{(n)}, e_1 \rangle|^2 \xrightarrow[n\to+\infty]{} \begin{cases} 0 & si\ \theta \leq 1; \\ 1 - \frac{1}{\theta^2} & si\ \theta > 1. \end{cases} \tag{1.20}$$

Mentionnons que la convergence (1.19) a d'abord été obtenue par Péché [Pé06] dans le cas d'une perturbation de rang fini du GUE. Dans cet article, l'auteure a également identifié les fluctuations de la plus grande valeur propre de ce modèle : elles suivent une loi gaussienne lorsque $\theta > 1$, la loi de Tracy-Widom lorsque $\theta < 1$ et une déformation de la loi de Tracy-Widom lorsque $\theta = 1$. De manière intéressante, les fluctuations de la plus grande valeur propre ne sont pas universelles dans le cas général, mais dépendent de la loi des coefficients de X_n : ce phénomène a d'abord été constaté dans [CDMF09], et a été récemment analysé en détail par Knowles et Yin [KY14].

Perturbation multiplicative de matrices de Wishart. Pour tout $n \geq 1$, soit $\Sigma_n \in \mathbf{R}^{n\times n}$ une matrice symétrique positive de taille $n \times n$, dont on notera $\gamma_1^{(n)}, \dots, \gamma_n^{(n)}$ les valeurs propres et $v_1^{(n)}, \dots, v_n^{(n)}$ les vecteurs propres normalisés associés. On suppose à nouveau que les valeurs propres sont bornées et que leur distribution empirique converge vers une mesure de probabilité μ_Σ :

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \xrightarrow[n\to+\infty]{} \mu_\Sigma,$$

au sens de la convergence étroite. On s'intéresse alors à la matrice de Wigner déformée

$$S_n := \frac{1}{n}\Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2}.$$

Dans ce contexte, S_n est la matrice de covariance d'un échantillon de m vecteurs i.i.d. de dimension n, dont la matrice de variance-covariance est donnée par Σ_n. Pour cette raison, on appelle parfois Σ_n la *matrice de population*. Soit $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ les valeurs propres de S_n et $\phi_1^{(n)}, \dots, \phi_n^{(n)}$ les vecteurs propres normalisés associés. Silverstein [Sil95] a identifié la limite des mesures spectrales empiriques de ce modèle en utilisant la méthode de la résolvante.

Théorème 14. *Presque-sûrement, la mesure spectrale empirique μ_{S_n} converge vers une mesure de probabilité déterministe $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$, caractérisée par l'équation*

$$\forall z \in \mathbf{C}_+, \quad s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu_\Sigma(\lambda)}{\lambda(\alpha - 1 - z s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(z)) - z}.$$

La notation $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ trouve son origine au sein de la théorie des probabilités libres et se lit "produit libre des mesures $\mu_{\mathrm{MP},\alpha}$ et μ_Σ".

Dans le contexte du modèle de Wishart déformé multiplicativement, Knowles et Yin [KY17] ont obtenu des lois locales analogues au Théorème 8. Le lecteur pourra consulter l'Équation (4.15) du Chapitre 4 pour une formulation précise.

On supposera désormais que $\gamma_1^{(n)} \equiv \theta$ pour un paramètre $\theta > 0$, et qu'aucune des autres valeurs propres de Σ_n n'est *atypique*, en ce sens que

$$\sup_{2\leq i\leq n} \mathrm{dist}\left(\gamma_i^{(n)}, \mathrm{Supp}(\mu_\Sigma)\right) \xrightarrow[n\to+\infty]{\mathbf{P}} 0.$$

Du point de vue de l'inférence statistique, on aimerait estimer le paramètre θ à partir de l'observation de S_n. Selon sa valeur, il peut être responsable de l'apparition d'une valeur propre atypique, un *outlier*, dans le spectre de W_n. On notera $F(x) = \alpha x - x - s_{\mu_{sc}\boxtimes\mu_\Sigma}(1/x)$.

Théorème 15. *Supposons qu'il existe $x_{\alpha,\theta} \notin \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$ tel que $1/F(1/x_{\alpha,\theta}) = \theta$. Soit $\delta > 0$ tel que $[x_\theta - \delta, x_\theta + \delta] \cap \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma) = \emptyset$. Alors pour tout n suffisamment grand, il existe une*

unique valeur propre de S_n dans l'intervalle $[x_{\alpha,\theta} - \delta, x_{\alpha,\theta} + \delta]$. En notant i_n l'indice de cette valeur propre, les deux convergences suivantes ont lieu en probabilité

$$\lambda_{i_n}^{(n)} \underset{n\to+\infty}{\longrightarrow} x_\theta \qquad et \qquad |\langle \phi_{i_n}^{(n)}, v_1^{(n)} \rangle|^2 \underset{n\to+\infty}{\longrightarrow} \frac{x_{\alpha,\theta} F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})}.$$

À l'inverse, si $1/F(1/x) = \theta$ n'admet pas de solution en dehors du support de $\mu_{sc} \boxtimes \mu_\Sigma$, le spectre de S_n n'admet pas de valeur propre atypique, en ce sens que

$$\sup_{1\le i\le n} \operatorname{dist}\left(\lambda_i^{(n)}, \operatorname{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)\right) \underset{n\to+\infty}{\overset{\mathbf{P}}{\longrightarrow}} 0.$$

Concernant la convergence de la valeur propre atypique, une première version de ce résultat a été obtenue par Rao et Silverstein [NS10] et par Bai et Yao [BY12]. Capitaine [Cap13] fût la première à obtenir la convergence de la projection du vecteur propre associé.

Le cas particulier $\Sigma_n = \operatorname{Diag}(\theta, 1, \dots, 1)$ a été introduit par Johnstone [Joh01b] et correspond en fait au premier modèle de matrices aléatoires déformées a avoir été considéré dans la littérature. Dans ce cas, la mesure spectrale empirique limite est la loi de Marchenko-Pastur puisque $\mu_{\mathrm{MP},\alpha} \boxtimes \delta_1 = \mu_{\mathrm{MP},\alpha}$.

Théorème 16. *Les convergences suivantes ont lieu en probabilité :*

$$\lambda_1^{(n)} \underset{n\to+\infty}{\longrightarrow} \begin{cases} b = (1+\sqrt{\alpha})^2 & si\ \theta \le 1 + \alpha^{-1/2}; \\ \frac{\theta(\alpha\theta - \alpha + 1)}{\theta - 1} > b & si\ \theta > 1 + \alpha^{-1/2}. \end{cases} \tag{1.21}$$

et

$$|\langle \phi_1^{(n)}, e_1 \rangle|^2 \underset{n\to+\infty}{\longrightarrow} \begin{cases} 0 & si\ \theta \le 1 + \alpha^{-1/2}; \\ \frac{1 - \frac{1}{\alpha(\theta-1)^2}}{1 + \frac{1}{\alpha(\theta+1)}} & si\ \theta > 1 + \alpha^{-1/2}. \end{cases} \tag{1.22}$$

La transition de phase de la plus grande valeur propre a été démontrée par Baïk, Ben Arous et Péché [BBAP05] dans le cas où les coefficients de X_n sont des gaussiennes complexes. Leur célèbre article correspond à la première occurrence d'un phénomène de transition de phase pour les spectres des matrices aléatoires déformées dans la littérature. Pour cette raison, les différentes transitions de phases concernant le spectres des matrices aléatoires déformées sont souvent appelées *transitions de phase BBP*. Dans ce même article, Baïk, Ben Arous et Péché étudient aussi les fluctuations de la plus grande valeur propre, lesquelles suivent une loi gaussienne dans le cas sur-critique, la loi de Tracy-Widom dans le cas sous-critique et une déformation de cette loi dans le cas critique. Finalement, mentionnons que la transition de phase à laquelle est assujetti le carré de la projection du plus grand vecteur propre dans la direction de la perturbation a d'abord été exhibée par Paul [Pau07] lorsque les coefficients de X_n sont des gaussiennes réelles.

Avant de poursuivre notre présentation et de discuter des vecteurs propres non-atypiques des matrices de Wigner perturbées additivement et des matrices de Wishart perturbées multiplicativement, nous tenons ici à mentionner l'existence de nombreux autres modèles de matrices aléatoires déformés. Parmi eux, on peut distinguer la famille des modèles invariants par l'action du groupe unitaire ; un exemple étant donné par les matrices de la forme $A_n + U_n B_n U_n^*$, où A_n et B_n sont déterministes et où U_n suit la loi de Haar sur le groupe unitaire. Nous renvoyons le lecteur curieux aux articles [BGN11, BBCF17] et à l'article d'exposition de Donati-Martin et Capitaine [CDM17] pour une bibliographie détaillée.

Vecteurs propres non-atypiques. Dans les modèles déformés que l'on vient de présenter, l'influence du paramètre θ est visible au niveau d'une valeur propre et de son vecteur propre associé, d'après les phénomènes de transitions de phases des Théorèmes 12 et 15. Plus généralement, peut-on étudier la corrélation $\langle \phi_i, v_j \rangle$ entre un vecteur propre de la matrice perturbée et un vecteur propre de la perturbation ? Une idée naturelle, d'abord proposée par Biane [Bia03], pour faire apparaître ces produits scalaires est d'introduire les quantités suivantes

$$\frac{1}{n}\mathbf{E}\left[\mathrm{Tr}\left((W_n - zI_n)^{-1} g(A_n)\right)\right] \qquad \text{et} \qquad \frac{1}{n}\mathbf{E}\left[\mathrm{Tr}\left((S_n - zI_n)^{-1} g(\Sigma_n)\right)\right],$$

où $g : \mathbf{R} \to \mathbf{R}$ est une fonction mesurable. En effet, celles-ci se ré-écrivent

$$\frac{1}{n}\sum_{i=1}^{n} \frac{|\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2}{\lambda_i^{(n)} - z} g\left(\gamma_i^{(n)}\right). \tag{1.23}$$

L'étude asymptotique de cette statistique a été réalisée par Péché et Ledoit [LP11] dans le cas des modèles de Wishart déformés multiplicativement, puis par Allez et Bouchaud [AB14] dans le cas des modèles de Wigner perturbés additivement. D'autres modèles déformés sont considérés avec le niveau de rigueur des physicens par Bun, Allez, Bouchaud et Potters dans l'article [BABP16], bien qu'une adaptation des arguments de [LP11] et [BABP16] semble possible comme le suggèrent les auteurs. Dans le cas particulier des fonctions indicatrices $g(\cdot) = \mathbf{1}_{\cdot \leq \gamma}$, Péché-Ledoit et Allez-Bouchaud obtiennent l'existence d'une fonction $\phi(\cdot, \cdot)$ telle que

$$\frac{1}{n}\sum_{i,j=1}^{n} |\langle \phi_i^{(n)}, v_i^{(n)} \rangle|^2 \mathbf{1}_{\lambda_i^{(n)} \leq \lambda} \mathbf{1}_{\gamma_i^{(n)} \leq \gamma} \underset{n \to +\infty}{\longrightarrow} \int_{-\infty}^{\lambda} \int_{-\infty}^{\gamma} \phi(x, y) \mathrm{d}\left((\mu_{sc} \boxplus \mu_A)(-\infty, x)\right) \mathrm{d}\left(\mu_A(-\infty, y)\right). \tag{1.24}$$

De manière informelle, au vu de l'Équation (1.24), $\phi(\lambda, \gamma)$ est une bonne approximation de la corrélation moyenne

$$\frac{1}{n^2} \sum_{i,j \in I(\lambda) \times J(\gamma)} n |\langle \phi_i^{(n)}, v_j^{(n)} \rangle|^2$$

où l'on somme sur une proportion *macroscopique* de vecteurs propres de W_n (resp. S_n) au voisinage de λ et de vecteurs propres de A_n (resp. Σ_n) au voisinage de γ. Cette limitation à des sommes sur des portions de tailles macroscopiques ne permet malheureusement pas d'obtenir d'information sur les valeurs propres et les vecteurs propres atypiques, qui sont par définition isolés dans le spectre. L'étude de la mesure spectrale empirique que l'on va maintenant présenter s'avère féconde pour pallier à ce problème.

Mesure spectrale dans la direction de la perturbation. La deuxième contribution de cette thèse concerne l'étude d'une statistique spectrale permettant d'éclairer différemment les résultats de transitions de phase qui précèdent, et d'obtenir de nouveaux résultats concernant les vecteurs propres non-atypiques des modèles déformés. Cette statistique est la *mesure spectrale dans la direction de la perturbation*, définie par

$$\mu_{(M_n, v_1^{(n)})} := \sum_{i=1}^{n} |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \delta_{\lambda_i^{(n)}}, \tag{1.25}$$

où $M_n = W_n$ ou S_n selon que l'on considère le modèle de Wigner perturbé additivement ou le modèle de Wishart perturbé multiplicativement. À la différence de la mesure spectrale

empirique, cette mesure conserve en un certain sens la direction favorisée par la perturbation, comme en témoigne par exemple l'expression de sa transformée de Stieltjes :

$$s_{\mu_{(M_n,v_1^{(n)})}}(z) = \langle v_1^{(n)}, (M_n - zI_n)^{-1} v_1^{(n)} \rangle.$$

L'analyse du terme de droite de l'égalité ci-dessus découle des lois locales générales obtenues par Knowles et Yin [KY17] au sujet des modèles matriciels déformés. Une formulation précise de leurs résultats est contenue dans les Équations (4.7) et (4.15). De manière informelle, il s'agit d'obtenir les *équivalents déterministes* suivants :

$$\begin{cases} (W_n - zI_n)^{-1} \approx \left(A_n - z - s_{\mu_{sc} \boxplus \mu_A}(z)\right)^{-1}, \\ (S_n - zI_n)^{-1} \approx -\left(z\left(1 + s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z) - \frac{\alpha-1}{z}\Sigma_n\right)\right)^{-1}. \end{cases}$$

Il est alors aisé d'obtenir le résultat suivant, qu'on énonce sous la forme d'un théorème.

Théorème C : Chapitre 4, Corollaires 4, 7 et Propositions 3, 5

Dans le cas des matrices de Wigner perturbées additivement, la mesure spectrale $\mu_{(W_n, v_1^{(n)})}$ converge étroitement, en probabilité, vers une mesure de probabilité $\mu_{sc,A,\theta}$ caractérisée par

$$s_{\mu_{sc,A,\theta}}(z) = \frac{1}{\theta - z - s_{\mu_{sc} \boxplus \mu_A}(z)}. \tag{1.26}$$

Dans le cas des matrices de Wishart perturbées multiplicativement, la mesure spectrale $\mu_{(S_n, v_1^{(n)})}$ converge étroitement, en probabilité, vers une mesure de probabilité $\mu_{\alpha,\Sigma,\theta}$ caractérisée par

$$s_{\mu_{\alpha,\Sigma,\theta}}(z) = \frac{1}{\theta(\alpha-1) - z s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z) - z}. \tag{1.27}$$

Dans le cas particulier des perturbations de rang 1, où $A_n = \theta e_1 e_1^T$ et $\Sigma_n = \mathrm{Diag}(\theta, 1, \ldots, 1)$, les mesures limites sont explicites, respectivement égales à :

$$\mu_{sc,\theta}(\mathrm{d}x) = \frac{\sqrt{4-x^2}}{2\pi(\theta^2 + 1 - \theta x)} \mathbf{1}_{|x| \leq 2} \mathrm{d}x + \mathbf{1}_{|\theta|>1} \left(1 - \frac{1}{\theta^2}\right) \delta_{\theta + \frac{1}{\theta}}(\mathrm{d}x), \tag{1.28}$$

et

$$\mu_{\mathrm{MP},\alpha,\theta}(\mathrm{d}x) = \frac{\theta\sqrt{(b-x)(x-a)}\mathbf{1}_{(a,b)}(x)}{2\pi x\left(x(1-\theta) + \theta(\alpha\theta - \alpha + 1)\right)} \mathrm{d}x + c_{\alpha,\theta}\delta_0(\mathrm{d}x) + d_{\alpha,\theta}\mathbf{1}_{|\theta-1|>\frac{1}{\sqrt{\alpha}}} \delta_{x_{\alpha,\theta}}(\mathrm{d}x), \tag{1.29}$$

où

$$c_\alpha = \frac{1-\alpha}{\alpha(\theta-1)+1}, \quad d_{\alpha,\theta} = \frac{1 - \frac{1}{\alpha(\theta-1)^2}}{1 + \frac{1}{\alpha(\theta-1)}} \quad \text{et} \quad x_{\alpha,\theta} = \frac{\theta(\alpha\theta - \alpha + 1)}{\theta - 1}.$$

Mentionnons ici que l'utilisation des lois locales de Knowles et Yin pourrait être contournée afin d'obtenir les convergences ponctuelles (1.26) et (1.27). Par exemple, dans le cas particulier où $v_1 = e_1$ est le premier vecteur de la base canonique et où $M_n = W_n$, la matrice déformée correspond à une matrice de Wigner classique à laquelle on a ajouté θ au coefficient $(1,1)$, et l'on peut utiliser la formule des compléments de Schur (1.2) que l'on a déjà présentée au cours de la preuve du Théorème de Wigner. Dans notre nouveau contexte, le coefficient $(W_n -$

$zI_n)_{11}$ converge vers $\theta - z$ lorsque n tend vers l'infini, et il est possible de démontrer que la forme quadratique $b^T C b$ possède la même limite que la transformée de Stieltjes de la mesure empirique μ_{W_n}, à savoir $s_{\mu_{sc} \boxplus \mu_A}(z)$. En revanche, l'utilisation des lois locales (4.7) et (4.15) sera essentielle pour obtenir le Théorème D à venir. Notons finalement que les mesures $\mu_{sc,\theta}$ et $\mu_{\mathrm{MP},\alpha,\theta}$ correspondent aux *lois de Meixner libres*, comme l'avait déjà remarqué Lenczewski [Len15] dans un contexte Gaussien.

Le Théorème C peut être considéré comme une généralisation des résultats de Péché et Ledoit (resp. Allez et Bouchaud) présentés précédemment puisque la quantité (1.23) n'est autre qu'une moyenne des transformées de Stieltjes des mesures spectrales de S_n (resp. W_n), poussées en avant par la fonction g. Ce résultat possède deux applications.

La première concerne l'étude de la transition de phase pour les modèles déformés. Rappelons les définitions des fonctions w et F, respectivement bien définies en dehors des supports de $\mu_{sc} \boxplus \mu_A$ et $\mu_{sc} \boxtimes \mu_\Sigma$:

$$w(x) = x + s_{\mu_{sc} \boxplus \mu_A}(x) \qquad \text{et} \qquad F(x) = \alpha x - x - s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(1/x),$$

et supposons, comme dans le Théorème 12, qu'il existe $x_\theta \notin \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ tel que $w(x_\theta) = \theta$. Alors l'Équation (1.26) entraîne que x_θ est un pôle de $s_{\mu_{sc,A,\theta}}$. Ceci correspond à un atome de $\mu_{sc,A,\theta}$ dont le poids est donné par l'opposé du résidu :

$$\mu_{sc,A,\theta}(\{x_\theta\}) = \lim_{z \to x_\theta} (x_\theta - z) s_{\mu_{sc,A,\theta}}(z) = \lim_{z \to x_\theta} \frac{x_\theta - z}{w(x_\theta) - w(z)} = \frac{1}{w'(x_\theta)}.$$

Dans un cas où $W_n - \theta v_1^{(n)} {v_1^{(n)}}^T$ ne possède pas de valeur propre atypique, ceci permet d'obtenir la première partie du Théorème 12. Un même raisonnement sur la mesure spectrale $\mu_{\alpha,\Sigma,\theta}$ permet d'obtenir la première partie du Théorème 15. Des énoncés précis font l'objet des Corollaires 5 et 8. On retiendra en particulier que les influences de la valeur propre atypique et du vecteur propre qui lui est associé sont présentes à l'échelle *macroscopique* au sein de la mesure spectrale dans la direction de la perturbation. Ceci diffère des approches proposées jusqu'ici pour l'étude des transitions de phases, qui reposaient sur une étude des fluctuations d'ordre $1/n$ de la mesure spectrale empirique.

Dans le cas particulier des perturbations de rang 1, les formules explicites de $\mu_{sc,\theta}$ et $\mu_{\mathrm{MP},\alpha,\theta}$ permettent de retrouver le cas sur-critique des Théorèmes 13 et 16 en remarquant que les positions et poids des atomes correspondent respectivement à la limite de la plus grande valeur propre et à la limite du carré de la projection du vecteur propre associé dans la direction e_1.

Une deuxième application de l'étude des mesures spectrales concerne les corrélations entre les vecteurs propres non-atypiques de W_n (resp. S_n) et v_1. Afin d'énoncer notre résultat, nous admettrons ici que les mesures $\mu_{sc} \boxplus \mu_A$ et $\mu_{sc,A,\theta}$ sont absolument continues par rapport à la mesure de Lebesgue, et nous noterons $f_{sc,A}$ et $f_{sc,A,\theta}$ leurs densités respectives. De même, les mesures $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ et $\mu_{\alpha,\Sigma,\theta}$ sont absolument continues par rapport à la mesure de Lebesgue sur $\mathbf{R}_+^*$, et nous noterons $f_{\alpha,\Sigma}$ et $f_{\alpha,\Sigma,\theta}$ leurs densités respectives.

Théorème D : Chapitre 4, Théorèmes 3, 5, 4 et 6

Soit $x \in \mathbf{R}$ tel que $f_{sc,A}(x) > 0$ dans le cas du modèle de Wigner déformé, et tel que $f_{\alpha,\Sigma}(x) > 0$ dans le cas du modèle de Wishart déformé. Pour tout $0 < \delta < 1/2$, il existe une suite de réels $(\varepsilon_n)_{n \geq 1}$ telle que $n^\delta / \sqrt{n} \ll \varepsilon_n \ll 1$ et telle que, en notant $\mathcal{I}_{\varepsilon_n}^{(n)}(x) := |\{1 \leq$

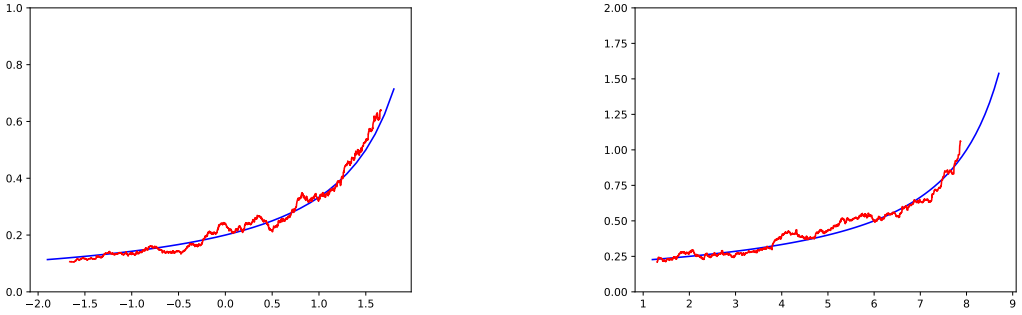
$i \leq n$, $|\lambda_i^{(n)} - x| \leq \varepsilon_n\}|$, la convergence suivante a lieu en probabilité :

$$\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}^{(n)}(x)} |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \underset{n \to +\infty}{\longrightarrow} \begin{cases} \frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)} & \text{dans le cas du modèle de Wigner,} \\ \frac{f_{\alpha,\Sigma,\theta}(x)}{f_{\alpha,\Sigma}(x)} & \text{dans le cas du modèle de Wishart.} \end{cases}$$

Dans le cas particulier des perturbations de rang 1 où $A_n = \theta e_1 e_1^T$ et $\Sigma_n = \mathrm{Diag}(\theta, 1, \ldots, 1)$, les formules sont explicites :

$$\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}^{(n)}(x)} |\langle \phi_i^{(n)}, e_1 \rangle|^2 \underset{n \to +\infty}{\longrightarrow} \begin{cases} \frac{1}{x^2 - \theta x + 1} & \text{dans le cas du modèle de Wigner,} \\ \frac{\theta}{x(1-\theta) + \theta(\alpha\theta - \alpha + 1)} & \text{dans le cas du modèle de Wishart.} \end{cases}$$

La Figure 1.3 ci-dessous illustre les formules explicites du Théorème D.



(a) Modèle de Wigner.

(b) Modèle de Wishart.

FIGURE 1.3 – En rouge : simulations des moyennes des carrés des projections des vecteurs propres associés à des valeurs propres dans le voisinage d'un point x fixé, où les moyennes sont réalisées sur des intervalles de taille typique $n^{0,1}/\sqrt{n}$. Chacune des simulations est réalisée à partir d'une unique matrice X_n, de taille 3000×3000 pour le modèle de Wigner et de taille 2000×8000 pour le modèle de Wishart ($\alpha = 4$). Dans chaque cas, $\theta = 2$.

Lorsque θ est une valeur propre à l'intérieur du support de μ_A (resp. μ_Σ), le Théorème D est une confirmation *microscopique* de la convergence (1.24) obtenue par Péché et Ledoit (resp. Allez et Bouchaud). Lorsque θ se situe à l'extérieur du support de μ_A (resp. μ_Σ), les moyennes macroscopiques réalisées par ces auteurs sur les corrélations $\langle \phi_i, v_i \rangle$ ne sont pas possibles, tandis que l'approche de la mesure spectrale fonctionne toujours.

La démonstration du Théorème D consiste à obtenir l'analogue des Théorèmes 5 et 9 dans le contexte des mesures spectrales des modèles matriciels déformés. Par exemple, dans le cas du modèle de Wigner perturbé additivement, en fixant $x \in \mathbf{R}$ tel que $f_{sc,A}(x) > 0$ et en notant $I_x(\varepsilon_n) = (x - \varepsilon_n, x + \varepsilon_n)$, on démontre que pour tout $\varepsilon > 0$, il existe $D > 0$ tel que,

$$\begin{cases} \mathbf{P}(|\mu_{(W_n, v_1^{(n)})}(I_x(\varepsilon_n)) - \mu_{sc,A,\theta}(I_x(\varepsilon_n))| \geq \omega_n) & \leq n^{-D}, \\ \mathbf{P}\left(|\mu_{W_n}(I_x(\varepsilon_n)) - \mu_{sc,A,\theta}(I_x(\varepsilon_n))| \geq \omega_n\right) & \leq n^{-D}, \end{cases}$$

où $(\omega_n)_{n \geq 1}$ est une suite bien choisie telle que $n^\delta/\sqrt{n} \ll \omega_n \ll \varepsilon_n$. Ces inégalités sont obtenues en utilisant la méthode de preuve du Théorème 5 laquelle, rappelons le, repose sur l'utilisation de la transformée de Helffer-Sjöstrand (1.8). Les détails de l'analyse correspondante font l'objet de la partie 4.5 du Chapitre 4. Ces inégalités en main, il reste alors à remarquer qu'avec probabilité au moins $1 - n^{-D}$,

$$\mu_{(W_n, v_1^{(n)})}(I_x(\varepsilon_n)) = 2\varepsilon_n f_{sc,A,\theta}(x) + \mathcal{O}(\omega_n),$$

et

$$\begin{aligned}\mu_{(W_n,v_1^{(n)})}(I_x(\varepsilon_n)) &= \sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)} |\langle \phi_i^{(n)}, v_1^{(n)}\rangle|^2 \\ &= \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2\right) \times \mu_{S_n}(I_{\varepsilon_n}(x)) \\ &= \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2\right) \times (2\varepsilon_n f_{sc,A}(x) + \mathcal{O}(\omega_n)).\end{aligned}$$

Nous allons maintenant nous tourner vers un calcul combinatoire des moments de la mesure $\mu_{\mathrm{MP},\alpha,\theta}$. Avant cela, mentionnons ici que l'étude des mesures spectrales a été exploitée dans d'autres contextes que les modèles matriciels déformés que l'on vient de considérer. Dans leur article [BGEM18], Benaych-Georges, Enriquez et Michaïl ont ainsi pu étudier l'évolution des vecteurs propres des perturbations de matrices diagonales. Par ailleurs, dans une série de travaux, Gamboa, Nagel et Rouault [GR11, GNR16, GNR17, GNR19] ont étudié les mesures spectrales de modèles matriciels solubles, dont le GUE fait partie. En explicitant de deux manières différentes la fonction de taux du principe de grande déviation vérifié par les mesures spectrales de tels modèles, les auteurs parviennent à obtenir une démonstration très élégante des célèbres *sum rules*, dont la compréhension est au cœur de certains travaux de théorie spectrale.

Un lien avec un modèle de physique statistique. Avec Nathanaël Enriquez, nous nous sommes aperçus que les moments de la mesure $\mu_{\mathrm{MP},1,\theta}$ sont reliés à la fonction de partition d'un modèle de physique statistique : l'accrochage de la marche aléatoire simple. Cette remarque d'apparence anodine est à l'origine d'une étude plus systématique des modèles d'accrochage généraux et des processus de renouvellement qui leur sont sous-jacents, que l'on présentera dans la partie suivante.

Il s'agit d'abord d'obtenir une formule combinatoire pour les moments de la mesure $\mu_{\mathrm{MP},1,\theta}$. Comme celle-ci est la limite faible des mesures spectrales $\mu_{(S_n,e_1)}$,

$$\forall N \geq 0, \quad \int_{\mathbf{R}} x^N \mathrm{d}\mu_{\mathrm{MP},\alpha,\theta}(x) = \lim_{n\to+\infty} \mathbf{E}\left[\int_{\mathbf{R}} x^N \mathrm{d}\mu_{(S_n,e_1)}(x)\right].$$

Par définition de $\mu_{(S_n,e_1)}$, le terme de droite de cette égalité est égal au coefficient $(1,1)$ de la matrice $\mathbf{E}[S_n^N]$. Notons $Y_n(i,j)$ le coefficient (i,j) de la matrice $\Sigma_n^{1/2}X_n$. Alors le moment d'ordre N de $\mu_{\mathrm{MP},\alpha,\theta}$ vérifie

$$\int_{\mathbf{R}} x^N \mathrm{d}\mu_{\mathrm{MP},\alpha,\theta}(x) = \lim_{n\to+\infty} \frac{1}{n^N} \sum_{\substack{1\leq i_2,\dots,i_N\leq n\\ 1\leq j_1,\dots,j_N\leq n}} \mathbf{E}\left[Y_n(1,j_1)Y_n(i_2,j_1)Y_n(i_2,j_2)\cdots Y_n(i_N,j_N)Y_n(1,j_N)\right].$$

Sans rentrer dans les détails, les contributions asymptotiques non nulles proviennent des mots $1j_1i_2j_2\cdots i_Nj_N1$ dont le graphe associé est un arbre possédant N arêtes. Ici, le nombre d'arêtes issues de la racine joue un rôle particulier puisque la variance de $Y_n(i,j)$ vaut θ si $i=1$ ou $j=1$. En notant $\mathscr{T}_N$ l'ensemble des arbres planaires enracinés possédant N arêtes et en exploitant les remarques précédentes, il est possible d'obtenir la formule suivante

$$\int_{\mathbf{R}} x^N \mathrm{d}\mu_{\mathrm{MP},1,\theta}(x) = \sum_{\mathrm{T}\in\mathscr{T}_N} \theta^{c_{\mathrm{T}}(\emptyset)}, \tag{1.30}$$

où $c_{\mathrm{T}}(\emptyset)$ désigne le nombre d'enfants de la racine de T. On utilise ensuite que les arbres planaires enracinés possédant N arêtes sont en bijection avec les excursions positives de longueurs $2N$ de la marche aléatoire simple via l'application de *contour*. Aussi, le moment d'ordre N de $\mu_{\mathrm{MP},1,\theta}$ peut s'écrire comme une somme sur les excursions positives de longueur $2N$ de la marche aléatoire simple, le poids associé à une excursion e étant donné par $\theta^{n_e(0)}$, où $n_e(0)$ désigne le nombre de retours en 0 de l'excursion e. Finalement, en remarquant qu'une excursion positive e donne lieu à $2^{n_e(0)}$ ponts de longueur $2N$ de la marche aléatoire simple, (1.30) se réécrit

$$\int_{\mathbf{R}} x^N \mathrm{d}\mu_{\mathrm{MP},1,\theta}(x) = \sum_{b \in \mathscr{B}_N} \left(\frac{\theta}{2}\right)^{n_e(0)}, \tag{1.31}$$

où $\mathscr{B}_N$ est l'ensemble des ponts de longueur $2N$. Sans risque de confusion, on notera encore $\mathbf{P}$ la loi de la marche aléatoire simple et $\mathbf{E}$ l'espérance sous $\mathbf{P}$. En posant $\beta = \log(\theta/2)$, on vient de démontrer que

$$Z_{\beta,N} := \mathbf{E}\left[\exp\left(\beta \sum_{i=1}^{2N} \mathbf{1}_{S_i=0}\right) \mathbf{1}_{S_{2N}=0}\right] = \int_{\mathbf{R}} \left(\frac{x}{4}\right)^N \mathrm{d}\mu_{\mathrm{MP},1,2\,\mathrm{e}^{\beta}}(x). \tag{1.32}$$

Le terme de gauche de l'égalité ci-dessus est la fonction de partition du modèle d'accrochage associé à la marche aléatoire simple. Pour tout $\beta \in \mathbf{R}$ et tout $N \geq 0$, ce modèle est décrit par la mesure de probabilité $\mathbf{P}_{\beta,N}$ sur les ponts de longueur $2N$, définie par

$$\frac{\mathrm{d}\mathbf{P}_{\beta,N}}{\mathrm{d}\mathbf{P}}(S) = \frac{1}{Z_{\beta,N}} \exp\left(\beta \sum_{i=1}^{2N} \mathbf{1}_{S_i=0}\right) \mathbf{1}_{S_{2N}=0}. \tag{1.33}$$

Le paramètre β pénalise (ou récompense, selon les points de vues) les retours en zéro et il est naturel d'étudier le temps moyen passé par le polymère le long de l'axe des abscisses. Celui-ci peut se calculer via la fonction de partition puisque

$$\frac{1}{2N}\frac{\partial Z_{\beta,N}}{\partial \beta} = \frac{1}{2N}\mathbf{E}_{\beta,N}\left[\sum_{i=1}^{2N} \mathbf{1}_{S_i=0}\right]. \tag{1.34}$$

Lorsque N tend vers l'infini, cette quantité converge vers la dérivée de l'*énergie libre* du modèle, définie par la limite thermodynamique suivante,

$$F(\beta) := \lim_{N \to +\infty} \frac{1}{2N} \log Z_{N,\beta}. \tag{1.35}$$

Il est possible de calculer explicitement cette fonction en utilisant l'Équation (1.32) et la définition de la mesure $\mu_{\mathrm{MP},1,2\,\mathrm{e}^{\beta}}$ donnée par (1.29). Après un petit calcul, on obtient :

$$F(\beta) = \begin{cases} \log\left(\sqrt{\frac{\mathrm{e}^{2\beta}}{2\,\mathrm{e}^{\beta}-1}}\right) & \text{si } \beta > 0, \\ 0 & \text{si } \beta \leq 0. \end{cases} \tag{1.36}$$

Au vu de (1.34), ceci entraîne que le modèle d'accrochage associé à la marche simple présente la transition de phase suivante :

- lorsque $\beta > 0$, le temps moyen passé sur l'axe des abscisses est strictement positif et le régime est dit *localisé* ;
- lorsque $\beta \leq 0$, le temps moyen passé sur l'axe des abscisses est nul et le régime est dit *délocalisé*.

Le lecteur aura remarqué l'analogie entre la transition de phase concernant la plus grande valeur propre d'une matrice de Wishart perturbée multiplicativement – qu'on a énoncée dans le Théorème 16 – et la transition de phase pour le modèle de polymère que l'on vient d'introduire. Ce lien est rendu possible grâce à la relation (1.32). Dans le régime sur-critique où $\beta > 0$, l'énergie libre correspond au logarithme de la limite de la plus grande valeur propre du modèle de matrices de covariance déformées. De plus, il est aisé de remarquer que le pré-facteur du terme exponentiel de l'équivalent asymptotique de $Z_{\beta,N}$ est donné par le carré de la projection du plus grand vecteur propre dans la direction e_1. Aussi,

$$\forall \beta > 0, \quad Z_{\beta,N} \underset{N\to+\infty}{\sim} \frac{1 - \frac{1}{(2\,\mathrm{e}^{\beta}-1)^2}}{1 + \frac{1}{(2\,\mathrm{e}^{\beta}-1)}} \left(\frac{\mathrm{e}^{2\beta}}{2\,\mathrm{e}^{\beta}-1}\right)^{N}. \tag{1.37}$$

Le modèle d'accrochage de la marche aléatoire simple que l'on vient de présenter est l'avatar d'une famille plus générale de modèles d'accrochage dont on donnera une définition précise dans la partie qui suit. De manière informelle, il s'agit de remplacer les temps de retour en zéro de la marche simple par un processus de renouvellement général, c'est-à-dire un ensemble aléatoire d'entiers dont les écarts consécutifs sont des variables aléatoires i.i.d. Ces modèles ont été étudiés en détails et l'on renvoie le lecteur à l'ouvrage de référence de Giacomin [Gia07] pour un exposé complet. En particulier, l'identification du point critique $\beta = 0$, l'expression de l'énergie libre (1.36), ou encore le calcul de l'équivalent asymptotique (1.37) étaient déjà accessibles. L'approche que nous venons de proposer diffère des méthodes classiques et se fonde sur la représentation (1.32) des fonctions de partitions comme moments d'une mesure de probabilité explicite qui, à notre connaissance, n'existait pas dans la littérature. Ceci nous a incité à étudier, toujours en collaboration avec Nathanaël Enriquez, la possible existence d'une famille de modèles d'accrochage dont les fonctions de partition admettraient une représentation analogue. Dans la partie suivante, nous allons voir qu'une telle famille existe et correspond à une classe soluble de processus de renouvellement.

1.5 Une classe soluble de processus de renouvellement et ses applications

De manière informelle, les processus de renouvellement modélisent des temps de réalisation de certains événements[3] en faisant l'hypothèse que leurs espacements sont des variables aléatoires i.i.d. Ces processus sont au cœur de divers objets probabilistes et ont été étudiés en profondeur par de nombreux mathématiciens. Nous renvoyons par exemple à l'ouvrage de Feller [Fel71, chap. XI] pour une introduction détaillée. Avec Nathanaël Enriquez, nous avons identifié une classe soluble de processus de renouvellement discrets, en ce sens que pour tout entier $N \geq 1$, la probabilité qu'un évènement ait lieu au temps N est donnée par le moment d'ordre N d'une mesure de probabilité explicite. Par ailleurs, nous avons démontré que les fonctions de partitions des modèles d'accrochage décrits par ces processus de renouvellement s'écrivent aussi comme les moments de mesures de probabilités explicites. Nos méthodes s'appliquent également dans le contexte des processus de renouvellement continus, où nous démontrons aussi l'existence d'une classe soluble. Ces résultats font l'objet du Chapitre 5, issu de l'article [EN20].

Définitions et notations. Soit K une mesure de probabilité sur $\{1, 2, \ldots\} \cup \{\infty\}$ et $(\eta_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi K. Cette suite jouera le rôle des temps d'inter-arrivées :

[3] on pensera par exemple à des séïsmes.

on définit $\tau_0 = 0$ et pour tout $n \geq 1$, $\tau_n = \sum_{1 \leq i \leq n} \eta_i$. L'ensemble des temps de renouvellement est alors défini par l'ensemble aléatoire $\tau := \{\tau_i,\, i \geq 0\}$. Sans risque de confusion, on notera $\mathbf{P}$ la loi de τ.

Notons $m_K := \sum_{i \geq 0} iK(i)$. Si le processus formé par les points de τ était stationnaire, la fréquence d'apparition des temps de renouvellement serait exactement égale à $1/m_K$. Ce n'est pas exactement le cas ici puisqu'on a imposé $0 \in \tau$. Il est en revanche possible de coupler τ avec sa version stationnaire et l'on obtient alors le Théorème du renouvellement qui établie la convergence suivante :

$$\mathbf{P}(N \in \tau) \underset{N \to +\infty}{\longrightarrow} \frac{1}{m_K}. \tag{1.38}$$

Bien entendu, l'étude des processus de renouvellement ne s'arrête pas à ce résultat : il est par exemple possible d'identifier le terme de reste dans la convergence (1.38) – voir [Rog73, Don97], ou encore d'étudier le processus des *temps de vie restants* $(\inf\{\tau_i,\, \tau_i \geq n\} - n)_{n \geq 0}$. Nous renvoyons le lecteur à l'ouvrage [Asm03] pour une exposition détaillée.

Une classe soluble. On suppose désormais qu'il existe une mesure de probabilité μ sur $[0,1]$ telle que

$$\begin{cases} K(n) = \int_0^1 (1-x)^{n-1} x \mathrm{d}\mu(x), \\ K(\{\infty\}) = \mu(\{0\}). \end{cases} \tag{1.39}$$

On dit que K est un mélange de lois géométriques et on appelle μ la mesure de mélange. Dans ce contexte, les probabilités d'appartenance au processus de renouvellement associé à K sont explicites.

Théorème E : Chapitre 5, Théorème 7

Il existe une unique mesure de probabilité ν sur $[0,1]$ telle que

$$s_\nu(z) s_\mu(1-z) = \frac{1}{z(1-z)}. \tag{1.40}$$

De plus, pour tout $N \geq 0$,

$$\mathbf{P}(N \in \tau) = \int_0^1 x^N \mathrm{d}\nu(x). \tag{1.41}$$

Mentionnons que dans le cas particulier où la mesure μ possède une densité continue, Nagaev [Nag15] avait déjà obtenu une représentation en moments des quantités $\mathbf{P}(N \in \tau)$. Cependant, les méthodes qu'il utilise se fondent sur des calculs analytiques directs et ne permettent pas d'obtenir un lien précis entre les mesures μ et ν de la forme de (1.40). Notre approche contourne ces calculs analytiques et consiste à démontrer que l'application du demi-plan complexe

$$S(z) := -\frac{1}{z} \sum_{N \geq 0} \mathbf{P}(N \in \tau) \frac{1}{z^N}$$

est la transformée de Stieltjes d'une mesure de probabilité. À cet effet, on exploite l'équation dite du renouvellement

$$\sum_{N \geq 0} \mathbf{P}(N \in \tau) z^N = \frac{1}{1 - \mathbf{E}[z^{\tau_1}]}$$

pour établir que $S(z)^{-1} = z(1-z) s_\mu(1-z)$. Cette formule permet alors de vérifier que l'application S préserve le demi-plan supérieur complexe et que $S(z) \sim -1/z$ lorsque $|z| \to +\infty$. Par un

théorème abstrait d'analyse complexe, ces propriétés entraînent finalement que S se représente comme la transformée de Stieltjes d'une mesure de probabilité.

Notons que la relation (1.40) définit une involution $\mu \mapsto \nu$ et que les lois de l'Arcsinus généralisées, définies pour tout $v \in (0,1)$ par

$$\mu_v(\mathrm{d}x) := \frac{\sin(\pi v)}{\pi} x^{-v}(1-x)^{v-1}\mathbf{1}_{v\in(0,1)}\mathrm{d}x,$$

sont des points fixes de cette involution, puisque leurs transformées de Stieltjes valent $z^{-v}(1-z)^{v-1}$. Nous conjecturons que ce sont les seuls points fixes, ce qui donnerait lieu à une caractérisation de ces lois.

Mentionnons finalement qu'il est possible de retrouver le Théorème du renouvellement (1.38) dans notre contexte. En effet, l'application du Théorème de convergence dominé à l'égalité (1.41) entraîne que $\mathbf{P}(N \in \tau)$ converge vers l'atome de ν en 1. Ce dernier est égal à l'opposé du résidu de s_ν en 1, lui même égal à $s_\mu(0)^{-1} = m_K^{-1}$ via la relation (1.40).

Le cas des processus continus. De manière intéressante, les idées développées précédemment s'adaptent aisément au cadre des processus de renouvellement continus, où les variables aléatoires $(\eta_n)_{n\geq 0}$ sont positives et admettent une densité f_η sur $\mathbf{R}_+$. Notre résultat stipule que, lorsque cette densité f_η est un mélange de lois exponentielles, l'intensité du processus de renouvellement se représente comme la transformée de Laplace d'une mesure positive explicitement reliée à la mesure de mélange.

Théorème F : Chapitre 5, Théorème 8

Supposons qu'il existe une mesure de probabilité μ sur $[0,+\infty)$ telle que, pour tout $x > 0$:

$$f_\eta(x) = \int_0^{+\infty} s\,\mathrm{e}^{-sx}\,\mathrm{d}\mu(s).$$

Soit $H(x) := \sum_{k\geq 1} f_{\eta_1+\cdots+\eta_k}(x)$ l'intensité du processus de renouvellement dont les inter-arrivées sont les variables aléatoires $(\eta_i)_{i\geq 1}$.
Alors il existe une unique mesure positive ν sur $[0,+\infty)$ telle que

$$(1+s_\nu(z))s_\mu(z) = -\frac{1}{z}. \tag{1.42}$$

De plus, pour tout $x > 0$,

$$H(x) = \int_0^{+\infty} e^{-xs}\mathrm{d}\nu(s).$$

À première vue, les équations reliant les mesures μ et ν (à savoir (1.40) et (1.42)) peuvent paraître abstraites, et l'on peut légitimement se demander si la mesure ν peut être *explicitement* décrite. Cela est effectivement le cas, au moins lorsque la mesure μ est purement atomique ou absolument continue vis à vis de la mesure de Lebesgue : nous renvoyons le lecteur aux Corollaires 10 et 11 pour plus de détails.

Modèles d'accrochages solubles. Soit $\beta \in \mathbf{R}$. On considère le processus de renouvellement *discret* dont la loi des inter-arrivées est K, définie au cours de l'Équation (1.39). Le polymère de taille N associé au processus de renouvellement τ est défini par la mesure de probabilité $\mathbf{P}_{\beta,N}$

sur les sous-ensembles de $\{0, \dots, N\}$ dont la densité par rapport à la mesure $\mathbf{P}$ est

$$\frac{\mathrm{d}\mathbf{P}_{\beta,N}}{\mathrm{d}\mathbf{P}}(S) = \frac{1}{Z_{\beta,N}} \exp\left(\beta \mathcal{N}_N(\tau)\right) \mathbf{1}_{N\in\tau}, \tag{1.43}$$

où $\mathcal{N}_N(\tau) = |\{1, \dots, N\} \cap \tau|$ et

$$Z_{\beta,N} := E_{\mathbf{P}}\left[\exp\left(\beta \mathcal{N}_N(\tau)\right) \mathbf{1}_{N\in\tau}\right]. \tag{1.44}$$

Comme dans le cas du modèle d'accrochage de la marche aléatoire simple présenté dans la partie précédente, la fonction de partition $Z_{N,\beta}$ joue un rôle essentiel. Nous renvoyons à la partie 5.3 du Chapitre 5 pour plus de détails, et à l'ouvrage de référence de Giacomin [Gia07] pour un exposé complet. Des idées de preuve similaires à celles du Théorème E permettent de représenter les fonctions de partitions comme les moments d'une mesure de probabilité explicite.

Théorème G : Chapitre 5, Théorème 9

Pour tout $\beta \in \mathbf{R}$, il existe une unique mesure de probabilité ν_β sur $\mathbf{R}$ telle que

$$s_{\nu_\beta}(z)\left(\mathrm{e}^{\beta}\, s_\mu(1-z) - \frac{1-\mathrm{e}^{\beta}}{1-z}\right) = \frac{1}{z(1-z)}. \tag{1.45}$$

De plus, pour tout $N \geq 0$,

$$Z_{N,\beta} = \int_{\mathbf{R}} x^N \mathrm{d}\nu_\beta(x). \tag{1.46}$$

Lorsque μ est une loi de l'Arcsinus généralisée, tous les calculs peuvent être menés à terme et nous renvoyons le lecteur au Théorème 8 du Chapitre 5 pour un énoncé précis.

Chapitre 2

Graphes aléatoires

Dans leur célèbre article *On the evolution of random graphs* (1960), Erdős et Rényi proposent la première étude d'un modèle de graphes aléatoires portant désormais leurs noms, établissant ainsi les fondations d'un domaine de recherche à la frontière entre probabilité et combinatoire encore extrêmement actif aujourd'hui. Nous nous intéressons ici à ce modèle historique et à l'une de ses généralisations, le *modèle de configuration*, introduit par Bollobás en 1980. La richesse de ces deux modèles contraste avec la simplicité de leurs définitions, et a motivé leur étude approfondie au cours de ces soixante dernières années. Nous renvoyons aux ouvrages de références [Bol01, Dur10, FK16, vdH17] pour un traitement détaillé.

Afin de familiariser le lecteur avec les graphes d'Erdős-Rényi et les modèles de configuration, nous commencerons par en proposer une description géométrique locale. Dans le régime *dilué* où le degré moyen des sommets est fini, les paysages locaux typiques de ces modèles sont des arbres, lesquels convergent vers des arbres de Galton-Watson lorsque le nombre de sommets tend vers l'infini. Cette convergence locale s'avère extrêmement féconde dans la compréhension heuristique de nombreux problèmes : on s'intéressera en particulier à la possible existence d'une composante connexe de taille macroscopique. Lorsqu'une telle composante géante existe, l'algorithme d'exploration en profondeur permet d'établir l'existence de chemins simples dont la longueur est linéaire en fonction du nombre de sommets du graphe. Dans le cas du modèle de configuration, l'analyse complète de cet algorithme fait l'objet du Chapitre 6 de cette thèse, issu d'un article réalisé en collaboration avec Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard [EFMN19]. Une adaptation des méthodes développées dans ce travail permet d'analyser une variante de l'algorithme du parcours en profondeur et d'établir en particulier l'existence de longs chemins induits dans les modèles de configuration. Nous discuterons brièvement des résultats associés, qui font l'objet du Chapitre 7 issu de l'article [EFMN20]. Sans transition, nous présenterons finalement une deuxième application du Théorème A pour l'étude des spectres de graphes d'Erdős-Rényi bipartis.

Un peu de vocabulaire sur les graphes. Avant d'aller plus loin, il convient de fixer quelques notations. Rappelons qu'un graphe $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ est la donnée d'un ensemble de sommets V et d'un ensemble d'arêtes $\mathrm{E} \subset \{\{u, v\} \in \mathrm{V} \times \mathrm{V},\ u \neq v\}$. L'ensemble des sommets d'un graphe est toujours muni d'une distance naturelle induite par la longueur des chemins au plus proche voisin. On dit que H est un sous-graphe de G et l'on note $H \subset G$ lorsque $\mathrm{V}(\mathrm{H}) \subset \mathrm{V}$ et $\mathrm{E}(\mathrm{H}) \subset \mathrm{E}$. Par ailleurs, étant donné un sous ensemble de sommets $\mathrm{V}' \subset \mathrm{V}$, on appelle graphe induit par V' le graphe formé par les sommets de V' et leurs arêtes à l'intérieur du graphe G.

Un graphe enraciné (G, ρ) est la donnée d'un graphe et de l'un de ses sommets, appelé racine. Pour tout $k \geq 0$, on notera $(\mathrm{G}, \rho)_k$ le graphe induit par les sommets à distance inférieure ou

égale à k de la racine. Deux graphes enracinés (G_1, v_1) et (G_2, v_2) sont dits isomorphes lorsqu'il existe une bijection $\phi : V(G_1) \to V(G_2)$ telle que $\phi(v_1) = v_2$ et

$$\forall u, v \in V(G_1), \quad \{u, v\} \in E(G_1) \quad \Leftrightarrow \quad \{\phi(u), \phi(v)\} \in E(G_2).$$

Avec un léger abus de notation, on notera $(G_1, v_1) = (G_2, v_2)$.

2.1 Le modèle d'Erdős-Rényi

Définition du modèle. En 1959, Gilbert [Gil59] propose l'étude d'un graphe aléatoire $\mathrm{ER}(N, p)$ possédant N sommets numérotés de 1 à N, et dont chaque paire de sommets est une arête du graphe avec une probabilité $p \in (0, 1)$ fixée, indépendamment des autres paires. La motivation de Gilbert, alors membre des *Bell Telephone Laboratories*, était de modéliser simplement un réseau de N téléphones. Dans son article, il montre que la probabilité que le réseau soit connecté est équivalente à $1 - N(1-p)^{N-1}$ lorsque N tend vers l'infini.

Un an après, Erdős et Rényi [ER60] s'intéressent aux propriétés géométriques d'un graphe aléatoire dont la loi est uniforme parmi les graphes ayant N sommets et M arêtes. Ce modèle correspond peu ou prou au modèle de Gilbert lorsque $p = M/\binom{N}{2}$, puisque le nombre d'arêtes de $\mathrm{ER}(N, p)$ suit une loi Binomiale de paramètres $\binom{N}{2}$ et p.

La profondeur des idées de l'article [ER60], et en particulier la découverte de l'existence d'une transition de phase pour l'existence d'une composante connexe de taille macroscopique, a souvent occulté l'article de Gilbert et l'on appelle aujourd'hui indifféremment *graphe d'Erdős-Rényi* chacun des deux modèles équivalents que l'on vient d'introduire. D'un point de vue probabiliste, il sera pourtant plus commode d'étudier le modèle historique de Gilbert. La raison principale de ce choix réside dans la possibilité de construire $\mathrm{ER}(N, p)$ de manière dynamique, en parcourant successivement chaque paire de sommets et en décidant de les relier par une arête avec probabilité p. En particulier, partant d'un sommet v fixé du graphe, deux types d'explorations se distinguent :

- les explorations en largeur, qui révèlent successivement les boules centrées en v dans le graphe, informent sur le paysage local du graphe autour de v ;
- les explorations en profondeur et leurs variantes, qui informent entre autre sur l'existence de longs chemins simples dans le graphe. Dans le contexte des modèles de configurations, l'étude approfondie de ces algorithmes fait l'objet de deux articles de cette thèse, présentés dans les Chapitres 6 et 7.

Dans le modèle d'Erdős-Rényi, la probabilité d'occurrence d'un graphe G ayant N sommets numérotés et M arêtes est donnée par

$$\mathbf{P}(\mathrm{ER}(N, p) = G) = p^M (1-p)^{\binom{N}{2} - M}.$$

Nous supposerons désormais que le degré moyen des sommets est constant, c'est-à-dire qu'il existe une constante $c > 0$ telle que $p = c/N$. Du point de vue de la modélisation des réseaux sociaux, ce choix s'avère pertinent : on pensera par exemple à Facebook où un utilisateur typique possède quelques dizaines d'amis parmi plus de deux milliards d'utilisateurs.

Géométrie locale. Soit v un sommet de $\mathrm{ER}(N, p)$ choisi uniformément. Quelle est la géométrie du graphe au voisinage de v ? Afin de répondre à cette question, on peut réaliser un parcours

en largeur de la composante connexe de v, ce qui consiste à révéler successivement les boules de rayon k, $k \geq 1$.

La première étape de l'algorithme correspond donc la découverte des voisins de v dans le graphe, que l'on notera $u_1, \ldots, u_{\deg v}$, où $\deg v$ désigne le degré de v. Ce dernier suit une loi Binomiale de paramètres $N-1$ et c/N, ce qui entraîne que lorsque N tend vers l'infini, le nombre de sommets découverts suit donc une loi de Poisson de paramètre c. Notons enfin que la probabilité qu'une arête existe entre deux sommets u_i et u_j voisins de v est d'ordre $1/N$. Ainsi, lorsque N tend vers l'infini, la boule de rayon 1 centrée en v dans le graphe est un arbre enraciné de profondeur 1, dont le nombre d'enfants de la racine suit une loi de Poisson de paramètre c.

La seconde étape de l'algorithme consiste à découvrir les voisins de $u_1, \ldots, u_{\deg v}$ parmi les sommets que l'on n'a pas encore explorés. Cela peut être réalisé de manière séquentielle, en révélant d'abord les voisins de u_1, puis ceux de u_2, etc. Puisque $\deg v = o_{\mathbf{P}}(N)$, le nombre de sommets non-explorés est d'ordre N et il est aisé de se convaincre que pour tout $1 \leq i \leq \deg v$, le nombre de voisins de u_i parmi les sommets non-explorés suit asymptotiquement une loi de Poisson de paramètre c. De plus, la probabilité qu'un cycle soit créé parmi les sommets visités est encore d'ordre $1/N$. Ces considérations entraînent que la boule de rayon 2 centrée en v dans le graphe est asymptotiquement isomorphe à un arbre enracinée de profondeur 2, où les nombres d'enfants des individus sont indépendants et suivent chacun une loi de Poisson de paramètre c.

L'arbre généalogique qui semble émerger de cette exploration est un arbre de Galton-Watson, dont on donne ici une définition formelle.

Definition 1. *Soit $\boldsymbol{\pi}$ une loi de probabilité sur $\mathbf{N}$. Un arbre de Galton-Watson de loi de reproduction $\boldsymbol{\pi}$, noté $\mathrm{GW}(\boldsymbol{\pi})$, est une variable aléatoire dans l'espace des arbres planaires enracinés[1] défini par les deux propriétés suivantes :*

- *le nombre d'enfants de la racine $c_\emptyset$ est de loi $\boldsymbol{\pi}$;*
- *conditionnellement à $c_\emptyset = i$, les arbres induits par les i enfants de la racine sont des arbres de Galton-Watson indépendants et de loi de reproduction $\boldsymbol{\pi}$.*

En menant à terme le raisonnement amorcé ci-dessus, il est alors possible d'obtenir le résultat suivant.

Théorème 17. *Soit $\mathrm{GW}(\mathcal{P}(c))$ un arbre de Galton-Watson dont la loi de reproduction est une loi de Poisson de paramètre c. Alors, pour tout $k \geq 1$, la convergence suivante a lieu au sens de la convergence faible de variables aléatoires :*

$$(\mathrm{ER}(N, c/N), v)_k \underset{N \to +\infty}{\longrightarrow} \mathrm{GW}(\mathcal{P}(c))_k.$$

Mentionnons ici qu'il est en fait possible de coupler l'exploration en largeur du voisinage d'un sommet fixé $v \in \mathrm{ER}(N, c/N)$ avec celle de la racine d'un arbre de Galton-Watson jusqu'à des boules dont les rayons sont d'ordre $o(\log N)$. Cette borne correspond en fait à la profondeur typique à partir de laquelle des cycles sont révélés au voisinage du sommet v. En effet, en notant k_c la première valeur de k telle qu'un cycle est révélé, on s'attend à ce que k_c soit tel que la boule de rayon k_c et centrée en v possède un nombre de sommets d'ordre N. Or, cette boule possède en moyenne c^{k_c} sommets, ce qui entraîne que $c^{k_c} \approx N$ et partant, $k_c \approx \log N$.

[1] rappelons au lecteur qu'une définition formelle de ces objets a été donnée dans la Définition 1.

Existence d'une composante géante. Quelle est la taille des composantes connexes d'un graphe d'Erdős-Rényi dont le degré moyen est fini ? Du point de vue de la modélisation des réseaux sociaux, on aimerait qu'il existe au moins une composante connexe dont le nombre de sommets est proportionnel au nombre de sommets du graphe. Un regard sur les simulations de la Figure 2.1 semble indiquer l'existence d'une telle composante lorsque c est suffisamment grand. Le cas échéant, est-il possible d'identifier un paramètre critique ? De calculer la proportion de sommets appartenant à la composante géante ?

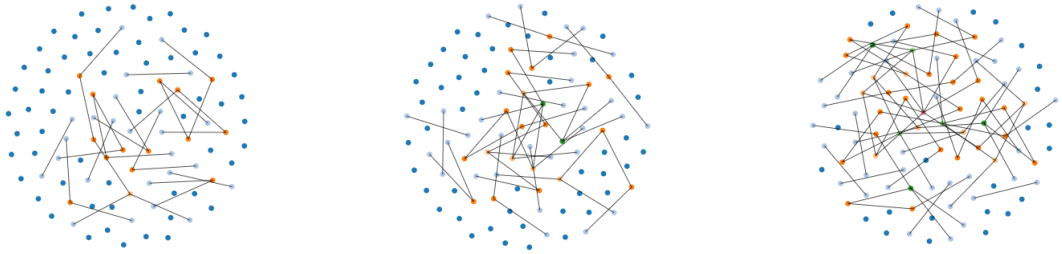


Figure 2.1 – Simulations de graphes d'Erdős-Rényi $\mathrm{ER}(100, c/100)$. De gauche à droite $c = 0.5, 1$ et $1,5$.

En étant *a priori* moins ambitieux, on peut d'abord se demander si, étant donné un sommet v du graphe, l'arbre de Galton-Watson obtenu comme limite locale autour de v survit avec probabilité positive. Ici, le paramètre $c = 1$ joue un rôle très particulier puisque :

- si $c \leq 1$, l'arbre de Galton-Watson $\mathrm{GW}(\mathcal{P}(c))$ s'éteint presque-sûrement,
- à l'inverse, si $c > 1$, cet arbre survit avec une probabilité positive $\rho_c > 0$.

En notant $f_c(s) = \exp(c(e^x - 1))$ la série génératrice des moments de la loi de Poisson de paramètre c, cette probabilité de survie est l'unique point fixe dans l'intervalle $(0, 1]$ de l'équation

$$1 - x = f_c(1 - x).$$

De manière remarquable, cette simplification du problème initial permet en fait d'identifier le bon paramètre critique.

Théorème 18. *Notons $\mathcal{C}_{\max}$ la composante connexe de $\mathrm{ER}(N, c/N)$ possédant le plus grand nombre de sommets. On dispose de la dichotomie suivante:*

- *si $c < 1$, alors $|\mathcal{C}_{\max}| = \mathcal{O}_{\mathbf{P}}(\log N)$ i.e. $\lim_{\kappa \to +\infty} \limsup_{N \to +\infty} \mathbf{P}(|\mathcal{C}_{\max}| \geq \kappa \log N) = 0$,*
- *si $c > 1$, alors $\frac{|\mathcal{C}_{\max}|}{N} \to \rho_c$ en probabilité lorsque $N \to +\infty$. De plus, les tailles des autres composantes connexes sont d'ordre $\mathcal{O}_{\mathbf{P}}(\log N)$.*

Le lecteur pourra consulter [vdH17] pour une preuve de ce résultat utilisant les méthodes du premier moment et du second moment. Nous allons esquisser une autre démonstration basée sur l'exploration de Łukasiewicz, très proche dans les idées de l'heuristique de la convergence locale. L'exploration consiste, à l'étape $n \geq 1$, à choisir uniformément un sommet u_n exploré et actif et à révéler ses voisins dans la partie du graphe non-explorée. Ceci étant fait, on déclare u_n exploré et retraité, et on déclare ses enfants explorés et actifs. On note X_n le nombre de voisins révélés et on considère la *marche de Łukasiewicz* S_n associée à l'exploration, définie par

$$S_n := \sum_{i=1}^{n} (X_n - 1),$$

et dont les premiers pas sont $S_0 = 0$ et $S_1 = \deg(v) - 1$. Cette marche aléatoire est positive et vaut -1 lorsqu'on termine l'exploration. Aussi, le premier temps d'atteinte de -1 est égal à la taille de la composante connexe de v. Afin d'étudier ce temps d'arrêt, il est possible de coupler (S_n) avec une marche (S'_n) dont les pas sont des variables i.i.d. de loi $\mathcal{P}(c) - 1$, qui correspond à la marche de Łukasiewicz associée à l'exploration d'un arbre de Galton-Watson de loi de reproduction $\mathcal{P}(c)$. De manière qualitative, la dichotomie du Théorème 18 correspond alors aux deux phénomènes suivants :

- lorsque $c < 1$, la marche aléatoire S'_n possède une dérive négative et sa première excursion positive est de taille finie ;
- lorsque $c > 1$, la marche aléatoire S'_n possède une dérive positive, donc la première excursion positive est de taille infinie avec probabilité positive.

Dans le cas sur-critique où $c > 1$, le fait que les composantes connexes différentes de $\mathscr{C}_{\max}$ sont de tailles logarithmiques provient de l'absence d'excursion positive de taille intermédiaire pour une marche aléatoire avec dérive positive.

Le monde est petit. Que peut-on dire des distances typiques au sein de la composante connexe géante ? De manière remarquable, celles-ci sont d'ordre logarithmique comme l'ont d'abord démontré van den Esker, van der Hofstad et Hooghiemstra [vdEvdHH08].

Théorème 19. *Supposons $c > 1$. Soient u_1 et u_2 deux sommets choisis indépendamment et uniformément dans la composante géante. Alors la convergence suivante a lieu en probabilité*

$$\frac{\mathrm{dist}(u_1, u_2)}{\log N} \xrightarrow[N\to+\infty]{} \frac{1}{\log c}.$$

Ce résultat est parfois appelé *small world phenomenon* que l'on peut traduire par l'expression courante "le monde est petit !". De ce point de vu, ceci fait du graphe d'Erdős-Rényi une bonne approximation des réseaux sociaux, où ce phénomène a été empiriquement vérifié. Par exemple, la distance moyenne entre deux utilisateurs de Facebook [BBR+12] est de 3,74 inter-amis.

Il est possible de deviner le Théorème 19 en utilisant la géométrie locale du graphe de la manière suivante. En explorant en largeur le voisinage de u_1, on s'attend à découvrir u_2 après avoir exploré un nombre de sommets d'ordre N. Puisque le voisinage de u_1 ressemble à un arbre de Galton-Watson T de loi de reproduction $\mathcal{P}(c)$, on peut faire l'approximation $|\mathrm{T}_k| \approx c^k$. Ainsi, on s'attend à ce que la distance de graphe d entre u_1 et u_2 vérifie $c^d \approx N$, autrement dit $d \approx \log N / \log c$.

Bien que les graphes d'Erdős-Rényi critiques ne constituent pas l'objet d'étude des contributions de cette thèse, nous terminons cette partie en proposant au lecteur une légère digression concernant leur comportement.

Comportement au point critique. Dans le cas critique où $c = 1$, Erdős et Rényi ont démontré que la taille de la plus grande composante connexe est d'ordre $N^{2/3}$. Au vu de la deuxième partie du Théorème 18, il est naturel d'étudier la possible unicité de la plus grande composante connexe. La réponse à ce problème a d'abord été apportée par Bollobás [Bol84], dont les résultats ont ensuite été précisés par Łuczak [Łuc90]. Au cours de la construction dynamique du graphe par l'ajout successif d'arêtes, notons $\mathscr{C}_{1,N}(t), \mathscr{C}_{2,N}(t), \ldots$ la suite des composantes connexes à l'étape t, rangées par ordre de tailles décroissantes. Notons par ailleurs $v_{1,N}(t)$ un sommet appartenant à la plus grande composante $\mathscr{C}_{1,N}(t)$. En écrivant $p = (1 + \lambda N^{-1/3})/N$, Łuczak

montre que la plus grande composante connexe est identifiable si et seulement si $|\lambda| \to +\infty$, en ce sens que

$$\exists N \geq 1,\ \exists t_0 = t_0(N),\ \forall t \geq t_0,\ v_{1,N}(t) \in \mathscr{C}_{1_N}(t) \quad \Leftrightarrow \quad |\lambda| \to +\infty.$$

Pour cette raison, le régime $p = (1 + \lambda N^{-1/3})/N$, $\lambda \in \mathbf{R}$ est appelé *fenêtre critique*.

Dans son célèbre article [Ald97], Aldous a étudié en détail le comportement de la marche de Łukasiewicz dans la fenêtre critique. Les idées développées dans ce papier irriguent encore tous les travaux portant sur le comportement critique de divers modèles de graphes aléatoires. Dans le cas des graphes d'Erdős-Rényi, l'argument principal consiste à remarquer que l'exploration du voisinage d'un sommet fixé du graphe peut être couplé avec celle d'un arbre de Galton-Watson dont la loi de reproduction est une loi de Poisson de paramètre $pN \approx 1$. Ceci suggère que la marche de Łukasiewicz devrait se comporter comme une marche aléatoire centrée et admettre une limite d'échelle. Plus précisément, les pas $(X_n - 1)_{n \geq 0}$ de la marche de Łukasiewicz sont correctement approximés de la manière suivante,

$$X_i - 1 \approx \mathrm{Bin}(N - i, (1 + \lambda N^{-1/3})/N) - 1 \approx \mathcal{P}\left(1 + \lambda N^{-1/3} - \frac{i}{N}\right) - 1.$$

On en déduit qu'à l'étape n, $S_n \approx \mathcal{P}(n + n\lambda N^{-1/3} - \frac{n^2}{2N}) - 1$. En choisissant $n = tN^{2/3}$ et après une mise à l'échelle $N^{1/3}$, on devine finalement les approximations suivantes

$$\frac{1}{N^{1/3}} S_{tN^{2/3}} \approx \frac{1}{N^{1/3}} \mathcal{P}\left(tN^{2/3} + t\lambda N^{1/3} - \frac{t^2}{2} N^{1/3}\right) - tN^{2/3} \approx W(t) + t\lambda - \frac{t^2}{2} \qquad (2.1)$$

où $(W(t))_t$ est un mouvement brownien. Cette heuristique peut être rendue rigoureuse et est à la source de l'article précité d'Aldous. Par la suite, d'autres investigations ont été menées au sujet des graphes d'Erdős-Rényi critiques, notamment sur le comportement des composantes connexes en tant qu'espaces métriques. Nous donnons ici le résultat le plus abouti à notre connaissance. On notera

- $\mathbf{M}_N$ la suite des composantes connexes, ordonnées par ordre de taille décroissante, et munies de la métrique de graphe ;
- $\mathbf{Z}_N = |\mathbf{M}_N|$ la suite du nombre de sommets de ces composantes ;
- $W^\lambda(t) := W(t) + t\lambda - \frac{t^2}{2}$, et $\mathbf{Z}$ la suite des longueurs des excursions du processus $W^\lambda(t) - \min_{0 \leq s \leq t} W^\lambda(s)$, rangées dans l'odre décroissant.

Théorème 20. *Il existe une suite d'espaces métriques $\mathbf{M} = (M_1, M_2, \ldots)$ telle que :*

$$\left(\frac{1}{N^{2/3}} \mathbf{Z}_n, \frac{1}{N^{1/3}} \mathbf{M}_n\right) \underset{n \to +\infty}{\longrightarrow} (\mathbf{Z}, \mathbf{M})$$

au sens de la convergence en loi, par rapport à la topologie de la convergence l^2 pour la première coordonnée et à la topologie de Gromov-Hausdorff pour la deuxième coordonnée.

La convergence de la première coordonnée a été démontrée par Aldous [Ald97]. Celle de la deuxième coordonnée a été obtenue ultérieurement par Addario-Berry, Broutin et Goldschmidt [ABBG12]. De manière informelle, on retiendra que dans le modèle d'Erdős-Rényi critique :

- les tailles des grandes composantes connexes sont d'ordre $N^{2/3}$,
- les distances typiques dans ces composantes sont d'ordre $N^{1/3}$.

2.2 Le modèle de configuration

Définition du modèle. Le modèle de configuration a été introduit par Bender and Canfield [BC78], Bollobás in [Bol80] et Wormald [Wor80] afin d'étudier certains problèmes de dénombrement sur les graphes aléatoires d-réguliers : nous reviendrons sur cette motivation historique plus tard. Rappelons ici qu'un multigraphe est un graphe où deux sommets peuvent être reliés par plusieurs arêtes, et où un sommet v peut posséder des boucles, c'est-à-dire des arêtes de la forme $\{v, v\}$. Un multigraphe est dit *simple* lorsqu'il ne possède ni arête multiple, ni boucle.

Soit $N \geq 1$ et $\mathbf{d} = (d_1, \ldots, d_N)$ une suite d'entiers positifs telle que la somme $d_1 + \cdots d_N$ est paire. On interprète la quantité d_i comme le nombre de demi-arêtes attachées au sommet numéro i. Le modèle de configuration $\mathscr{C}(\mathbf{d})$ associé à la suite $d_1, \ldots, d_N$ est défini comme le multigraphe aléatoire obtenu après l'appariement uniforme des demi-arêtes.

Figure 2.2 – Simulations de deux réalisations du modèle de configuration sur $N = 20$ sommets et dont les degrés sont choisis uniformément parmi $\{1, 2, 3\}$.

Soit G un multigraphe sur N sommets dont la suite des degrés est donnée par $d_1, \ldots, d_N$. Pour tout $1 \leq i \leq N$, on peut décomposer le degré de i dans G de la manière suivante : $d_i = d_{ii} + \sum_{j \neq i} d_{ij}$, où d_{ii} est le nombre de boucles issues du sommet i et d_{ij} le nombre d'arêtes entre le sommet i et le sommet j. Alors la probabilité que $\mathscr{C}(\mathbf{d})$ soit égal à G est donnée par

$$\mathbf{P}(\mathscr{C}(\mathbf{d}) = \mathrm{G}) = \frac{1}{(\sum_{1 \leq i \leq N} d_i)!!} \frac{\prod_{1 \leq i \leq N} d_i!}{\prod_{1 \leq i \leq N} 2^{\frac{d_{ii}}{2}} \left(\frac{d_{ii}}{2}\right)! \prod_{1 \leq i < j \leq N} d_{ij}!}, \tag{2.2}$$

où $(2n)!! := (2n-1) \times (2n-3) \cdots 3 \times 1$. En particulier, cette formule entraîne que le modèle de configuration conditionné à être un graphe simple suit la loi uniforme parmi les graphes dont la suite des degrés est $\mathbf{d}$.

Nous allons étudier des suites de modèles de configuration dont la taille N tend vers l'infini. À cet effet, on se fixe une suite de degrés pour tout $N \geq 1$,

$$\mathbf{d}^{(N)} := (d_1^{(N)}, \ldots, d_N^{(N)}) \in \mathbf{N}^N.$$

Nous supposerons qu'il existe une mesure de probabilité $\boldsymbol{\pi}$ sur les entiers positifs telle que $\sum_{i \geq 1} i^2 \boldsymbol{\pi}_i < +\infty$ et telle que la mesure empirique des degrés converge vers $\boldsymbol{\pi}$ au sens de la convergence presque sûre :

$$\forall k \geq 0, \quad \frac{1}{N} \sum_{i=0}^{N} \mathbf{1}_{d_i^{(N)} = k} \underset{N \to +\infty}{\longrightarrow} \boldsymbol{\pi}_k, \tag{A1}$$

Nous supposerons de plus que les seconds moments des mesures empiriques de cette suite convergent :

$$\lim_{N \to +\infty} \frac{{d_1^{(N)}}^2 + \cdots + {d_N^{(N)}}^2}{N} = \sum_{k \geq 0} k^2 \boldsymbol{\pi}_k. \tag{A2}$$

Finalement, nous supposerons qu'il existe une constante $\gamma > 2$ telle que, pour tout $N \geq 1$:

$$\max\left\{d_1^{(N)}, \dots, d_N^{(N)}\right\} \leq N^{1/\gamma}. \tag{A3}$$

Dans ce qui suit, on notera

$$\nu := \frac{\sum_{i\geq 1} i(i-1)\boldsymbol{\pi}_i}{\sum_{i\geq 1} i\boldsymbol{\pi}_i}.$$

Quelle est la probabilité qu'un modèle de configuration soit en fait un graphe simple ? En notant respectivement B_N et M_N le nombre de boucles et le nombre d'arêtes multiples, il est possible de montrer que (B_N, M_N) converge en loi vers un couple de variables aléatoires indépendantes, respectivement de loi de Poisson de paramètre $\nu/2$ et $\nu^2/4$. En particulier,

$$\mathbf{P}(\mathscr{C}(\mathbf{d}^{(N)}) \text{ est simple}) = \exp\left(-\frac{\nu}{2} - \frac{\nu^2}{4}\right)(1 + o(1)). \tag{2.3}$$

Par ailleurs, en utilisant (2.2), cette probabilité est aussi égale à

$$\mathbf{P}(\mathscr{C}(\mathbf{d}^{(N)}) \text{ est simple}) = \frac{\prod_{1\leq i\leq N} d_i^{(N)}!}{\left(\sum_{1\leq i\leq N} d_i^{(N)}\right)!!}\mathrm{N}(\mathbf{d}^{(N)}), \tag{2.4}$$

où $\mathrm{N}(\mathbf{d}^{(N)})$ désigne le nombre de graphes simples dont la suite des degrés est $\mathbf{d}^{(N)}$. En combinant les Équations (2.3) et (2.4), on obtient finalement un équivalent asymptotique du nombre de graphes simples ayant pour suite de degré $\mathbf{d}^{(N)}$:

$$\mathrm{N}(\mathbf{d}^{(N)}) = \frac{\left(\sum_{1\leq i\leq N} d_i^{(N)}\right)!!}{\prod_{1\leq i\leq N} d_i^{(N)}!}\exp\left(-\frac{\nu}{2} - \frac{\nu^2}{4}\right)(1 + o(1)). \tag{2.5}$$

Notons que cette formule a d'abord été obtenue par Bender et Canfield [BC78] avec une méthode différente. En particulier, on obtient l'équivalent suivant pour le cardinal $\mathrm{N}(d, N)$ des graphes d-réguliers ayant N, lorsque N tend vers l'infini,

$$\mathrm{N}(d, N) \sim \sqrt{2}\left(\frac{(dN)^{d/2}}{e^{d/2}d!}\right)^N \exp\left(\frac{-(d^2-1)}{4}\right).$$

L'apport de la méthode probabiliste que l'on vient de décrire ne s'arrête pas là ! Notons $\mathscr{G}(\mathbf{d}^{(N)})$ un graphe aléatoire simple, uniforme parmi les graphes simples dont la suite des degrés est $\mathbf{d}^{(N)}$. Par ailleurs, considérons $\mathscr{P}$ une propriété concernant les graphes, un exemple possible étant $\mathscr{P}$ = "être connexe". Comme l'Équation (2.4) implique que le modèle de configuration est simple avec une probabilité positive, on dispose de l'implication suivante :

$$\mathbf{P}(\mathscr{C}(\mathbf{d}^{(N)}) \text{ vérifie } \mathscr{P}) \to 1 \qquad \Rightarrow \qquad \mathbf{P}(\mathscr{G}(\mathbf{d}^{(N)}) \text{ vérifie } \mathscr{P}) \to 1.$$

Cette implication est précieuse, notamment parce qu'il est plus aisé de raisonner sur le modèle de configuration. Cela provient du fait qu'il est possible de construire ce modèle de manière dynamique, en effectuant séquentiellement les appariements des demi-arêtes. De manière analogue au modèle d'Erdős-Rényi, deux types d'algorithmes d'exploration/construction se distinguent : les parcours en largeur, qui révèlent le voisinage local d'un sommet donné, et les parcours en profondeur que nous décrirons en détails ultérieurement.

Géométrie locale. Comme dans le cas du modèle d'Erdős et Rényi, nous allons étudier la géométrie d'un modèle de configuration au voisinage d'un sommet v choisi uniformément. Pour cela, on utilise un algorithme d'exploration/construction en largeur qui, étant donné la suite des degré $\mathbf{d}^{(N)}$, construit successivement les boules centrées en v et de rayons k, $k \geq 1$.

La première étape de cet algorithme consiste à effectuer successivement les appariements uniformes de chacune des demi-arêtes issues de v. Par l'Hypothèse (**A1**) et comme la probabilité de créer une boucle $\{v, v\}$ est d'ordre $1/N$, le nombre de voisins de v ainsi révélés est asymptotiquement distribué selon la loi $\boldsymbol{\pi}$. Notons ces sommets $u_1, \ldots, u_{\deg v}$.

La deuxième étape de l'algorithme consiste à révéler successivement les voisins de $u_1, \ldots, u_{\deg v}$, ce qui revient à effectuer successivement les appariements uniformes des demi-arêtes de u_1, puis de u_2, etc. Détaillons l'analyse de la construction des voisins de u_1. Quelle est la probabilité que le degré sortant de u_1 soit égal à k ? Par construction, l'arête $\{v, u_1\}$ provient de l'appariement uniforme de la première demi-arête de v. Autrement dit, si le degré sortant de u_1 est égal à k, la probabilité d'avoir révélé $\{v, u_1\}$ était proportionnelle à $k+1$. Plus précisément :

$$\begin{aligned}\mathbf{P}(\text{degré sortant de } w \ = k) &= \frac{(k+1)|\{i \neq v,\, d_i^{(N)} = k+1\}|}{\sum_{i \neq v} i d_i^{(N)}} \\ &\underset{N \to +\infty}{\longrightarrow} \frac{(k+1)\boldsymbol{\pi}_{k+1}}{\sum_{i \geq 0} i \boldsymbol{\pi}_i}.\end{aligned}$$

Nous noterons $\widehat{\boldsymbol{\pi}}$ la mesure de probabilité définie par

$$\forall k \geq 0, \quad \widehat{\boldsymbol{\pi}}_k = \frac{(k+1)\boldsymbol{\pi}_{k+1}}{\sum_{i \geq 0} i \boldsymbol{\pi}_i}.$$

Lors de l'appariement des demi-arêtes de u_1, la probabilité de créer une boucle $\{u_1, u_1\}$ est d'ordre $1/N$. Ainsi, le nombre de sommets voisins de u_1 révélés lors de la deuxième étape de l'algorithme est asymptotiquement distribué selon la loi $\widehat{\boldsymbol{\pi}}$.

En poursuivant les considérations précédentes, il est possible d'établir le résultat suivant.

Théorème 21. *Notons* $\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$[2] *l'arbre aléatoire tel que :*

- *le nombre d'enfants de la racine* $c_\emptyset$ *est de loi* $\boldsymbol{\pi}$ *;*
- *conditionnellement à* $c_\emptyset = i$*, les arbres induits par les* i *enfants de la racine sont des arbres de Galton-Watson indépendants et de loi de reproduction* $\widehat{\boldsymbol{\pi}}$*.*

Alors

$$\forall k \geq 0, \quad (\mathscr{C}(\mathbf{d}^{(N)}), v)_k \underset{N \to +\infty}{\longrightarrow} \mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})_k,$$

au sens de la convergence en loi.

Comme dans le cas des graphes d'Erdős-Rényi, ce résultat peut se démontrer rigoureusement en couplant l'exploration en largeur du voisinage de v avec l'exploration en largeur du voisinage de la racine d'un arbre de Galton-Watson $\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$.

[2]par abus mais sans risque de confusion, on appellera encore arbre de Galton-Watson un tel arbre. Pour des raisons que nous ne détaillerons pas ici, l'appellation consacrée est en fait "arbre de Galton-Watson *unimodulaire*".

Un mot sur la convergence locale faible. Les Théorèmes 17 et 21 sont des avatars de *convergences locales faibles* de graphes aléatoires. De manière informelle, étant donné un graphe aléatoire enraciné (G_∞, ρ), on dit qu'une suite de graphes $(G_N)_{N\geq 1}$ converge vers (G_∞, ρ) au sens de la convergence locale faible lorsque,

$$\forall k \geq 0, \quad (G_N, v_N)_k \underset{N\to+\infty}{\longrightarrow} (G_\infty, \rho)_k, \tag{2.6}$$

au sens de la convergence en loi, et où v_N désigne un sommet choisi uniformément parmi les sommets de G_N. On parle aussi de convergence au sens de Benjamini et Schramm. Nous indiquons les références [BS01] et [AL07] au lecteur curieux. De manière inattendue, de nombreuses propriétés des graphes sont continues vis à vis de la convergence locale faible : on pourra consulter la thèse de Justin Salez [Sal11] pour quelques exemples. Dans la dernière partie de cette introduction, nous présenterons en particulier le cas des spectres de graphes. Même lorsqu'ils ne sont pas rigoureux, les raisonnements réalisés sur l'objet local limite fournissent bien souvent de puissantes heuristiques.

Existence d'un composante géante. De manière analogue au cas du modèle d'Erdős-Rényi, on commence par simplifier le problème de l'existence d'une composante connexe de taille macroscopique en le remplaçant par l'étude de la survie de l'arbre de Galton-Watson $\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$. Notons ξ la probabilité de survie d'un tel arbre, et ρ la probabilité de survie d'un arbre de Galton-Watson homogène de loi de reproduction $\widehat{\boldsymbol{\pi}}$. Par ailleurs, notons f_π et $f_{\widehat{\pi}}$ les séries génératrices des lois $\boldsymbol{\pi}$ et $\widehat{\boldsymbol{\pi}}$. La probabilité de survie d'un arbre de Galton-Watson homogène $\mathrm{GW}(\widehat{\boldsymbol{\pi}})$ est le plus grand réel dans l'intervalle $[0,1]$, point fixe de l'équation

$$1-\rho = f_{\widehat{\pi}}(1-\rho). \tag{2.7}$$

Par ailleurs, l'arbre de Galton-Watson $\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$ s'éteint si et seulement si tous les arbres induits par les enfants de la racine s'éteignent, autrement dit :

$$1-\xi = f_{\boldsymbol{\pi}}(1-\rho). \tag{2.8}$$

Ainsi, la probabilité de survie ξ est strictement positive si et seulement si ρ est strictement positif, ce qui est le cas si et seulement si la moyenne de la loi $\widehat{\boldsymbol{\pi}}$ est strictement plus grande que 1. En notant $\mathbf{E}_\pi$ l'espérance sous la loi $\boldsymbol{\pi}$, on vient de démontrer que

$$\xi = \mathbf{P}(\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}}) \text{ survit}) > 0 \qquad \Leftrightarrow \qquad \nu = \frac{\mathbf{E}_\pi[X(X-1)]}{\mathbf{E}_\pi[X]} > 1.$$

Comme dans le cas du modèle d'Erdős-Rényi, le problème local simplifié coïncide en fait avec le problème initial.

Théorème 22. *Soit $\mathscr{C}_{\max}$ la plus grande composante connexe de $\mathscr{C}(\mathbf{d}^{(N)})$. On dispose de la dichotomie suivante:*

- *si $\nu < 1$, alors $|\mathscr{C}_{\max}| = \mathscr{O}_{\mathbf{P}}(\log N)$ i.e. $\lim_{\kappa\to+\infty} \limsup_{N\to+\infty} \mathbf{P}(|\mathscr{C}_{\max}| > \kappa \log N) = 0$,*
- *si $\nu > 1$, alors $\frac{|\mathscr{C}_{\max}|}{N} \to \xi$ en probabilité lorsque $N \to +\infty$. De plus, les tailles des autres composantes connexes sont d'ordre $\mathscr{O}_{\mathbf{P}}(\log N)$.*

Une preuve rigoureuse peut être obtenue en étudiant la marche de Łucasiewicz comme pour le Théorème 18. Une autre approche très élégante a été proposée par Molloy et Reed [MR95, MR98]. Nous présentons ici une adaptation de leurs arguments due à Janson et Łuczak [JL09]. Il

s'agit d'abord d'introduire un algorithme de construction du modèle de configuration, que l'on pourra analyser précisément par la suite. Munissons les demi-arêtes d'horloges exponentielles i.i.d. de moyenne 1.

Initialement, tous les sommets et toutes les demi-arêtes sont déclarés non-explorés. L'algorithme construit une à une les composantes connexes du graphe de la manière suivante. On choisit d'abord un sommet v uniformément parmi tous les sommets non-explorés, et on déclare v comme étant exploré.

À chaque étape, on choisit uniformément une demi-arête non-explorée, disons e_1, de la composante connexe en cours d'exploration et on l'apparie à la première demi-arête non-explorée dont l'horloge sonne, disons e_2. Ceci-étant fait, on déclare les demi-arêtes e_1 et e_2 comme étant explorées, et l'on déclare le sommet auquel e_2 est rattaché comme étant exploré. Notons que ce sommet peut tout à fait déjà appartenir à la composante connexe en cours d'exploration, auquel cas l'on vient de révéler une arête au sein de cette composante. On réitère cette procédure jusqu'à avoir construit la composante connexe de v. Lorsque celle-ci est effectivement construite, on recommence la procédure à partir d'un sommet v' choisi uniformément parmi les sommets non-explorés. Soulignons que le graphe ainsi construit possède bien la loi du modèle de configuration puisque les appariements effectués au cours de l'algorithme sont tous uniformes. Pour tout temps $t \geq 0$, on définit:

- $A(t)$ le nombre de demi-arêtes non-explorées attachées à un sommet de la composante connexe en cours d'exploration,
- $V_k(t)$ le nombre de sommets de degré k non-explorés,
- $S(t) = \sum_{k\geq 0} kV_k(t)$ le nombre de demi-arêtes attachées à un sommet non-exploré,
- $L(t) := A(t) + S(t)$ le nombre total de demi-arêtes non-explorées.

De manière informelle, $L(t)$ décroît de 2 chaque fois qu'une horloge sonne, et à chacune de ces sonneries, $V_k(t)$ décroît de 1 proportionnellement à $kV_k(t)$. On s'attend donc à ce que ces fonctions vérifient les équations différentielles suivantes :

$$\begin{cases} L'(t) = -2L(t), \\ V_k'(t) = -kV_k(t). \end{cases} \tag{2.9}$$

Les solutions sont données par $L(t) = L(0)\,\mathrm{e}^{-2t}$ et $V_k(t) = V_k(0)\,\mathrm{e}^{-kt}$, d'où l'on déduit que la proportion de demi-arêtes attachées à des sommets de la composante connexe en cours d'exploration vaut

$$\frac{A(t)}{N} = \frac{L(t)}{N} - \frac{S(t)}{N} = \sum_{i\geq 0} i\boldsymbol{\pi}_i\,\mathrm{e}^{-2t} - \sum_{i\geq 0} i\boldsymbol{\pi}_i\,\mathrm{e}^{-it} =: H(\mathrm{e}^{-t}).$$

Une étude de la fonction $t \mapsto H(\mathrm{e}^{-t})$ montre qu'elle admet une excursion positive sur l'intervalle $[0, -\log(1-\rho)]$, où ρ est défini par l'Équation (2.7). Ceci a effectivement lieu si et seulement si $\nu > 0$. Dans ce cas, la proportion de sommets explorés dans la composante de taille macroscopique associée à cette excursion est donnée par :

$$1 - \sum_{i\geq 0} \frac{V_i(0)}{N}\,\mathrm{e}^{-it} = 1 - \sum_{i\geq 0} \boldsymbol{\pi}_i(1-\rho)^i = 1 - f_{\boldsymbol{\pi}}(1-\rho),$$

qui n'est autre que ξ au vu de l'Équation (2.8). Bien sûr, les détails techniques permettant de transformer cette heuristique en preuve rigoureuse requerraient quelques pages de plus, en particulier pour démontrer que les équations (2.9) sont en fait vérifiées par les limites fluides des fonctions $t \mapsto L(t)/N$ et $t \mapsto V_k(t)/N$.

Le monde est petit. Lorsque $\nu > 1$, nous avons vu que le modèle de configuration possède une unique composante connexe de taille macroscopique. Que peut-on dire des distances typiques au sein de cette composante ? Comme dans le cas du modèle d'Erdős-Rényi, celles-ci sont d'ordre logarithmique [vdHHVM05].

Théorème 23. *Soient u_1 et u_2 deux sommets choisis indépendamment et uniformément dans la composante géante. Alors la convergence suivante a lieu en probabilité*

$$\frac{\mathrm{dist}(u_1, u_2)}{\log N} \underset{N\to+\infty}{\longrightarrow} \frac{1}{\log \nu}.$$

De manière analogue au Théorème 19, une heuristique utilisant l'approximation locale au voisinage de u_1 est possible.

Avant de nous tourner vers une analyse plus approfondie de la géométrie de la composante connexe géante, nous réalisons ici une légère digression concernant le comportement des modèles de configuration critiques.

Comportement au point critique. Nous considérons ici un modèle de configuration critique, c'est-à-dire un modèle de configuration tel que $\nu = 1$. Dans ce cas la marche de Łukasiewicz associée à l'exploration du voisinage d'un sommet du graphe peut être couplée avec une marche aléatoire dont les pas sont des variables aléatoires i.i.d. de loi $\widehat{\boldsymbol{\pi}} - 1$, dont l'espérance vaut $\nu - 1 = 0$. La limite d'échelle d'une telle marche dépend de la queue de distribution de $\boldsymbol{\pi}$: lorsque $\boldsymbol{\pi}$ possède un troisième moment, on s'attend à obtenir un processus relié au mouvement brownien. En revanche, si $\boldsymbol{\pi}$ ne possède pas de troisième moment, on s'attend à obtenir des processus de Lévy. Cette prédiction a été confirmée par Joseph [Jos14] qui a démontré le résultat suivant.

Théorème 24. *Notons $\mu := \sum_{i\geq 1} i\boldsymbol{\pi}_i$, et $\mathbf{Z}_N$ la suite des tailles des composantes connexes de $\mathscr{C}(\mathbf{d}^{(N)})$, rangées dans l'ordre décroissant.*

1. *Supposons que $\boldsymbol{\pi}_2 < 1$ et $\beta := \sum_{i\geq 1} i(i-1)(i-2)\boldsymbol{\pi}_i < +\infty$. Dans ce cas, fixons W un mouvement brownien standard et définissons le processus*

$$W^{\boldsymbol{\pi}}(t) := \sqrt{\frac{\beta}{\mu}} W(t) - \frac{\beta}{2\mu^2} t^2.$$

Introduisons finalement $\mathbf{Z}$ la suite des longueurs des excursions du processus réfléchi $W^{\boldsymbol{\pi}}(t) - \min_{0\leq s\leq t} W^{\boldsymbol{\pi}}(s)$, rangées dans l'ordre décroissant. Alors

$$\frac{1}{N^{2/3}} \mathbf{Z}_N \underset{N\to+\infty}{\longrightarrow} \mathbf{Z}.$$

2. *Supposons qu'il existe $c > 0$ et $\alpha \in (1,2)$ tel que $\boldsymbol{\pi}_k \sim ck^{-(\alpha+2)}$. Dans ce cas, on introduit X, l'unique processus d'incréments indépendants et de transformée de Laplace*

$$\mathbf{E}[\exp(-\lambda X(t))] = \exp\left(\int_0^t \mathrm{d}s \int_0^{+\infty} \mathrm{d}x \left(\mathrm{e}^{-\lambda x} - 1 + \lambda x \right) \frac{c}{\mu} \frac{1}{x^{\alpha+1}} \mathrm{e}^{-xs/\mu} \right),$$

et

$$X^{\boldsymbol{\pi}}(t) = X(t) - \frac{c\Gamma(2-\alpha)}{\alpha(\alpha-1)\mu^{\alpha}} t^{\alpha}.$$

Alors, en notant $\mathbf{Z}$ la suite des excursions du processus réfléchi $X^{\boldsymbol{\pi}}(t) - \min_{0\leq s\leq t} X^{\boldsymbol{\pi}}(s)$, rangées dans l'ordre décroissant, on a :

$$\frac{1}{N^{\alpha/(\alpha+1)}} \mathbf{Z}_N \underset{N\to+\infty}{\longrightarrow} \mathbf{Z}.$$

Notons que la première partie de ce Théorème a été obtenue indépendamment par Riordan [Rio12], sous des conditions légèrement différentes. Concernant la convergence des plus grandes composantes connexes au sens de Gromov-Hausdorff, nous indiquons au lecteur le récent travail de Conchon-Kerjan et Goldschmidt [GCK20] – dont l'énoncé ci-dessus est tiré, où les auteurs obtiennent l'analogue de la convergence de la deuxième composante dans le Théorème 20. De manière informelle, on retiendra que

- si π possède un troisième moment, les tailles des composantes connexes sont d'ordre $N^{2/3}$ et les distances typiques au sein de ces composantes sont d'ordre $N^{1/3}$,
- s'il existe $c > 0$ et $\alpha \in (1,2)$ tels que $\pi_k \sim ck^{-(\alpha+2)}$, les tailles des composantes connexes sont d'ordre $N^{\alpha/(\alpha+1)}$ et les distances typiques au sein de ces composantes sont d'ordre $N^{(\alpha-1)/(\alpha+1)}$.

2.3 Longs chemins simples et parcours en profondeur

La géométrie des composantes connexes macroscopiques des modèles sur-critiques a fait l'objet de nombreux travaux. Nous avons par exemple vu que les distances typiques sont d'ordres logarithmiques, ce qui suggère que la composante géante se replie sur elle-même. Ce phénomène peut-il empêcher l'apparition de chemins simples de taille linéaire en le nombre de sommets dans le graphe ? Une manière naturelle d'obtenir une borne inférieure sur la taille du plus long chemin simple d'un graphe est d'estimer la longueur de la plus longue branche obtenue par l'algorithme de *parcours en profondeur*. Après avoir introduit cet algorithme, nous énoncerons les résultats principaux connus concernant son étude asymptotique dans le cas du modèle d'Erdős et Rényi. Nous présenterons ensuite les résultats du Chapitre 6, issu de l'article [EFMN19] réalisé en collaboration avec Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard, qui concerne l'étude asymptotique de l'algorithme de parcours en profondeur dans un modèle de configuration sur-critique. Les méthodes développées dans cet article nous ont également permis d'étudier une variante de l'algorithme du parcours en profondeur, permettant de construire de longs chemins induits dans le graphe. Nous nous contenterons ici d'énoncer les résultats correspondants et renvoyons le lecteur au Chapitre 7, issu de l'article [EFMN20] pour consulter les détails.

L'algorithme de parcours en profondeur. Soit G un graphe ou multigraphe possédant N sommets numérotés. De manière informelle, l'algorithme de parcours en profondeur, qu'on désignera désormais par DFS – pour *Depth First Search*, explore la composante connexe du sommet initialement exploré en effectuant une marche au plus proche voisin dans le graphe. À chaque étape, celle-ci visite un voisin non-exploré du sommet courant si cela est possible, et rebrousse chemin dans le cas contraire. Afin d'éviter toute confusion, nous en donnons maintenant une définition formelle. À chaque étape n de l'algorithme, on considèrera :

- A_n l'ensemble des sommets actifs, munis d'un ordre induit par l'exploration ;
- a_n le dernier élément de la liste A_n, qui correspond au sommet courant ;
- S_n l'ensemble des sommets non-explorés ;
- $R_n = \{1, \dots, N\} \setminus (A_n \cup S_n)$ l'ensemble des sommets retraités.

Disons que l'algorithme commence par explorer le sommet 1. Dans ce cas, ces quantités sont initialement données par $A_0 = (1)$, $S_0 = \{2, \dots, N\}$ et $R_0 = \emptyset$. L'algorithme s'arrête lorsque

$A_n = \emptyset$, ce qui arrive lorsque $n = 2|\mathscr{C}(1)| - 1$, où $\mathscr{C}(1)$ désigne la composante connexe du sommet 1. L'étape $n \to n+1$ de l'algorithme est obtenue de la manière suivante :

- Si a_n possède un voisin dans S_n, alors la marche associée au DFS visite le voisin de a_n ayant le plus petit numéro et

$$\begin{cases} a_{n+1} = \inf\{i \in S_n,\ \{a_n, i\} \in \mathrm{E(G)}\}, \\ A_{n+1} = A_n \cup a_{n+1} \text{ au sens de la concaténation (à droite)}, \\ S_{n+1} = S_n \setminus \{a_{n+1}\}, \\ R_{n+1} = R_n. \end{cases}$$

- Si a_n n'a pas de voisin dans S_n, alors la marche rebrousse chemin et

$$\begin{cases} A_{n+1} = A_n \setminus \{a_n\}, \\ S_{n+1} = S_n, \\ R_{n+1} = R_n \cup \{a_n\}. \end{cases}$$

Soulignons ici que, dans le deuxième cas, a_{n+1} correspond au dernier élément de A_{n+1}. Ainsi définie, la suite des sommets $(a_n)_n$ est une marche au plus proche voisin sur la composante connexe $\mathscr{C}(1)$, dont la trace en exhibe un arbre couvrant. Cet arbre est naturellement muni d'une structure d'arbre planaire enraciné puisque l'algorithme du DFS possède un sommet de départ et un historique d'exploration. Il sera plus commode d'étudier le processus de contour $(X_n)_{0 \le n \le 2|\mathscr{C}(1)|-1}$ associé à cet arbre, qui est une marche restant positive, de pas $+1$ lorsque l'exploration "monte" dans l'arbre, et de pas -1 lorsque celle-ci "descend". On consultera la Figure 2.3 pour une illustration de ces définitions.

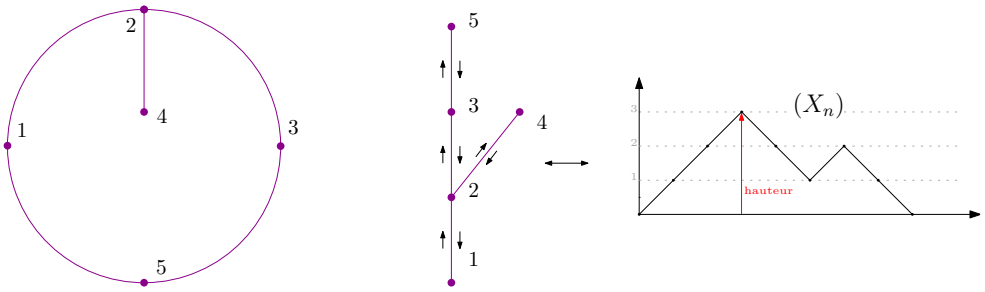


Figure 2.3 – Exemple de parcours en profondeur d'un graphe connexe (à gauche). À droite, l'arbre couvrant construit par le DFS et le processus de contour associé.

Par définition de l'algorithme, la hauteur du processus de contour est la longueur de la plus longue branche découverte par le DFS, ce qui fournit une borne inférieur sur la longueur du plus long chemin simple dans le graphe.

Cas du modèle d'Erdős-Rényi. Dans le cas des graphes d'Erdős-Rényi sur-critiques – où $c > 1$, l'existence de longs chemins simples de taille linéaire en le nombre de sommets a été conjecturée par Erdős [Erd75]. Fernandez de la Vega [FdlV79] et Ajtai, Komlós et Szemerédi [AKS81] ont démontré indépendamment cette conjecture. Leurs preuves consistent à vérifier que l'algorithme de parcours en profondeur possède un nombre linéaire de "montées" consécutives. Cette idée a été exploitée par Sudakov et Krivelevich [KS13] afin de donner une preuve relativement simple de la transition de phase pour l'existence d'une composante macroscopique. Une analyse

complète du DFS n'a été réalisée que très récemment par Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard [EFM20]. Rappelons que lorsque $c > 1$, la probabilité de survie d'un arbre de Galton-Watson homogène de loi de reproduction $\mathcal{P}(c)$ est l'unique réel $0 < \rho_c < 1$ tel que $1 - \rho_c = \exp(c(\mathrm{e}^{1-\rho_c} - 1))$. Notons par ailleurs Li_2 la fonction Dilogarithme.

Théorème 25. *Soit (X_n) le processus de contour du DFS, réalisé à partir d'un sommet choisi uniformément dans la composante géante. Alors la convergence suivant a lieu au sens de la topologie uniforme*

$$\forall t \in [0, 2\rho_c], \quad \lim_{N\to+\infty} \frac{X_{\lfloor tN \rfloor}}{N} = h(t),$$

où la fonction h est bien définie et continue sur l'intervalle $[0, 2\rho_c]$. Le graphe $(t, h(t))_{t\in[0,2\rho_c]}$ peut être divisé en une première partie de montée et une deuxième partie de descente. Ces deux parties sont respectivement paramétrées par

$$\begin{aligned}(t, h(t))_{0\le t\le f(0)} &= (f(\rho), g(\rho))_{0\le\rho\le\rho_c},\\ (t, h(t))_{f(0)\le t\le 2\rho_c} &= \left(f(\rho) + 2\rho\left(1 - \tfrac{f(\rho)+g(\rho)}{2}\right), g(\rho)\right)_{0\le\rho\le\rho_c},\end{aligned}$$

où les fonctions f et g sont définies par

$$\begin{aligned}f(\rho) &= \tfrac{1}{c}\left(Li_2(\rho_c) - Li_2(\rho) + \log\tfrac{1-\rho_c}{1-\rho} - 2\left(\tfrac{\log(1-\rho_c)}{\rho_c} - \tfrac{\log(1-\rho)}{\rho}\right)\right),\\ g(\rho) &= \tfrac{1}{c}\left(Li_2(\rho) - Li_2(\rho_c) + \log\tfrac{1-\rho}{1-\rho_c}\right).\end{aligned}$$

Le contraste avec le comportement critique est frappant : ici, la limite d'échelle est *déterministe*. En évaluant la fonction g en 0, on obtient en particulier une borne inférieure sur la taille $\mathcal{H}_N$ du plus long chemin simple dans le graphe $\mathrm{ER}(N, c/N)$:

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N}{N} \ge \rho_c - \frac{Li_2(\rho_c)}{c} - \varepsilon\right) \underset{N\to+\infty}{\longrightarrow} 1. \tag{2.10}$$

Le DFS étant un algorithme glouton, cette borne n'est pas atteinte en générale comme nous allons le voir dans le régime des grands c. À notre connaissance, il n'existe pas d'algorithme de complexité polynomiale en le nombre de sommets et fournissant une meilleure borne.

Pour obtenir de meilleurs estimées sur $\mathcal{H}_N$, il faut avoir recours à des arguments plus abstraits de théorie des graphes. L'idée principale est d'identifier un sous-graphe Hamiltonien, c'est-à-dire contenant un cycle simple parcourant tous les sommets, dans le graphe d'Erdős-Rényi. Cette méthode a d'abord été utilisée par Bollobás [Bol82] puis par Bollobás, Fenner et Frieze [BFF84]. La meilleure borne a été obtenue plus tard par Frieze [Fri86], qui montre que que $\mathcal{H}_N/N$ est asymptotiquement minoré par $1 - (1 + \varepsilon_c)(1 + c)\,\mathrm{e}^{-c}$, où $\varepsilon_c \to 0$ lorsque $c \to +\infty$. Comme le nombre de sommets de degré 1 du graphe est par ailleurs égal à $c\,\mathrm{e}^{-c}\,N$, cette borne est la meilleure possible lorsque $c \to +\infty$. Si $c > 1$ est une constante suffisamment grande, Anastos et Frieze [AF20] ont récemment démontré que la suite de variables aléatoires $(\mathcal{H}_N/N)_{N\ge1}$ converge presque sûrement vers une constante $f(c)$. De ce point de vue, la conjecture suivante semble raisonnable.

Conjecture. *Pour tout $c > 1$, il existe une constante $f(c)$ telle que la convergence suivante a lieu presque sûrement*

$$\frac{\mathcal{H}_N}{N} \underset{N\to+\infty}{\longrightarrow} f(c).$$

Comme le lecteur pourra facilement s'en convaincre, il n'existe pas d'arguments de sous-additivité pour ce problème, le rendant par la même occasion notoirement difficile. Une piste prometteuse consisterai à rendre rigoureuse la *méthode de la cavité* développée par Marinari et Semerjian [MS06].

Cas du modèle de configuration. Soit $(\mathbf{d}^{(N)})_{N\geq 1}$ une suite satisfaisant les hypothèses (**A1**), (**A2**), (**A3**) et telle que $\nu > 1$. Dans ce contexte, nous avons vu que la suite des modèles de configuration $\mathscr{C}(\mathbf{d}^{(N)})$ est sur-critique en ce sens qu'il existe une unique composante connexe de taille macroscopique. Le seul résultat concernant l'existence de longs chemins simples à l'intérieur de cette composante géante est dû à Frieze et Jackson [FJ87]. Sous l'hypothèse supplémentaire que les degrés sont uniformément bornés et supérieurs à 3, les auteurs établissent l'existence de longs cycles induits dans le graphe, c'est-à-dire de longs cycles tels que toute paire de sommets à distance plus grande que deux ne forme pas une arête du graphe. Leur résultat sera présenté dans le paragraphe suivant.

Nous nous attacherons ici à présenter un travail effectué en collaboration avec Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard, où nous avons analysé en détail l'algorithme de parcours en profondeur réalisé sur un modèle de configuration sur-critique, et établi en particulier une borne inférieure sur la longueur du plus long chemin simple. L'article correspondant [EFMN19] constitue le Chapitre 6 de cette thèse. Dans les lignes qui suivent, nous commençons par détailler les ingrédients principaux de notre analyse avant d'énoncer précisément les résultats que nous avons obtenus.

La première étape de notre démarche consiste à introduire un algorithme qui, étant donnée la suite des degrés $\mathbf{d}^{(N)}$, construit simultanément le modèle de configuration $\mathscr{C}(\mathbf{d}^{(N)})$ et un parcours en profondeur de ce graphe. Cette construction consiste à réaliser un appariement séquentiel convenable des demi-arêtes. À chaque étape, le lecteur se représentera un marcheur associé à l'algorithme se trouvant en un sommet du graphe en cours de construction. De manière informelle,

- ou bien le sommet courant ne possède pas de demi-arête dont on n'a pas encore effectué l'appariement et le marcheur rebrousse chemin,
- ou bien certaines demi-arêtes de ce sommet n'ont pas encore été appariées, auquel cas on les apparie de manière uniforme avec des demi-arêtes non-explorées, et le marcheur se déplace sur l'un des sommets ainsi révélé.

Le lecteur pourra consulter la partie 6.4 pour une définition précise. Puisque les appariements effectués au cours de cet algorithme sont uniformes, le graphe construit à la fin des temps est bien un modèle de configuration. De plus, l'historique des sommets visités par le marcheur fournit un arbre couvrant de chaque composante connexe de ce graphe, et la concaténation des processus de contour de ces arbres suit la loi du processus de contour associé au DFS. La définition de cet algorithme de construction/exploration permet également de s'assurer qu'au cours du temps, les graphes *induits* par les sommets non-explorés sont des modèles de configuration. Nous allons voir que l'étude de l'évolution de la loi du degré empirique des sommets de ces graphes induits est au cœur de notre analyse.

Une première remarque est que l'étude directe du processus de contour associé au parcours en profondeur n'est pas possible. En revanche, il existe une suite de temps non-markoviens $(T_k)_{k\geq 0}$, appelés temps de pseudo-renouvellement, auxquels le DFS se prête à une analyse détaillée. Le premier temps vaut $T_0 = 0$ et, si T_k est construit, on définit T_{k+1} comme le premier temps $i > T_k$ tel que le processus de contour reste au dessus du niveau $k+1$ pendant un temps d'ordre au moins $\sqrt{N}$ à partir de ce temps i. Intuitivement, la différence $T_{k+1} - T_k$ correspond au temps effectif écoulé pour que le processus de contour monte de 1. Aussi, dans l'hypothèse où ce processus possède une limite fluide, on s'attend à ce que

$$T_{k+1} - T_k \approx \frac{1}{\text{pente du profil limite}},$$

où $\approx$ désigne une moyenne mésoscopique réalisée sur des entiers successifs.

Donnons maintenant une idée du calcul de $T_{k+1} - T_k$. Notons v_k le sommet courant au temps T_k et supposons qu'une proportion positive mais suffisamment petite des sommets a été explorée. Dans ce cas, le graphe induit par les sommets non-explorés au temps T_k est encore un modèle de configuration sur-critique dont on notera $\mathscr{C}_{\max}(k)$ l'unique composante connexe géante. À partir du sommet v_k, le marcheur associé à l'algorithme du DFS explore les composantes connexes des voisins de v_k dans le graphe induit par les sommets non-explorés. En notant $(p_k)_{k\geq 0}$ la mesure empirique des degrés de ce graphe, l'exploration en profondeur de ces composantes connexes peut être couplée avec l'exploration en profondeur d'arbres de Galton-Watson indépendants et homogènes de loi de reproduction $(\widehat{p}_k)_{k\geq 0}$. En notant $\rho_k > 0$ la probabilité de survie de tels arbres, ceci entraîne que chaque voisin de v_k possède une probabilité approximativement donnée par ρ_k d'appartenir à $\mathscr{C}_{\max}(k)$. En conséquence, si $\mathfrak{e}_k$ désigne le nombre de fois qu'une petite composante connexe est visitée avant de visiter l'unique composante géante $\mathscr{C}_{\max}(k)$ au cours de l'exploration, alors $\mathfrak{e}_k + 1$ suit approximativement une loi géométrique de paramètre ρ_k, conditionnée à être inférieure au nombre de voisins de v_k. De plus, les composantes connexes correspondant aux $\mathfrak{e}_k$ premiers voisins visités devraient être correctement approximées par des arbres de Galton-Watson indépendants de loi de reproduction $(\widehat{p}_k)_{k\geq 0}$ et conditionnés à l'extinction. Toutes ces considérations peuvent être rendues rigoureuses et en menant les calculs à terme, il est possible d'établir la formule suivante :

$$\mathbf{E}[T_{k+1} - T_k \,|\, \mathcal{F}_k] = \frac{2-\rho_k}{\rho_k} + o(1). \tag{2.11}$$

où $(\mathcal{F}_k)$ désigne la filtration naturelle associée à l'exploration.

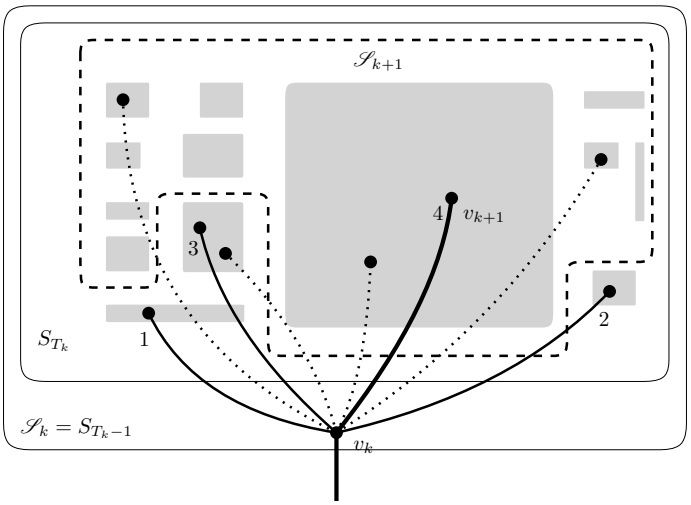


Figure 2.4 – Figure issue du Chapitre 6 illustrant la structure du graphe à un temps de pseudo-renouvellement de l'algorithme. Les demi-arêtes issues de v_k sont numérotées dans l'ordre de leur exploration. En particulier, $\mathfrak{e}_k = 3$. Les demi-arêtes qui seront explorées après l'exploration de la composante géante sont représentées en pointillés.

Ce raisonnement a d'abord été mis en exergue par Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard [EFM20] dans le contexte des graphes d'Erdős-Rényi. Tout l'enjeu réside donc dans la compréhension des quantités ρ_k, $k \geq 0$. Comme ρ_k est caractérisé par l'équation de point fixe $1 - \rho_k = \sum_{i\geq 0} \widehat{p}_i (1-\rho_k)^i$, il s'agit en fait d'identifier l'évolution de la mesure empirique des degrés dans les graphes induits par les sommets non-explorés. Dans le cas du modèle

d'Erdős-Rényi, comprendre cette évolution est immédiat puisque par définition du modèle, après avoir exploré une proportion α des sommets, le graphe restant à explorer est un graphe d'Erdős-Rényi ayant $(1-\alpha)N$ sommets et de probabilité de connexion c/N. Dans le cas du modèle de configuration, l'étude de l'évolution de la mesure empirique des degrés s'avère plus délicate et est au cœur de l'article [EFMN19]. Nous décrivons les étapes principales de notre démarche dans les lignes qui suivent.

Pour tout $i \geq 0$, notons $N_i(k)$ le nombre de sommets non-explorés au temps T_k et dont le degré vaut i dans le graphe qu'ils induisent. Une analyse structurelle du graphe entre deux temps de pseudo-renouvellement permet d'établir que pour tout $i \geq 0$,

$$\mathbf{E}[N_i(k+1) - N_i(k) \mid \mathcal{F}_k] = f_i\left(\frac{N_0(k)}{N}, \frac{N_1(k)}{N}, \dots\right) + o(N),$$

où $f_i : \mathbf{R}^{\mathbf{N}} \to \mathbf{R}$ est une fonction explicite, dont une expression est donnée dans l'Équation (6.8). Dans ce cas, une variante de la méthode de l'équation différentielle de Wormald [Wor95] entraîne que les quantités $N_i(k)/N$ possèdent des limites fluides $z_i(k/N)$, uniques solutions du système infini formé par les équations différentielles $z_i'(t) = f_i((z_j(t))_{j\geq 0})$[3]. La fin de la démonstration est jalonnée de trois petits miracles que l'on liste ici.

1. Il existe des fonctions $\widetilde{f}_i((z_j(t))_{j\geq 0})$ dépendant *explicitement* des $z_j(t)$ et telles que

$$f_i((z_j(t))_{j\geq 0}) = \frac{1}{\rho_{(z_j(t))_{j\geq 0}}} \widetilde{f}_i((z_j(t))_{j\geq 0}).$$

   La présence de la quantité implicite ρ n'est pas de bonne augure, mais il est possible de contourner son influence en effectuant un changement de temps qui nous ramène à l'étude du système infini formé par les équations différentielles

$$\zeta_i'(t) = \widetilde{f}_i((\zeta_j(t))_{j\geq 0}).$$

2. En introduisant la série génératrice $f(t,s) = \sum_{i\geq 0} \zeta_i(t) s^i$, ce système infini peut s'écrire de manière concise sous la forme :

$$\frac{\partial f}{\partial t}(t,s) = \frac{\frac{\partial f}{\partial s}(t,s)}{\frac{\partial f}{\partial s}(t,1)} \left((1-s) \frac{\frac{\partial^2 f}{\partial s^2}(t,1)}{\frac{\partial f}{\partial s}(t,1)} - 1 \right). \tag{2.12}$$

   En particulier, l'évaluation en $s = 1$ montre que $\sum_{i\geq 0} \zeta_i'(t) = -1$ et par suite que $\sum_{i\geq 0} \zeta_i(t) = 1 - t$. Ainsi, la nouvelle échelle de temps, qui était un artefact *a priori* purement technique, a un sens physique puisqu'elle correspond à la proportion de sommets explorés !

3. Finalement, l'Équation (2.12) admet une solution explicite, donnée par

$$f(t,s) = f_\pi\left(f_\pi^{-1}(1-t) - (1-s) \frac{f_\pi'(f_\pi^{-1}(1-t))}{f_\pi'(1)} \right).$$

Notons $S_n^{(N)}$ le graphe induit par les sommets non-explorés à l'étape n de l'algorithme du parcours en profondeur. Nous venons d'esquisser la démonstration du résultat suivant.

[3] Pour être plus précis, la convergence a lieu sous réserve que le système infini admette une unique solution... ce que l'on peut démontrer : le lecteur consultera la partie 6.6.2 pour les détails techniques.

Théorème H : Chapitre 6, Théorème 10

Soit α_c la plus petite solution positive de l'équation suivante

$$\frac{f''_\pi\left(f_\pi^{-1}(1-\alpha)\right)}{f'_\pi(1)} = 1.$$

Pour tout $\alpha \in [0, \alpha_c]$, on définit la mesure de probabilité $\boldsymbol{\pi}_\alpha$ sur $\mathbf{N}$, dont la série génératrice vaut

$$g(\alpha, s) = \frac{1}{1-\alpha} f_\pi\left(f_\pi^{-1}(1-\alpha) - (1-s)\frac{f'_\pi(f_\pi^{-1}(1-\alpha))}{f'_\pi(1)}\right). \tag{2.13}$$

Alors, pour tout $\alpha \in [0, \alpha_c]$, si l'on note $\tau^{(N)}(\alpha) = \inf\{k \geq 1 : |S_k^{(N)}| \leq (1-\alpha)N\}$, la mesure empirique des degrés dans les graphes $S^{(N)}_{\tau^{(N)}(\alpha)}$ converge vers $\boldsymbol{\pi}_\alpha$ en probabilité.

De plus, certains choix de loi initiale permettent d'effectuer des calculs explicites :

- lorsque $\boldsymbol{\pi}$ est une loi de Poisson de paramètre $c > 1$, ce qui correspond au modèle d'Erdős-Rényi, π_α est une loi de Poisson de paramètre $c(1-\alpha)$,
- lorsque $\boldsymbol{\pi} = \delta_d$ pour un entier $d \geq 3$, ce qui correspond au modèle d-régulier, π_α est une loi Binomiale de paramètres d et $(1-\alpha)^{(d-2)/d}$,
- lorsque $\boldsymbol{\pi}$ est une loi Géométrique partant de 0 et de moyenne $0 \leq p < 2/3$, π_α est une loi Géométrique de paramètre $p(\alpha) = \frac{p}{p+(1-p)(1-\alpha)^3}$.

À partir du Théorème H, il est possible d'obtenir une description de ce profil. Notons d'abord que la relation (2.11) entraîne que T_k/N admet une limite fluide $z(k/N)$, ce qui fournit effectivement l'existence d'un profil limite $(t, z(t))$ pour le processus de contour renormalisé associé au parcours en profondeur.

Introduisons ρ_π le plus grand réel dans l'intervalle $[0,1]$ solution de l'Équation $1-\rho = f_{\widehat{\pi}}(1-\rho)$, et définissons, pour tout $\rho \in (0, \rho_\pi]$, la quantité implicite $\alpha(\rho)$ caractérisée par l'équation

$$\forall \rho \in (0, \rho_\pi], \quad 1-\rho = \frac{\partial_s g(\alpha(\rho), 1-\rho)}{\partial_s g(\alpha(\rho), 1)}.$$

En notant $(x(\rho), y(\rho))_{\rho \in (0, \rho_\pi]}$ une paramétrisation du profil limite du DFS en fonction de la "probabilité de survie locale", les fonctions x et y vérifient les relations différentielles suivantes,

$$\begin{cases} \frac{x'(\rho)}{y'(\rho)} = \frac{2-\rho}{\rho}, \\ \frac{x'(\rho)+y'(\rho)}{2} = \alpha'(\rho). \end{cases}$$

La première égalité est le pendant continu de la relation (2.11) tandis que la seconde est la version continue de l'égalité *déterministe* $n + X_n = 2 \times$ "nombre de sommets explorés". En intégrant ces équations, on obtient le résultat suivant.

Théorème I : Chapitre 6, Théorème 10

Soit (X_n) le processus de contour du DFS, réalisé à partir d'un sommet choisi uniformément dans la composante géante. Alors la convergence suivant a lieu au sens de la topologie

uniforme

$$\forall t \in [0, 2\xi], \quad \lim_{N\to+\infty} \frac{X_{\lfloor tN \rfloor}}{N} = h(t),$$

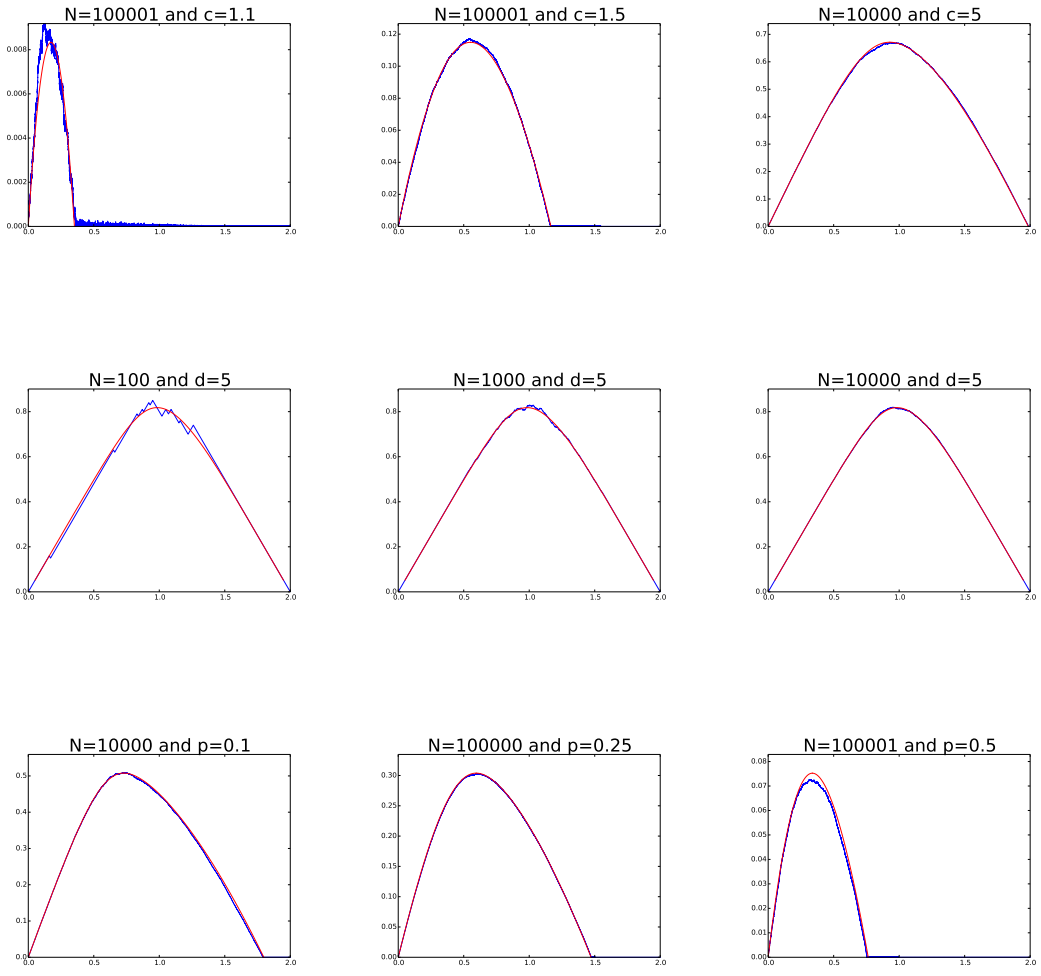
où la fonction h est bien définie et continue sur l'intervalle $[0, 2\xi]$. Le graphe $(t, h(t))_{t\in[0,2\rho_c]}$ peut être divisé en une première partie de montée et une deuxième partie de descente. Ces deux parties peuvent être respectivement paramétrée par $(x^{\uparrow}(\rho), y^{\uparrow}(\rho))_{\rho\in[0,\rho_\pi]}$ et $(x^{\downarrow}(\rho), y^{\downarrow}(\rho))_{\rho\in[0,\rho_\pi]}$, où :

$$\begin{cases} x^{\uparrow}(\rho) & := (2-\rho)\,\alpha(\rho) - \int_\rho^{\rho_\pi} \alpha(u)\mathrm{d}u, \\ y^{\uparrow}(\rho) & := \rho\,\alpha(\rho) + \int_\rho^{\rho_\pi} \alpha(u)\mathrm{d}u, \end{cases}$$

et

$$\begin{cases} x^{\downarrow}(\rho) := x^{\uparrow}(\rho) + 2\,(1-\alpha(\rho))\left(1 - g\big(\alpha(\rho), 1-\rho\big)\right), \\ y^{\downarrow}(\rho) := y^{\uparrow}(\rho). \end{cases}$$

Ci-dessous le lecteur trouvera des simulations numériques issues du Chapitre 6 et illustrant le Théorème I dans les cas où π est une loi de Poisson de paramètre $c > 1$, un Dirac en $d \geq 3$ ou une loi Géométrique partant de 0 et de paramètre $0 \leq p < 2/3$. Les courbes rouges correspondent à nos prédictions théoriques.



Finalement, l'évaluation de la hauteur du profil en $\rho = 0$ fournit une borne inférieure sur la longueur du plus long chemin simple dans le graphe.

Théorème J : Chapitre 6, Théorème 10

La longueur $\mathcal{H}_N$ du plus long chemin vérifie:

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N}{N} \geq y^{\uparrow}(0) - \varepsilon = \int_0^{\rho_\pi} \alpha(u)\mathrm{d}u - \varepsilon\right) \underset{N\to+\infty}{\longrightarrow} 1.$$

Dans les cas particuliers mentionnés au cours du Théorème H, on obtient des formules explicites :

- lorsque $\boldsymbol{\pi}$ est une loi de Poisson de paramètre $c > 1$, on retrouve (2.10) :

$$y^{\uparrow}(0) = \rho_c - \frac{Li_2(\rho_c)}{c},$$

- lorsque $\boldsymbol{\pi} = \delta_d$ pour un entier $d \geq 3$,

$$y^{\uparrow}(0) = 1 - \int_0^1 \left(\frac{1 - x^{\frac{1}{d-1}}}{1-x}\right)^{\frac{d}{d-2}} \mathrm{d}x,$$

- lorsque $\boldsymbol{\pi}$ est une loi Géométrique partant de 0 et de moyenne $0 < p < 2/3$,

$$y^{\uparrow}(0) = \rho_{\boldsymbol{\pi}} - \left(\frac{p}{1-p}\right)^{1/3} \int_0^{\rho_\pi} \left(\frac{1}{x+\sqrt{x}}\right)^{1/3} \mathrm{d}x.$$

Avec Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard, nous nous sommes rendu compte que les techniques développées au cours de l'article [EFMN19] s'appliquaient également à l'étude de variantes de l'algorithme de parcours en profondeur. L'une de ces variantes permet de construire de longs chemins induits dans le graphe et est analysée en détail au cours du Chapitre 7, issu de l'article en préparation [EFMN20]. Nous présentons brièvement les résultats correspondants dans le paragraphe qui suit.

Longs chemins induits. Un chemin simple $u_1, \ldots, u_l$ est appelé *induit* lorsque u_i et u_j ne sont pas reliés par une arête pour tout $|i-j| > 1$. Au cours du paragraphe précédent, nous avons brièvement mentionné les travaux de Frieze et Jackson [FJ87] au sujet de l'existence de tels chemins dans un modèle de configuration. Dans leur travail, les auteurs supposent uniquement que les degrés des sommets sont supérieurs à 3 et uniformément bornés. En effectuant l'hypothèse supplémentaire (**A1**), leur résultat peut s'énoncer de la manière suivante.

Théorème 26. *Supposons qu'il existe $\Delta \geq 3$ tel que pour tout $N \geq 1$ et tout $1 \leq i \leq N$, $3 \leq d_i^{(N)} \leq \Delta$. Notons $\mathcal{H}_N^{(1)}$ la longueur du plus long chemin induit au sein de $\mathscr{C}(\mathbf{d}^{(N)})$. Alors, en notant $q = q(\Delta) := (2\Delta-3)(2\Delta-4)$,*

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N^{(1)}}{N} \geq \frac{\sum_{i\geq 0} i\pi_i}{2\Delta} \frac{1}{\Delta - 2}\left(1 - q\log\left(1+\frac{1}{q}\right) - \varepsilon\right)\right) \underset{N\to+\infty}{\longrightarrow} 1. \tag{2.14}$$

Afin de démontrer ce Théorème, Frieze et Jackson ont introduit et étudié une variante de l'algorithme du DFS, où le marcheur rebrousse chemin dès qu'il réalise qu'un chemin induit a été révélé. Les lignées ancestrales de l'arbre construit par cette algorithme correspondent alors à des chemins induits dans le graphe, et les auteurs démontrent qu'avec grande probabilité, le marcheur réalise un nombre linéaire de "montées" successives.

Avec Nathanaël Enriquez, Gabriel Faraud et Laurent Ménard [EFMN20], nous avons introduit un autre algorithme construisant de longs chemins induits et se prêtant à une analyse détaillée. Ce dernier est une modification locale de l'algorithme du DFS, qui consiste à alterner parcours en profondeur et parcours en largeur. Plus précisément, à chaque fois qu'un nouveau sommet u est visité, la boule de taille deux et centrée en u au sein du graphe induit par les sommets non-explorés est révélée, et le marcheur explore ensuite les feuilles de cette boule selon un parcours en profondeur. Le lecteur se convaincra aisément que cet algorithme construit un arbre couvrant, planaire et enraciné de chaque composante connexe et que les lignées ancestrales de ces arbres correspondent à des chemins induits dans le graphe. À la différence de l'algorithme de Frieze et Jackson qui rebrousse chemin lorsqu'un sommet déjà exploré est revisité , notre algorithme est plus prévoyant puisqu'il exclu d'emblée les voisins des sommets visités du réservoir des sommets actifs. Pour cette raison, notre algorithme semble *a priori* moins performant que celui de Frieze et Jackson.

Cependant, les techniques développées dans le paragraphe précédent s'adaptent à ce nouveau contexte et permettent en particulier d'obtenir l'analogue des Théorèmes H, I et J. De manière remarquable, l'évolution de la loi des degrés au sein des graphes induits par les sommets non-explorés est la même que dans le cas de l'algorithme du DFS, dans l'échelle de temps "proportion de sommets explorés".

Théorème K : Chapitre 7, Théorèmes 14, 15, 17

On se place sous les hypothèses (**A1**), (**A2**), (**A3**) et $\nu > 1$. Soit $g(\alpha, s)$ la fonction définie au cours de l'Équation (2.13). Alors, avec probabilité tendant vers 1 lorsque N tend vers l'infini, on dispose des résultats suivants.

1. Pour tout $\alpha \in [0, \alpha_c)$, lorsque l'algorithme a révélé une proportion α de sommets, la loi asymptotique π_α des degrés au sein du graphe induit par les sommets non-explorés a pour série génératrice $g(\alpha, s)$.

2. Le processus $\left(X_{\lfloor tN \rfloor}/N\right)_{t\in[0,2]}$ converge vers un profile déterministe $(t, h(t))_{t\in[0,2]}$ au sens de la topologie uniforme. Ce profil limite peut être divisé en une phase ascendante et une phase descendante, respectivement paramétrisées par

$$\begin{cases} x^{\uparrow}(\rho) := \int_{\rho}^{\rho_\pi} \frac{(2-r)\alpha'(r)}{\partial_s \hat{g}(\alpha(r),1)} \mathrm{d}r, \\ y^{\uparrow}(\rho) := \int_{\rho}^{\rho_\pi} \frac{r\alpha'(r)}{\partial_s \hat{g}(\alpha(r),1)} \mathrm{d}r, \end{cases}$$

et

$$\begin{cases} x^{\downarrow}(\rho) := x^{\uparrow}(\rho) + 2(1-\alpha(\rho))\left(1 - g(\alpha(\rho), 1-\rho)\right), \\ y^{\downarrow}(\rho) := y^{\uparrow}(\rho), \end{cases}$$

où ρ varie au sein de l'intervalle $(0, \rho_\pi]$ pour chaque arc.

3. La longueur $\mathcal{H}_N^1$ du plus long chemin induit vérifie:

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N^1}{N} \geq y^{\uparrow}(0) = \int_0^{\rho_\pi} \frac{u\alpha'(u)}{\partial_s \hat{g}(\alpha(s),1)} \mathrm{d}u - \varepsilon \right) \underset{N\to+\infty}{\longrightarrow} 1. \tag{2.15}$$

Dans le cas particulier où π est une loi de Poisson de paramètre $c > 1$, on obtient :

$$y^{\uparrow}(0) = \frac{\rho_c}{-\ln(1-\rho_c)} \left(\gamma + \rho_c + \ln(-\ln(1-\rho_c)) - \mathrm{Li}_2(1-\rho_c)\right),$$

où $\gamma \approx 0.577\ldots$ est la constant d'Euler et Li_2 est la fonction Dilogarithme.
Lorsque $\pi = \delta_d$ pour un entier $d \geq 3$, on obtient :

$$y^{\uparrow}(0) = \frac{d}{2(d-1)} \left(1 - \int_0^1 \left(\frac{1 - x^{\frac{1}{d-1}}}{1-x} \right)^{\frac{2}{d-2}} \mathrm{d}x \right).$$

Bien que notre algorithme diffère de celui de Frieze et Jackson, il est possible de démontrer (voir Chapitre 7, partie 7.4) qu'avec grande probabilité, ces deux algorithmes construisent le même long chemin induit sur des modèles de configuration satisfaisant les hypothèses du Théorème K. Ainsi, la borne inférieure (2.15) est également valable pour l'algorithme de Frieze et Jackson dans le cas où l'hypothèse (**A1**) est satisfaite, et semble améliorer considérablement la minoration (2.14). Par exemple, dans le cas où $\pi = \delta_3$, (2.15) fournit une borne inférieure environ égale à 0,45 tandis que (2.14) fournit une borne inférieure environ égale à 0,07.

Mentionnons que la minoration (2.15) est aussi valable pour la longueur du plus long cycle induit. Cela provient du fait qu'avec grande probabilité, le long chemin induit construit au cours de notre algorithme peut être "refermé". Plus précisément, il est possible de démontrer qu'au cours de la phase ascendante de l'algorithme, si $L_n = \{v_1, \ldots, v_n\}$ désigne le long chemin induit en cours de construction, un réservoir de taille macroscopique de sommets distants de 1 de $\{v_1, \ldots, v_{\varepsilon N}\}$ est encore non-exploré par l'algorithme. Les détails techniques de ce raisonnement font l'objet de la partie 7.3.5 du Chapitre 7.

Terminons finalement cette partie en énonçant un dernier résultat du Chapitre 7, obtenu par une analyse analogue d'un algorithme alternant parcours en profondeur et parcours en largeur, et construisant de longs chemins *m-induits*, $m \geq 1$. Les détails font l'objet de la partie 7.5.

Théorème L : Chapitre 7, Proposition 10

Soit $m \geq 1$. Un chemin simple $u_1, \ldots, u_l$ est dit m-induit lorsque, pour tous sommets u_i et u_j séparés par k arêtes du chemin, la distance entre u_i et u_j au sein du graphe est minorée par $\inf\{m, k\}$. Notons $\mathcal{H}_N^m$ la longueur du plus long chemin m-induit au sein d'un modèle de configuration $\mathcal{C}(\mathbf{d}^{(N)})$ satisfaisant les hypothèses (**A1**), (**A2**), (**A3**) et $\nu > 1$. Alors,

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N^m}{N} \geq m \int_0^{\rho_\pi} \frac{r\,\alpha'(u)}{\sum_{j=1}^m (\partial_s \hat{g}(\alpha(u),1))^j} \mathrm{d}u - \varepsilon \right) \underset{N\to+\infty}{\longrightarrow} 1.$$

2.4 Spectre des graphes d'Erdős-Rényi

Soit $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ un graphe. La matrice d'adjacence $A = A(\mathrm{G})$ de G est une matrice symétrique indexée par V et dont les coefficients sont donnés par

$$\forall u, v \in \mathrm{V}, \quad A_{uv} = \mathbf{1}_{\{u,v\}} \in \mathrm{E}.$$

Autrement dit, le coefficient (u, v) vaut 1 si et seulement si $\{u, v\}$ est une arête du graphe. Bien entendu, la donnée de la matrice d'adjacence A est équivalente à celle de G. En revanche, la seule connaissance des valeurs propres de A ne permet pas de reconstruire G. Tout l'enjeu de la *théorie algébrique des graphes* consiste à étudier les propriétés capturées par leurs spectres. Le lecteur curieux pourra consulter les ouvrages de références [Big93, CDS95, BH12]. Dans cette dernière partie, nous nous intéresserons à l'étude du spectre des graphes d'Erdős-Rényi bipartis, où le Théorème A possède une application naturelle. Avant d'énoncer notre résultat, on commence par présenter quelques généralités sur les spectres de graphes aléatoires dilués.

Convergence locale et spectres de graphes. Pour tout $N \geq 1$, soit G_N un graphe aléatoire possédant N sommets. On définit la mesure spectrale empirique de G_N par

$$\mu_{\mathrm{G}_N} := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i^{(N)}},$$

où $\lambda_1^{(N)} \geq \cdots \geq \lambda_N^{(N)}$ sont les valeurs propres de $A(\mathrm{G}_N)$. Pour de nombreux modèles de graphes *denses*, où le nombre d'arêtes est d'ordre supérieur au nombre de sommets, on retrouve le régime classique des matrices aléatoires et la suite des mesures spectrales empiriques converge vers la loi du demi-cercle de Wigner. Dans le cas des graphes aléatoires *dilués*, où le nombre d'arêtes est proportionnel au nombre de sommets, la convergence de la mesure spectrale empirique peut être déduite de la convergence locale faible[4] de la suite $(\mathrm{G}_N)_{N\geq 1}$. Bien que cette convergence puisse être établie par d'autres méthodes, la limite locale faible permet en outre de caractériser la mesure limite, comme l'ont démontré Bordenave et Lelarge [BL10]. En particulier, lorsque le graphe enraciné limite est un arbre de Galton-Watson $\mathrm{T} = GW(\mathcal{P}(c))$ dont la loi de reproduction est une loi de Poisson de paramètre $c > 0$, il est possible de montrer que, en probabilité, μ_{G_N} converge faiblement vers

$$\mu_c := \mathbf{E}[\mu_{(\mathrm{T},\varnothing)}],$$

où $\mu_{(\mathrm{T},\varnothing)}$ est la mesure spectrale de T enracinée en la racine $\varnothing$. Celle-ci est caractérisé par ses moments : pour tout $k \geq 0$, le moment d'ordre k de $\mu_{(\mathrm{T},\varnothing)}$ est égal au nombre de chemins au plus proche voisin de $\varnothing$ à $\varnothing$ et de longueur k dans l'arbre. La propriété de récursivité des arbres de Galton-Watson permet en particulier de caractériser $\mu_{(\mathrm{T},\varnothing)}$ via une équation de récursion en loi sur sa transformée de Stieltjes. Soit $(s_i)_{i\geq 1}$ une famille de variables aléatoires indépendantes et de même loi que $s_{\mu_{(\mathrm{T},\varnothing)}}$, et N_c une variable aléatoire suivant une loi de Poisson de paramètre c, indépendante de $s_{\mu_{(\mathrm{T},\varnothing)}}$ et de la famille $(s_i)_{i\geq 1}$. On dispose de l'égalité en loi suivante :

$$\forall z \in \mathbf{C}_+, \quad s_{\mu_{(\mathrm{T},\varnothing)}}(z) = \frac{-1}{z + \sum_{i=1}^{N_c} s_i(z)}. \tag{2.16}$$

Notons que lorsque la limite locale est un arbre de Galton-Watson $GW(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$, une récursion similaire à (2.16) existe. En utilisant les Théorèmes 17 et 21, on en déduit que, en probabilité,

[4] rappelons au lecteur que cette notion a été introduite au cours de l'Équation (2.6)

- la mesure spectrale empirique $\mu_{\mathrm{ER}(N,c/N)}$ converge vers μ_c ;
- la mesure spectrale empirique $\mu_{\mathscr{C}(\mathbf{d}^{(N)})}$ converge vers une mesure de probabilité dépendant uniquement de $\boldsymbol{\pi}$, notée $\mu_{\boldsymbol{\pi}}$ dans ce qui suit.

L'étude des équations distributionnelles de la forme de (2.16) s'avère ardue et peu de propriétés des mesures de probabilités associées sont connues.

Le possible atome en 0 joue ici un rôle particulier pour deux raisons. D'abord, parce que le résidu en 0 de $s_{\mu_{(\mathrm{T},\varnothing)}}$ admet une récursion distributionnelle agréable que l'on peut déduire de (2.16). Ensuite, parce que l'atome en zéro d'un graphe peut être estimé à l'aide d'un algorithme d'élagage de ce graphe, d'abord proposé par Karp et Sipser [KS81]. En exploitant ces deux remarques, Bordenave, Lelarge et Salez [BLS11] ont pu expliciter les valeurs de $\mu_c(\{0\})$ et de $\mu_{\boldsymbol{\pi}}(\{0\})$. Bien qu'un tel calcul semble hors de portée pour les autres atomes, de très élégants arguments exploitant certaines invariances des arbres de Galton-Watson ont récemment permis à Salez [Sal20] d'étudier la partie atomique de la mesure $\mu_{\boldsymbol{\pi}}$. Dans un article précédent [Sal15], le même auteur avait démontré que pour tout $c > 0$, l'ensemble des atomes de μ_c correspond à l'ensemble des valeurs propres des arbres finis.

Que peut-on dire de la partie non-atomique de μ_c (resp. $\mu_{\boldsymbol{\pi}}$) ? Une conséquence des travaux de Bordenave, Virág et Sen [BSV17] est que celle-ci existe lorsque, avec probabilité positive, la limite locale $\mathrm{GW}(\mathcal{P}(c))$ (resp. $\mathrm{GW}(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})$) possède un sous-graphe isomorphe à $\mathbf{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ et passant par la racine, ce qui correspond exactement à la transition de phase pour l'existence d'une composante connexe de taille macroscopique. En particulier, la mesure μ_c admet une partie continue lorsque $c > 1$ et il est naturel d'étudier sa régularité. Dans cette direction, Coste et Salez [CS18] ont récemment démontré que μ_c possède des *états étendus* en 0 lorsque $c > \mathrm{e}$, autrement dit que le rapport $(\mu_c([-\varepsilon, \varepsilon]) - \mu_c(\{0\}))/\varepsilon$ ne tend pas vers 0 lorsque $\varepsilon \to 0$. Il est en fait communément conjecturé que lorsque $c > 1$, la partie continue de μ_c est absolument continue par rapport à la mesure de Lebesgue, bien qu'aucun résultat ne soit disponible à ce sujet pour le moment.

Dans ce qui suit, nous laissons cette question délicate de côté et considérons une autre manière d'interroger la mesure μ_c.

Le régime des grands c. Un autre angle d'attaque possible pour l'étude de μ_c consiste à analyser cette mesure lorsque c tend vers l'infini. En utilisant la loi des grands nombres dans l'Équation (2.16), il est aisé de déduire que $s_c(z) := c^{-1/2} s_{\mu_c}(c^{-1/2} z)$ converge ponctuellement vers l'unique solution de $s^2 + zs + 1 = 0$ satisfaisant $s(z) \sim -1/z$ lorsque $|z| \to +\infty$. Autrement dit, le poussé en avant de la mesure μ_c par la dilatation $x \mapsto xc^{-1/2}$, noté $\widetilde{\mu}_c$, converge faiblement vers la loi du demi-cercle. Cela pouvait se deviner puisque, plus c est grand, plus la matrice d'adjacence $A_{c,N}$ du graphe d'Erdős-Rényi $\mathrm{ER}(N, c/N)$ est "pleine" et se rapproche du régime classique des matrices aléatoires. Une question naturelle concerne alors l'étude de la suite formée par les mesures signées $\mu_{sc} - \widetilde{\mu}_c$. En exploitant la méthode des moments pour l'étude de la mesure spectrale empirique de la matrice $c^{-1/2} A_{c,N}$, Nathanaël Enriquez et Laurent Ménard [EM16] sont parvenus à identifier le terme d'ordre $1/c$.

Théorème 27. *Définissons $\mu_{sc}^{(1)}$ la mesure signée de masse nulle suivante :*

$$\mu_{sc}^{(1)}(\mathrm{d}x) := \frac{1}{2\pi} \frac{x^4 - 4x^2 + 2}{\sqrt{4 - x^2}} \mathbf{1}_{|x| \leq 2}.$$

Alors, lorsque $c \to +\infty$,

$$\forall k \geq 0, \quad \int_{\mathbf{R}} x^k \mathrm{d}\widetilde{\mu}_c(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_{sc}(x) + \frac{1}{c} \int_{\mathbf{R}} x^k \mathrm{d}\mu_{sc}^{(1)}(x) + o\left(\frac{1}{c}\right).$$

Spectres des graphes d'Erdős-Rényi bipartis. Nous allons maintenant présenter une seconde application du Théorème A permettant d'obtenir l'analogue du Théorème 27 dans le contexte des graphes d'Erdős-Rényi bipartis. Soient $N, M \in \mathbf{N}$ et $p \in [0,1]$. Le graphe d'Erdős-Rényi biparti de paramètres N, M, p est le graphe aléatoire obtenu après avoir effectué une percolation de paramètre p sur le graphe biparti complet de paramètres N et M. La matrice d'adjacence d'un tel graphe peut s'écrire sous la forme

$$A = A(N, M, p) = \begin{pmatrix} 0 & X_N^T \\ X_N & 0 \end{pmatrix},$$

où $X_N \in \mathbf{R}^{N \times M}$ est une matrice rectangulaire de taille $N \times M$ dont les coefficients sont i.i.d. et de loi de Bernoulli de paramètre p. Comme $\det(A - \lambda I_{N+M}) = \det(X_N X_N^T - \lambda^2 I_N)$, l'étude du spectre de la matrice de covariance aléatoire $X_N X_N^T$ est équivalente à celle du spectre de A. Lorsque $p = c/N$, la matrice renormalisée

$$W_N = \frac{1}{c} X_N X_N^T$$

satisfait les hypothèses du Théorème A et les quantités asymptotiques A_{2i} correspondantes vérifient :

$$\forall i \geq 1, \quad A_{2i} = c^{1-i}.$$

Notons $\mu_{\alpha,c}$ la limite des mesures spectrales empiriques μ_{W_N}. Il est possible d'obtenir l'analogue du Théorème B, qui dans ce nouveau contexte correspond à un développement asymptotique des moments de $\mu_{\alpha,c}$ lorsque c tend vers l'infini :

$$\forall k \geq 0, \quad \int_{\mathbf{R}} x^k \mathrm{d}\mu_{\alpha,c}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_{\mathrm{MP},\alpha}(x) + \frac{1}{c} \int_{\mathbf{R}} x^k \mathrm{d}\mu^{(1)}_{\mathrm{MP},\alpha}(x) + o\left(\frac{1}{c}\right),$$

où $\mu_{\mathrm{MP},\alpha}$ est la loi de Marchenko-Pastur, définie dans l'Équation (1.9), et où $\mu^{(1)}_{\mathrm{MP},\alpha}$ est la mesure signée de masse nulle définie dans l'Équation (1.18). En introduisant $\nu_{\alpha,c}$ la limite des mesures spectrales empiriques $\mu_{c^{-1}A(N,M,c/N)}$, il est finalement possible d'établir l'analogue du Théorème 27 dans le contexte des graphes d'Erdős-Rényi bipartis.

Théorème M : Chapitre 3, Corollaire 2

Lorsque $c \to \infty$:

$$\forall k \geq 0, \quad \int_{\mathbf{R}} x^k \mathrm{d}\nu_{\alpha,c}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\nu_{\alpha}(x) + \frac{1}{c} \int_{\mathbf{R}} x^k \mathrm{d}\nu^{(1)}_{\alpha}(x) + o\left(\frac{1}{c}\right),$$

où ν_α est la mesure de probabilité définie par :

$$\nu_\alpha(\mathrm{d}x) = \frac{\sqrt{(b-x^2)(x^2-a)}}{2\pi x} \mathbf{1}_{\sqrt{a} \leq |x| \leq \sqrt{b}} \mathrm{d}x + 1_{\alpha<1}(1-\alpha)\delta_0(\mathrm{d}x),$$

et $\nu^{(1)}_\alpha$ la mesure signé de masse nulle définie par :

$$\nu^{(1)}_\alpha(\mathrm{d}x) := \frac{x^5 - 2(\alpha+1)x^3 + (\alpha^2+1)x}{2\alpha\pi\sqrt{(b-x^2)(x^2-a)}} \mathbf{1}_{\sqrt{a}<|x|<\sqrt{b}}.$$

Part II

Publications and prepublications

Chapter 3

Spectral asymptotic expansion of Wishart matrices with exploding moments

This chapter corresponds to the publication [Noi18].

We study random covariance matrices whose entries have exploding moments meaning that the ratio between their k-th moment and the k-th power of their standard deviation goes to infinity with the size of the matrix. We compute an asymptotic expansion of the limiting spectral measure in the critical regime when this measure is close to the Marchenko-Pastur distribution. Explicit computations are given in the two classical cases of Bernoulli and truncated heavy tailed entries.

3.1 Introduction

Let X_n be a real random matrix of size $n \times m$ with i.i.d. entries which are centered and with second moment M_2. We define the Wishart matrix $W_n = \frac{1}{nM_2} X_n X_n^T$, where X_n^T is the transpose of X_n. The spectral measure of W_n is the random probability law:

$$\mu_{W_n} = \frac{1}{n} \sum_{\lambda \in \mathrm{Spec}(W_n)} \delta_\lambda,$$

where $\mathrm{Spec}(W_n)$ is the spectrum of W_n and δ_λ the Dirac delta function at λ. Since W_n is a positive symmetric matrix, its eigenvalues are nonnegative reals. The work of Marchenko and Pastur [MP67] implies that μ_{W_n} weakly converges to a probability law μ_α as $n, m \to +\infty$ and $m/n \to \alpha > 0$. The law μ_α is given by:

$$\mu_\alpha(\mathrm{d}x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x} \mathrm{d}x + \mathbf{1}_{\alpha<1} (1-\alpha)\, \delta_0(\mathrm{d}x),$$

where $a = (1-\sqrt{\alpha})^2$ and $b = (1+\sqrt{\alpha})^2$.

When the ratio between the k-th moment and the k-th power of the standard deviation of the entries goes to infinity with the size n, the limiting spectrum may not be μ_α. However, when this ratio is of order $n^{k/2-1}$, the spectral measure converges to a limiting distribution (see the work of Benaych-Georges and Cabanal-Duvillard [BGCD12] and Male [Mal17]) that can still be close to the Marchenko-Pastur law.

In this paper, we focus on two models. When X_n has Bernoulli entries with parameter c/n, $c > 0$, the limiting probability law $\mu_{\alpha,c}$ depends only on α and c. When the entries of X_n are in the domain of attraction of a β-stable law, one can truncate them at the n-th lowest and largest quantiles times a parameter $B > 0$. The resulting spectral law $\mu_{\alpha,\beta,B}$ depends only on α, β and B. In each case, except for the existence, the limiting spectral law remains poorly understood.

The main concern of this paper is to obtain an asymptotic expansion of $\mu_{\alpha,c}$ and $\mu_{\alpha,\beta,B}$ as $c \to +\infty$ and $B \to 0$. We propose first order formulas in term of moments. In each case, the leading term is the Marchenko-Pastur law with parameter α. More interestingly, the order one perturbation term involves in each case a signed measure $\mu_\alpha^{(1)}$ with total mass 0 for which we are able to obtain an explicit expression:

$$\mu_\alpha^{(1)}(\mathrm{d}x) = \frac{x^2 - 2(\alpha+1)x + (\alpha^2+1)}{2\alpha\pi\sqrt{(b-x)(x-a)}}\mathbf{1}_{(a,b)}(x)\mathrm{d}x. \tag{3.1}$$

These results are the content of Theorems 1 and 2. They suggest that somehow, $\mu_\alpha^{(1)}$ is a typical perturbative term when a sequence of measure converges to the Marchenko-Pastur law. In the Bernoulli case, $\mu_{\alpha,c}$ can be interpreted as some transform of the spectrum of a large bipartite Erdős-Rényi random graph with parameters n, $m = \alpha n$ and c/n. Therefore, our method provides a first order expansion of the limiting spectrum of sparse bipartite random graphs when c becomes large. This is the content of Corollary 1.

The work of Benaych-Georges and Cabanal-Duvillard [BGCD12, Theorem 3.2] provides a characterization of the limiting spectra in terms of moments. The combinatorics of the formula has the flavor of free probability and can be compared to the work of Ryan [Rya98] in the symmetric case. However, it does not lead to direct computations as we do here. Our proof is based on an alternative formula for the limiting moments obtained in Proposition 1. It involves a certain class of walks on rooted plane trees more amenable to analysis and can be compared to the work of Zakharevich [Zak06] for Wigner matrices.

Enriquez and Ménard [EM16] derived similar developments of moments for the limiting spectra of diluted random graphs. Their proof relies on a combinatorial analysis of the so-called local limit of the sequence of graphs. Although an analogous analysis could be done in the setting of bipartite random graphs, our method bypasses this argument and allows us to study both bipartite random graphs *and* truncated heavy tailed covariance matrices as particular cases of a more general formula, namely the combinatorial expression obtained for the limiting moments. We also cover the *weighted* cases, studied in the setting of bipartite random graphs by Vengerovsky [Ven14], where he obtained recursive expressions for the limiting spectral moments, which cannot lead to our expansions. This is the content of Corollaries 2 and 3.

Organization of the paper. In section 3.2, we state our main results concerning the first order asymptotic expansion of the limiting spectra of Bernoulli and truncated heavy-tailed covariance matrices. In section 3.3 we derive a new formula for the limiting moments of Wishart matrices with exploding moments and briefly explain how this leads to a new proof of the convergence of the spectral measure, given by Benaych-Georges and Cabanal-Duvillard [BGCD12, Theorem 3.2]. Section 3.4 is devoted to the proof of the results of section 3.2.

3.2 Main results

3.2.1 Spectra of Bipartite Erdős-Rényi Random Graphs

A bipartite Erdős-Rényi random graph with parameters n, m and p is a percolation with parameter p on the bipartite complete graph having parts of respective sizes n and m. Up to a relabeling of the vertices, its adjacency matrix can be written

$$A = A(n,m,p) = \begin{pmatrix} 0 & X_n^T \\ X_n & 0 \end{pmatrix},$$

where X_n is a $n \times m$ matrix with i.i.d. entries having Bernoulli law with parameter p. The associated empirical spectral measure is the random probability law

$$\nu_A = \frac{1}{n} \sum_{\lambda \in \mathrm{Spec}(A)} \delta_\lambda$$

that puts mass $1/n$ at each eigenvalue of A, counting multiplicities. Notice that

$$\mathrm{Spec}(A) = \left\{ \pm \sqrt{\lambda_i(X_n X_n^T)} \right\}_{1 \leq i \leq n}$$

where $\lambda_1(X_n X_n^T), \ldots, \lambda_n(X_n X_n^T)$ are the eigenvalues of $X_n X_n^T$. Let f be the bijection of $\mathbf{R}_+$: $f(x) = \sqrt{x}$. For a measure ν on $\mathbf{R}_+$, define $\mathrm{Sym}(\nu)(\cdot) = (\nu(\cdot) + \nu(-\cdot))/2$ the symmetrized version of ν. Then

$$\nu_A = (\mathrm{Sym} \circ f_*) \mu_{X_n X_n^T},$$

where $f_* \mu$ is the pushforward of a measure μ by f. This one-to-one correspondence between probability measures on $\mathbf{R}_+$ and symmetric probability measures on $\mathbf{R}$ allows us to directly work in Wishart's setting and study $\mu_{X_n X_n^T}$.

We are interested in the dilute regime $m = \alpha n$, $\alpha > 0$ and $p = c/n$, $c > 0$. In that case, the object of interest is the renormalized adjacency matrix $\frac{1}{\sqrt{c}} A$, so we study the following sequence of Wishart matrices:

$$W_n = \frac{1}{c} X_n X_n^T.$$

In this regime, a lot of entries of X_n are equal to zero, which prevents μ_{W_n} from converging to the Marchenko-Pastur law μ_α defined in the introduction. However, there still exists a limiting probability law $\mu_{\alpha,c}$ that only depends on α and c. This is a consequence of Benaych-Georges and Cabanal-Duvillard [BGCD12, Theorem 3.2]. To see this, it suffices to show that the centered version $W_n' = c^{-1} \left(X_n - \mathbb{E}[X_n]\right) \left(X_n - \mathbb{E}[X_n]\right)^T$ has the same limiting spectrum. Denoting respectively F and F' the cumulative distribution functions of μ_{W_n} and $\mu_{W_n'}$, a consequence of Lidskii's inequalities is that:

$$||F - F'||_\infty \leq \frac{\mathrm{rk}(W_n - W_n')}{n} = \frac{\mathrm{rk}\left(c^{-1}\mathbb{E}[X_n]^2\right)}{n} = \frac{1}{n},$$

where rk is the rank operator. Therefore, μ_{W_n} has the same limit as $\mu_{W_n'}$.

Another enlightening approach due to Bordenave and Lelarge [BL10] is directly concerned with $\nu_{\alpha,c} = (\mathrm{Sym} \circ f_*)\mu_{\alpha,c}$. The two authors observe that the spectrum of a graph is continuous with respect to the local convergence introduced in [AL07, BS01]. In our setting, it can be shown that the limiting local law $\mathscr{L}_{\alpha,c}$ is supported on unimodular random trees. More precisely,

$$\mathscr{L}_{\alpha,c} = \frac{1}{1+\alpha} \delta_{\mathbf{GW}(c,\alpha c)} + \frac{\alpha}{1+\alpha} \delta_{\mathbf{GW}(\alpha c, c)},$$

where $\mathbf{GW}(x, y)$ stands for a Galton-Watson tree where individuals in even (resp. odd) generations reproduce with Poisson law with parameter x (resp. y). The measure $\nu_{\alpha,c}$ is then the unique probability measure whose k-th moment is given by:

$$\int x^k \mathrm{d}\nu_{\alpha,c} = \mathbf{E}\left[|\{\text{neirest neighbor path of length } k \text{ from the root to the root of } \mathscr{L}_{\alpha,c}\}|\right]$$

Apart from this characterization, the measure $\nu_{\alpha,c}$ remains poorly understood. It can be proved, using same arguments as in the work of Salez [Sal15], that its set of atoms is dense in $\mathbf{R}$. Besides, a consequence of the work of Bordenave, Sen and Virág [BSV17] is that $\nu_{\alpha,c}$ possesses a continuous part if and only if $c > 1$. When $c \to +\infty$, $\nu_{\alpha,c}$ converges to $\nu_\alpha := (\mathrm{Sym} \circ f_*)\mu_\alpha$. A natural question is then to describe how $\nu_{\alpha,c}$ differs from ν_α as c becomes large. Our result provides a characterization of the perturbation of order $1/c$ in terms of moments of the Wishart counterpart $\mu_{\alpha,c}$.

Theorem 1. *For all $k \geq 1$, as $c \to \infty$:*

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\alpha,c}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha(x) + \frac{1}{c}\int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha^{(1)}(x) + o\left(\frac{1}{c}\right),$$

where $\mu_\alpha^{(1)}$ is defined in Equation (3.1).

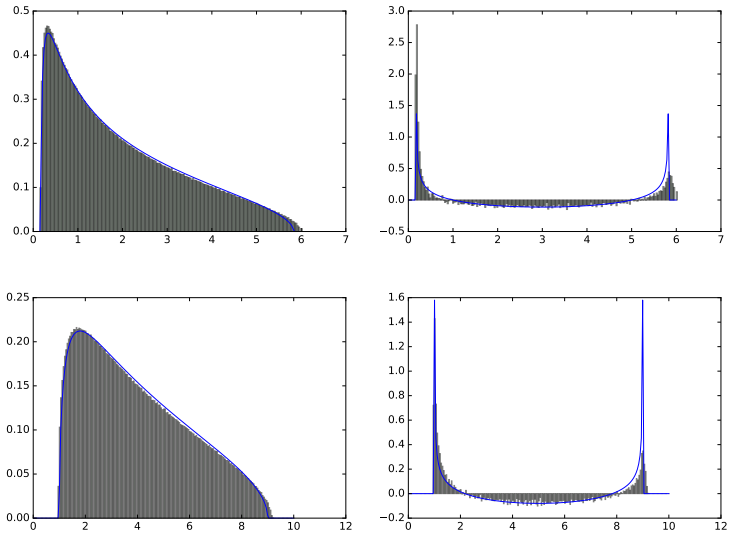


Figure 3.1 – Numerical simulations for the spectrum of 100 Wishart matrices associated to random matrices of size $n \times \alpha n$ with i.i.d. entries with Bernoulli law of parameter c/n, with $c = 20$ and $n = 3000$. The theoritical densities of μ_α and $\mu_\alpha^{(1)}$ are drawn in blue. The top diagrams correspond to $\alpha = 2$ whereas the bottom diagrams correspond to $\alpha = 4$.

It leads to an asymptotic expansion of the spectrum of large bipartite Erdős-Rényi random graph at large intensity c:

Corollary 1. *For all $k \geq 1$, as $c \to \infty$:*

$$\int_{\mathbf{R}} x^k \mathrm{d}\nu_{\alpha,c}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\nu_\alpha(x) + \frac{1}{c}\int_{\mathbf{R}} x^k \mathrm{d}\nu_\alpha^{(1)}(x) + o\left(\frac{1}{c}\right),$$

where $\nu_\alpha^{(1)} = (\mathrm{Sym} \circ f_*)\mu_\alpha^{(1)}$ *is given by the density:*

$$\frac{x^5 - 2(\alpha+1)x^3 + (\alpha^2+1)x}{2\alpha\pi\sqrt{(b-x^2)(x^2-a)}}\mathbf{1}_{\sqrt{a}<|x|<\sqrt{b}}.$$

Finally, let us mention that our method also applies in the *weighted* case where we study $Y_n(i,j) = X_n(i,j) \times \xi_n(i,j)$ with $\xi_n(i,j)$'s i.i.d. with a law that does not depend on n and having all moments finite. In that case, the perturbation term is just multiplied by $\mathbb{E}[\xi^4]/\mathbb{E}[\xi^2]$, where ξ has the same law as the $\xi_n(i,j)$'s.

Corollary 2. *In the weighted case, for all* $k \geq 1$, *as* $c \to \infty$*:*

$$\int_{\mathbf{R}} x^k \mathrm{d}\nu_{\alpha,c}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\nu_\alpha(x) + \frac{1}{c} \cdot \frac{\mathbb{E}[\xi^4]}{\mathbb{E}[\xi^2]} \int_{\mathbf{R}} x^k \mathrm{d}\nu_\alpha^{(1)}(x) + o\left(\frac{1}{c}\right).$$

3.2.2 Truncated heavy tailed random matrices

For all $n \geq 1$, let X_n be a rectangular random matrix of size $n \times m$ having i.i.d. entries with heavy tailed law P which has cumulative distribution function F. As before, we suppose that the ratio m/n converges to $\alpha > 0$. We suppose that P is in the domain of attraction of a β-stable law, $\beta < 2$. By [Bre92, Theorem 9.34], this implies that there exist two reals $M^-, M^+ \geq 0$ such that $M^- + M^+ > 0$ and

$$\frac{F(-x)}{1-F(x)} \underset{x\to+\infty}{\longrightarrow} \frac{M^-}{M^+}, \tag{3.2}$$

We restrict our study to the case $M^+ > 0$: the following arguments easily adapt when $M^- > 0$ by considering $P(-\cdot)$.

Theorem 1.10 in [BDG09] ensures that the spectral measure of

$$n^{-\frac{2}{\beta}} X_n X_n^T$$

almost surely weakly converges to a deterministic probability law $\mu_{\alpha,\beta}$ which only depends on α and β. We are here concerned with a truncated version. More precisely, define the quantiles q_n^- and q_n^+ by:

$$\begin{cases} F(q_n^-) &= 1/n \\ 1 - F(q_n^+) &= 1/n. \end{cases}$$

For all $B > 0$, we consider the sequence of random matrices $X_n^{(B)}$ of size $n \times m$ where

$$X_n^{(B)}(i,j) = X_n(i,j)\mathbf{1}_{Bq_n^- \leq X_n(i,j) \leq Bq_n^+} + Bq_n^- \mathbf{1}_{X_n(i,j) < Bq_n^-} + Bq_n^+ \mathbf{1}_{X_n(i,j) > Bq_n^+}.$$

Let $P_n^{(B)}$ be the law of the entries of $X_n^{(B)}$. This forms a sequence of probability measures that can be viewed as an approximation of P. The sequence of Wishart matrices:

$$\frac{1}{nM_2(P_n^{(B)})} X_n^{(B)} {X_n^{(B)}}^T,$$

where $M_2(P_n^{(B)})$ is the second moment of $P_n^{(B)}$, satisfies the hypothesis of [BGCD12, Theorem 3.2]. Therefore, the associated sequence of spectral measures converges to a probability law $\mu_{\alpha,\beta,B}$ which, as we will see, only depends on α, β and B.

The quantities q_n^- and q_n^+ correspond to the lowest and largest n-th quantile of P. Therefore, our choice of law P_n can be interpreted as a truncation of the largest entries in each row of X_n. If

one had chosen a smaller order of truncation, one would have retrieved the Marchenko-Pastur regime. On the contrary, if one had chosen a larger order of truncation, the A_k's defined in Equation (3.3) would have been all infinite, meaning that the truncation is not large enough to apply Theorem 3.2 of [BGCD12]. In this spirit, the parameter $B > 0$ can be seen as a finer adjustment of the truncation.

When $B \to 0$, we are able to obtain a first order expansion in terms of moments. Interestingly, it involves the signed measure $\mu_\alpha^{(1)}$ that also appears in the Bernoulli case.

Theorem 2. *For all $k \geq 1$, as $B \to 0$:*

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\alpha,\beta,B}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha(x) + \quad B^\beta \cdot C(\beta, M^+, M^-) \int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha^{(1)}(x) + o\left(B^\beta\right),$$

where $C(\beta, M^+, M^-) = \frac{(2-\beta)^2}{4-\beta} \cdot \frac{1}{1+(M^-/M^+)^{1/\beta}}$.

As in the Bernoulli case, our method directly applies to the weighted setting where we consider rectangular matrices $Y_n^{(B)}(i,j) = X_n^{(B)} \times \xi_n(i,j)$ with $\xi_n(i,j)$'s i.i.d. random variables, independent of $X_n^{(B)}$, whose law does not depend on n and has all moments finite. If ξ has the same law as this family, the corollary may be written in the following way.

Corollary 3. *In the weighted case, for all $k \geq 1$, as $B \to 0$:*

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\alpha,\beta,B}(x) = \int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha(x) + \quad B^\beta \cdot C(\beta, M^+, M^-) \frac{\mathbb{E}[\xi^4]}{\mathbb{E}[\xi^2]} \int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha^{(1)}(x) + o\left(B^\beta\right).$$

3.3 Spectral moments of generalized Wishart matrices

Our main results are obtained from a general formula that we derive for the limiting moments of size-dependent Wishart matrices. More precisely, we consider the following setting. Let X_n be an $n \times m$ matrix having i.i.d. entries with centered law P_n which has k-th moment $M_k(P_n) < +\infty$. We make the following assumption:

$$\forall k \geq 2, \quad \frac{M_k(P_n)}{n^{k/2-1} M_2(P_n)^{k/2}} \underset{n\to+\infty}{\longrightarrow} A_k \in [0, \infty). \tag{3.3}$$

In the regime $m/n \to \alpha > 0$, Benaych-Georges and Cabanal-Duvillard proved the convergence of the empirical spectral measures associated to the Wishart sequence:

$$W_n := \frac{1}{n M_2(P_n)} X_n X_n^T.$$

The limiting measure $\mu_{\mathcal{A}}$ only depends on $\mathcal{A} := (A_k)_{k\geq 2}$. They obtained a formula for the moments of $\mu_{\mathcal{A}}$ which has the flavor of free probability theory. We propose here an alternative formula which turns out to be more amenable to analysis for our purpose.

In order to properly state our result, we need to introduce the notion of word on a labeled graph. A labeled graph is a graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ together with a labeling of the vertices, that is a one-to-one function from V to $\{1, \ldots, |\mathrm{V}|\}$. A relabeling of a labeled graph is a new choice of bijection between V and $\{1, \ldots, |\mathrm{V}|\}$. Note that there are $|\mathrm{V}|!$ choices of labelings for a given graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$. A word of length $k \geq 1$ on a labeled graph G is a sequence of labels $i_1, i_2, \ldots, i_k$ such that $\{i_j, i_{j+1}\}$ is a pair of adjacent labels (that is the associated vertices are neighbours in G) for all $1 \leq j \leq k-1$. A word of length k is said to be closed if $i_1 = i_k$. Let $\mathbf{i} = i_1, \ldots, i_k$ and $\mathbf{i}' = i'_1, \ldots, i'_k$ be two words of length k on two labeled graphs G and G′ having the same number

of vertices. Then, $\mathbf{i}$ and $\mathbf{i}'$ are said to be equivalent if there exists a bijection σ of $\{1,\dots,|\mathrm{V}|\}$ such that $\sigma(i_j) = i'_j$ for all $1 \leq j \leq k$. In words, $\mathbf{i}$ and $\mathbf{i}'$ are equivalent if there exists a relabeling of a G such that the word associated to $\mathbf{i}$ is exactly $\mathbf{i}'$. One can check that this defines an equivalence relation on words on labeled graphs.

Recall that a rooted plane tree is a connected graph without cycles embedded in the plane, with a distinguished vertex called the root. In this context, we will always think of edges as oriented away from the root. A vertex at odd (resp. even) distance from the root will be called an odd (resp. even) vertex. An edge with an odd (resp. even) origin vertex will be called an even (resp. odd) edge. For instance, the edges attached to the root are odd.

Proposition 1. *For all $k \geq 1$, the k-th moment of $\mu_{\mathcal{A}}$ is*

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_{\mathcal{A}}(x) = \sum_{a=1}^{k} \sum_{l=1}^{a} \alpha^l \sum_{\substack{\mathbf{b}=(b_1,\dots,b_a) \\ b_1 \geq b_2 \geq \dots \geq b_a \geq 2 \\ b_1+b_2+\cdots+b_a=2k}} |\mathcal{W}_k(a, a+1, l, \mathbf{b})| \prod_{i=1}^{a} A_{b_i}. \tag{3.4}$$

where $\mathcal{W}_k(a, a+1, l, \mathbf{b})$ is a set of representatives of the equivalence classes of closed words on labeled rooted plane trees having "a" edges, of which l are odd edges, starting from the root and such that for all $1 \leq i \leq a$, one edge is traversed b_i times.

Remark 1. *Let $w \in \mathcal{W}_k(a, a+1, l, \mathbf{b})$. Since it is a representative walk on a tree, starting and ending at the root, the multiplicity of each edge has to be even. In particular, $\mathbf{b}$ must be an a-tuple of even integers summing to $2k$, and the sequence of even parameters $(A_{2k})_{k\geq 1}$ characterizes the limiting law.*

Proof. For all $k \geq 1$, denote by $M_k(\mu_{W_n})$ the k-th moment of μ_{W_n}. Its expected value $\mathbb{E}\left[M_k(\mu_{W_n})\right]$ is given by:

$$\frac{1}{n^{k+1} M_2(P_n)^k} \sum_{\substack{1 \leq i_1,\dots,i_k \leq n \\ 1 \leq j_1,\dots j_k \leq m}} \mathbb{E}[X(i_1,j_1)X(i_2,j_1)\cdots X(i_k,j_k)X(i_1,j_k)]. \tag{3.5}$$

Denote by $(\mathbf{i},\mathbf{j})$ the generic word $i_1 j_1 i_2 \dots i_1 j_k$ appearing in (3.5). We define the bipartite graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ associated to the word $(\mathbf{i},\mathbf{j})$ by:

$$\begin{cases} \mathrm{V} = \{(i_r, \mathbf{i}), (j_r, \mathbf{j});\ 1 \leq r \leq k\}, \\ \mathrm{E} = \big\{\{(i_r, \mathbf{i}), (j_r, \mathbf{j})\}, \{(i_{r+1}, \mathbf{i}), (j_r, \mathbf{j})\};\ 1 \leq r \leq k\big\}, \end{cases}$$

where we used the convention $k+1 = 1$. The abstract symbols $\mathbf{j}$ and $\mathbf{j}$ are needed to obtain a bipartite graph since the i_r's and j_r's can have common values (see Figure 3.2 for an illustration). We will refer to (**i**-) and (**j**-)letters. In words, the vertices of G are the letters of the word $(\mathbf{i},\mathbf{j})$ and two vertices are linked by an edge when they are consecutive in $(\mathbf{i},\mathbf{j})$. Denote by s the number of vertices, a the number of edges, l the number of **j**-vertices and $\bar{l}$ the number of **i**-vertices in the word. Since G is connected, $s \leq a+1$. Moreover, since P_n has zero mean, each edge must appear at least twice in the word to give a non-zero contribution in (3.5). As a consequence we obtain the bound $a \leq k$ because $i_1 j_1 \dots j_k$ possesses $2k$ edges counted with multiplicity.

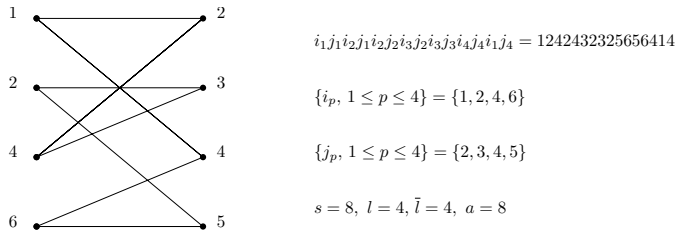


Figure 3.2 – Example of a word $(\mathbf{i},\mathbf{j})$ with its associated graph and quantities.

Two words $(\mathbf{i},\mathbf{j})$ and $(\mathbf{i}',\mathbf{j}')$ are said to be equivalent if one can find a permutation σ of $\{1,\dots,n\}$ and another one τ of $\{1,\dots,m\}$ such that

$$\forall p \in \{1,\dots k\}, \quad \sigma(i_p) = i'_p \text{ and } \tau(j_p) = j'_p.$$

One can check that this is an equivalence relation on the words appearing in (3.5). Note that $(\mathbf{i},\mathbf{j})$ has

$$C(s,l) = n(n-1)\cdots(n-s+l+1) \times m(m-1)\cdots(m-l+1) \sim \alpha^l n^s$$

equivalents. Fix $a \in \{1,\dots,k\}$, $1 \leq s \leq a+1$ and $1 \leq l \leq a$. Let $\mathcal{B}_{a,k}$ be the set of a-tuples $\mathbf{b} = (b_1,\dots,b_a)$ of integers such that

1. $b_1 \geq b_2 \geq \cdots \geq b_a \geq 2$;
2. $b_1 + \cdots + b_a = 2k$.

For all $k \geq 1$ and $\mathbf{b} \in \mathcal{B}_{a,k}$, we introduce $\mathcal{W}_k(a,s,l,\mathbf{b})$ a set of representatives of the equivalence classes of words $(\mathbf{i},\mathbf{j})$ such that the associated graph has a edges and s vertices, of which l are $\mathbf{j}$-vertices and such that for all $1 \leq i \leq a$ there is an edge which has multiplicity b_i in $(\mathbf{i},\mathbf{j})$. We can rewrite (3.5) as:

$$\sum_{a=1}^{k} \sum_{s=1}^{a+1} \sum_{l=1}^{s} \frac{C(s,l)}{n^{a+1}} \sum_{\mathbf{b}\in\mathcal{B}_{a,k}} \sum_{(\mathbf{i},\mathbf{j})\in\mathcal{W}(a,s,l,b)} \prod_{1\leq i\leq a} \frac{M_{b_i}(P_n)}{n^{b_i/2-1} M_2(P_n)^{b_i/2}}. \tag{3.6}$$

Since $C(s,l)n^{-a-1} \sim \alpha^l n^{s-a-1}$ when $n \to +\infty$ we deduce that when $s < a+1$ the asymptotic contribution is zero. Hence a possible non-zero contribution arises only when $s = a+1$. The Proposition is a consequence of (3.3). □

Note that our method could lead to an alternative proof of Benaych-Georges and Cabanal-Duvillard's result. We only give a sketch of the remaining steps and do not enter the details here as it is not our main purpose. The variance of $M_k(\mu_{W_n})$ can be written

$$\frac{1}{n^{2(k+1)} M_2(P_n)^{2k}} \sum_{(\mathbf{i},\mathbf{j}),(\mathbf{i}',\mathbf{j}')} \Big(\mathbb{E}[P(\mathbf{i},\mathbf{j})P(\mathbf{i}',\mathbf{j}')] - \mathbb{E}[P(\mathbf{i},\mathbf{j})]\mathbb{E}[P(\mathbf{i}',\mathbf{j}')] \Big),$$

where $P(\mathbf{i},\mathbf{j})$ is the product $X(i_1,j_1)X(i_2,j_1)\cdots X(i_1,j_k)$. A combinatorial analysis of the contributions associated to words $(\mathbf{i},\mathbf{j})$ and $(\mathbf{i}',\mathbf{j}')$ can lead to $\mathrm{Var}(M_k(\mu_{W_n})) = O(1/n)$. Then, it can be shown that the limiting k-th moment does not grow faster than k^{ck} for a constant $c > 0$, which ensures that the limiting law is characterized by its moments.

3.4 Proof of the main results

The formula for the moments of the limiting law can be rewritten:

$$\sum_{l=1}^{k} \alpha^l |\mathcal{W}_k(k,k+1,l,(2,\dots,2))| + A_4 \sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k(k,k+1,l,(4,2,\dots,2))| + \cdots \tag{3.7}$$

We used that $A_2 = 1$ and that:

- a word of length $2k+1$ on a rooted plane tree having k edges that starts and ends at the root and traverses every edge *has to* traverse every edge exactly twice;

- a word of length $2k+1$ on a rooted plane tree having $k-1$ edges that starts and ends at the root and traverses every edge *has to* traverse every edge twice except for one that is traversed 4 times.

The (numerous) remaining terms (represented by the "$\cdots$") involve the A_k's for $k > 4$.
The first term of (3.7) is known to be the k-th moment of the Marchenko-Pastur law with parameter α. We will show that the second term

$$\sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k(k, k+1, l, (4, 2, \dots, 2))|$$

is the k-th moment of $\mu_\alpha^{(1)}$. Finally, we will see that A_4 is equal to $1/c$ in the Bernoulli case and to $B^\beta \cdot C(\beta, M^+, M^-)$ in the truncated heavy-tailed case, by identifying the asymptotic coefficients of the A_k's in both settings.

3.4.1 Identification of the signed measure

Although this is a known result, we briefly prove that

$$\int_{\mathbf{R}} x^k \mathrm{d}\mu_\alpha(x) = \sum_{l=1}^{k} \alpha^l |\mathcal{W}_k(k, k+1, l, (2, \dots, 2))|$$

in order to introduce some notation.

Recall that $\mathcal{W}_k(k, k+1, l, (2, \dots, 2))$ is a set of representatives of closed words starting at the root, of length $2k+1$ on labeled rooted plane trees having k edges, l of these being odd edges. Therefore,

$$\sum_{l=1}^{k} \alpha^l |\mathcal{W}_k(k, k+1, l, (2, \dots, 2))| = \sum_{\mathrm{T} \in \mathcal{T}_k} \alpha^{l(\mathrm{T})},$$

where $\mathcal{T}_k$ is the set of rooted plane trees having k edges. For convenience, we introduce

$$a_k := \sum_{\mathrm{T} \in \mathcal{T}_k} \alpha^{l(\mathrm{T})} \quad \text{and} \quad b_k := \sum_{\mathrm{T} \in \mathcal{T}_k} \alpha^{\bar{l}(\mathrm{T})},$$

where $\bar{l}(\mathrm{T})$ is the number of even edges of a given tree $\mathrm{T} \in \mathcal{T}_k$, which satisfies $\bar{l}(\mathrm{T}) = k - l(\mathrm{T})$.

It turns out that the a_k's are the moments of μ_α. To obtain the term of order $1/c$ we will need to compute the generating series of the a_k's and b_k's.

Let T be a rooted plane tree having $k+1$ edges. Let T_1 be the tree induced by the first child of the root and T_2 the connected component of the root after removing the edge between the root and its first child (see Figure 3.3).

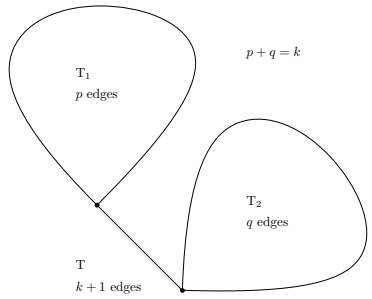


Figure 3.3 – Decomposition of a rooted plane tree.

Denoting p (resp. q) the number of edges of T_1 (resp. T_2), we have $p+q=k$. It is straightforward to obtain the relations $l(\mathrm{T}) = 1 + \bar{l}(\mathrm{T}_1) + l(\mathrm{T}_2)$ and $\bar{l}(\mathrm{T}) = l(\mathrm{T}_1) + \bar{l}(\mathrm{T}_2)$. Therefore

$$\begin{cases} a_{k+1} = \alpha \sum\limits_{p+q=k} a_p b_q \\ b_{k+1} = \sum\limits_{p+q=k} a_p b_q. \end{cases}$$

Denoting $A(z) = \sum_{k\geq 0} a_k z^k$ and $B(z) = \sum_{k\geq 0} b_k z^k$ the generating functions of the a_k's and the b_k's we obtain the functional relations:

$$\begin{cases} A = 1 + \alpha z A B \\ B = 1 + z A B. \end{cases} \tag{3.8}$$

These imply that $zA^2 + (\alpha z - z - 1)A + 1 = 0$. If we denote by $S(z) := -z^{-1}A(z^{-1})$ the Stieltjes transform of the measure with moments a_k's, then S satisfies the equation:

$$zS^2 - (\alpha - z - 1)S + 1 = 0. \tag{3.9}$$

The function S of the variable $z \in \mathbf{C}_+$ is the limit of the Stieltjes transform of the μ_{W_n} when $c \to +\infty$. The imaginary part of a Stieltjes transform is positive: this allows us to choose the right solution for equation (3.9). For a complex z, if we denote $\sqrt{z}$ the square root having a positive imaginary part on the upper half plane:

$$S(z) = \frac{\alpha - z - 1 + \sqrt{(z-b)(z-a)}}{2z},$$

where $a = (1-\sqrt{\alpha})^2$ and $b = (1+\sqrt{\alpha})^2$. This is the Stieltjes transform of the Marchenko-Pastur law μ_α, as announced.

The second term involves $\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))$ a set of equivalence classes of closed words of length $2k+1$ on labeled rooted plane tree having $k-1$ edges, starting at the root and such that each edge is traversed exactly two times except one which is traversed four times. Let us denote

$$a_k^{(1)} = \sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))|,$$

and

$$b_k^{(1)} = \sum_{\bar{l}=1}^{k-1} \alpha^l |\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))|.$$

The associated generating series will be denoted $A^{(1)}$ and $B^{(1)}$. Notice that by definition $a_0^{(1)} = a_1^{(1)} = b_0^{(1)} = b_1^{(1)} = 0$. We are going to obtain a recursion linking the four generating series $A, B, A^{(1)}$ and $B^{(1)}$. The idea is to use a first generation decomposition of the rooted plane tree on which the words are written, and then to distinguish whether or not the quadruple edge is an edge of this generation. For all $k \geq 1$, we partition $\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))$ into two parts:

$$\mathcal{W}_k^{(0)}(k-1, k, l, (4, 2, \ldots, 2)) \bigsqcup \mathcal{W}_k^{(1)}(k-1, k, l, (4, 2, \ldots, 2)),$$

where $\mathcal{W}_k^{(0)}(k-1, k, l, (4, 2, \ldots, 2))$ is the set of representative belonging to $\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))$ such that the quadruple edge is not a first generation edge, and $\mathcal{W}_k^{(1)}(k-1, k, l, (4, 2, \ldots, 2))$ is the set of representatives belonging to $\mathcal{W}_k(k-1, k, l, (4, 2, \ldots, 2))$ such that the quadruple edge

is a first generation edge. The associated quantities will be denoted $a_k^{(1,0)}, a_k^{(1,1)}, A^{(1,0)}, \dots$ For example:

$$a_k^{(1,0)} = \sum_{l=1}^{k-1} \alpha^l |\mathcal{W}_k^{(0)}(k-1, k, l, (4, 2, \dots, 2))|.$$

A representative word $(\mathbf{i}, \mathbf{j}) \in \mathcal{W}_k(k-1, k, l, (4, 2, \dots, 2))$ can be written:

$$(\mathbf{i}, \mathbf{j}) = i_1 \mathbf{S}_1 \zeta \xi \mathbf{S}_2 \xi \zeta \mathbf{S}_3 \zeta \xi \mathbf{S}_4 \xi \zeta \mathbf{S}_5 i_1,$$

where:

1. $i_1 \mathbf{S}_1 \zeta \mathbf{S}_5 i_1$ is the contour of a rooted plane tree having p_1 edges;
2. $\xi \mathbf{S}_2 \xi$ is the contour of a rooted plane tree having p_2 edges;
3. $\zeta \mathbf{S}_3 \zeta$ is the contour of a rooted plane tree having p_3 edges;
4. $\xi \mathbf{S}_4 \xi$ is the contour of a rooted plane tree having p_4 edges;
5. $\xi \mathbf{S}_2 \xi \mathbf{S}_4 \xi$ is the contour of a rooted plane tree having $p_2 + p_4$ edges.

The above integers satisfy $p_1 + p_2 + p_3 + p_4 = k - 2$. See Figure 3.4 for an illustration.

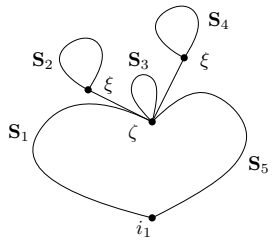


Figure 3.4 – The writing $(\mathbf{i}, \mathbf{j})$ and its quadruple edge $\{\zeta, \xi\}$.

All of these conditions are sufficient to define a class of canonical representatives. Let T be the rooted plane tree on which a representative word $(\mathbf{i}, \mathbf{j})$ is written. Denote e_4 the quadruple edge, $\mathrm{T} \setminus e_4$ the connected component of the root after removing e_4 and T^{e_4} the rooted plane tree formed by the descendants of e_4. Then, the above conditions ensures that $(\mathbf{i}, \mathbf{j})$ is such that $\mathrm{T} \setminus e_4$ and T^{e_4} are respectively traversed in lexicographic order.

Let $(\mathbf{i}, \mathbf{j}) \in \mathcal{W}_k^{(0)}(k-1, k, l, (4, 2, \dots, 2))$. The underlying tree can have $p \in \{1, \dots, k-2\}$ edges which are all traversed twice by $(\mathbf{i}, \mathbf{j})$. One of the trees induced by the children of the root contains the quadruple edge, leading to p different choices. On the other hand, if $(\mathbf{i}, \mathbf{j}) \in \mathcal{W}_k^{(1)}(k-1, k, l, (4, 2, \dots, 2))$ then the underlying tree can have $p \in \{1, \dots, k-1\}$ edges out of which one is the quadruple edge. There are $\binom{p+1}{2}$ choices for the locations of the the visits of the quadruple edge. See Figure 3.5 for an illustration.

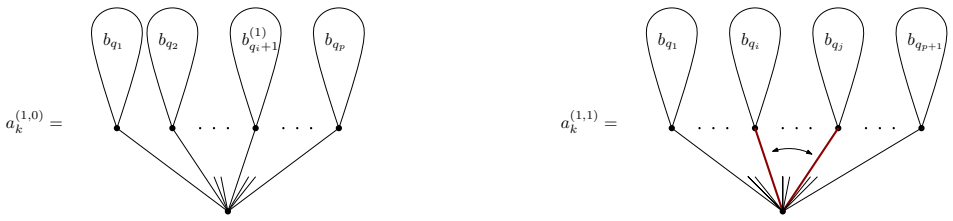


Figure 3.5 – First edge decomposition of a word respectively in $\mathcal{W}_k^{(0)}(k-1, k, l, (4, 2, \dots, 2))$ on the left and in $\mathcal{W}_k^{(1)}(k-1, k, l, (4, 2, \dots, 2))$ on the right, where the quadruple edge is in red.

As a consequence, we get the following recursions:

$$a_k^{(1,0)} = \sum_{p=1}^{k-2} \alpha^p p \sum_{q_1+\cdots+q_p=k-p-1} b^{(1)}_{q_1+1} b_{q_2} \cdots b_{q_p},$$

and

$$a_k^{(1,1)} = \sum_{p=1}^{k-2} \alpha^p \binom{p+1}{2} \sum_{q_1+\cdots+q_{p+1}=k-p-1} b_{q_1} b_{q_2} \cdots b_{q_{p+1}}.$$

This yields

$$A^{(1,0)} = \frac{\alpha z B^{(1)}}{(1-\alpha z B)^2} = \alpha z A^2 B^{(1)}$$

and

$$A^{(1,1)} = \frac{\alpha z^2 B^2}{(1-\alpha z B)^3} = \alpha z^2 A^3 B^2,$$

where we used equation (3.8). The same arguments and computations give $B^{(1,0)} = zA^{(1)}B^2$ and $B^{(1,1)} = z^2A^2B^3$, to finally obtain

$$\begin{cases} A^{(1)} = \alpha z A^2 B^{(1)} + \alpha z^2 A^3 B^2 \\ B^{(1)} = zA^{(1)} B^2 + z^2 A^2 B^3. \end{cases}$$

We deduce, using equation (3.8), that $A^{(1)}$ is given by:

$$A^{(1)} = \frac{\alpha (zAB)^2}{1-\alpha(zAB)^2}(zA^2B + A) = \frac{AB}{1-\alpha(zAB)^2}\alpha(zAB)^2. \tag{3.10}$$

To obtain a more explicit formula for $A^{(1)}$, one can compute $\alpha(zAB)^2$ using first that $B = (A+\alpha-1)/\alpha$ and then that $zA^2 = (1-(\alpha-1)z)A - 1$. After simplifications:

$$\begin{aligned} \alpha(zAB)^2 &= \frac{(1-\alpha z - z)A + z - 1}{\alpha z} \\ &= \frac{(\alpha^2+1)z^2 - 2z(\alpha-1) + 1 - (1-\alpha z - z)\sqrt{\delta}}{2\alpha z^2}, \end{aligned} \tag{3.11}$$

since $A = (2z)^{-1}(1-(\alpha-1)z - \sqrt{\delta})$. Using $\sqrt{\delta} = -2zA - (\alpha-1)z + 1$, one can then check that $\sqrt{\delta}AB = 1 - \alpha(zAB)^2$. From (3.11), we can finally rewrite (3.10) as

$$A^{(1)} = \frac{1}{\sqrt{\delta}} \frac{(\alpha^2+1)z^2 - 2z(\alpha+1) + 1 - (1-\alpha z - z)\sqrt{\delta}}{2\alpha z^2}.$$

Therefore, the function $S^{(1)}(z) = -\frac{1}{z}A^{(1)}(\frac{1}{z})$ is given by

$$S^{(1)}(z) = -\frac{z^2 - 2z(\alpha+1) + (\alpha^2+1)}{2\alpha\sqrt{(z-b)(z-a)}} + \frac{z-\alpha-1}{2\alpha}.$$

This corresponds to the Stieltjes transform of the measure $\mu_\alpha^{(1)}$ with density:

$$\frac{1}{\pi}\lim_{\varepsilon\to 0} \operatorname{Im}\big(S^{(1)}(x+i\varepsilon)\big) = \frac{x^2 - 2x(\alpha+1) + (\alpha^2+1)}{2\alpha\pi\sqrt{(b-x)(x-a)}} \mathbf{1}_{(a,b)}.$$

3.4.2 Asymptotic coefficients

By Remark 1, it is sufficient to take l even.

The Bernoulli case. In this setting the computation is direct. As explained in Section 3.2.1, it suffices to consider the centered version of Bernoulli laws, which corresponds to taking:

$$P_n = \frac{c}{n}\delta_{1-c/n} + \left(1 - \frac{c}{n}\right)\delta_{-c/n}.$$

Therefore, for all $k \geq 1$:

$$\frac{M_{2k}(P_n)}{n^{k-1}M_2(P_n)^k} \underset{n\to+\infty}{\longrightarrow} c^{1-k}.$$

In particular $A_4 = 1/c$, which, combined with the above identification of $\mu_\alpha^{(1)}$, provides the proof of Theorem 1.

The truncated heavy-tailed case. Recall that P is in the domain of attraction of a β-stable law, which implies that its cumulative distribution function F satisfies Equation (3.2), where we supposed that $M^+ > 0$. In this regime, Theorem 9.34 of the book of Breiman [Bre92] provides

$$\frac{1 - F(\xi x)}{1 - F(x)} \underset{x\to+\infty}{\longrightarrow} \xi^{-\beta},$$

meaning that $1 - F(\cdot)$ varies regularly with exponent $-\beta$. This implies that for all $k \geq 2$, the truncated moment function

$$U_k(x) := \int_0^x t^k \mathrm{d}P(t)$$

varies regularly with exponent $k - \beta$. Theorem 2 of [Fel71, VIII.9], known as Karamata's estimate, yields the following asymptotic as $x \to +\infty$:

$$U_k(x) \sim \frac{\beta}{k-\beta} x^k (1 - F(x)).$$

The behavior of the left truncated moment $\overline{U}_k(x) = \int_{-x}^0 t^k \mathrm{d}P(t)$ can be obtained in the same way. As $x \to +\infty$:

$$\overline{U}_k(x) \sim \frac{\beta}{k-\beta}(-x)^k F(-x)\mathbf{1}_{M^->0}.$$

Therefore, for all $k \geq 1$:

$$\begin{aligned} M_{2k}\left(P_n^{(B)}\right) &= U_{2k}(Bq_n^+) + \frac{(Bq_n^+)^{2k}}{n} + \overline{U}_{2k}(Bq_n^-) + \frac{(Bq_n^-)^{2k}}{n} \\ &\sim \frac{(Bq_n^+)^{2k}}{n} B^{-\beta}\left(1 + \frac{\beta}{2k-\beta}\right) + \frac{(Bq_n^-)^{2k}}{n} B^{-\beta}\left(1 + \frac{\beta}{2k-\beta}\right)\mathbf{1}_{M^->0} \\ &\sim \frac{(q_n^+)^{2k}}{n} \cdot B^{-\beta} \cdot \frac{2k}{2k-\beta}\left(1 + \left(\frac{M^-}{M^+}\right)^{1/\beta}\right) \end{aligned}$$

because on the event $M^- > 0$,

$$\frac{1 - F(q_n^+)}{1 - F\left(\left(\frac{M^+}{M^-}\right)^{1/\beta} q_n^-\right)} \underset{n\to+\infty}{\longrightarrow} 1,$$

which yields $q_n^- \sim (M^-/M^+)^{1/\beta} q_n^+$. We finally obtain:

$$A_{2k} = B^{\beta(k-1)} \cdot \frac{2k}{2k-\beta} \left(\frac{2-\beta}{2}\right)^k \left(1 + \left(\frac{M^-}{M^+}\right)^{1/\beta}\right)^{1-k}.$$

In particular, $A_4 = B^\beta \cdot C(\beta, M^+, M^-)$ which leads to Theorem 2.

Chapter 4

Spectral Measures Of Spiked Random Matrices

This chapter corresponds to the publication [Noi20].

We study two spiked models of random matrices under general frameworks corresponding respectively to additive deformation of random symmetric matrices and multiplicative perturbation of random covariance matrices. In both cases, the limiting spectral measure in the direction of an eigenvector of the perturbation leads to old and new results on the coordinates of eigenvectors.

4.1 Introduction

The study of deformed models of random matrices has been the subject of tremendous number of papers in the last decades. In this paper, we study two of them. The first one corresponds to an additive perturbation of a symmetric random matrix (or Wigner matrix):

$$W_n = \frac{1}{\sqrt{n}} X_n + A_n$$

where

- X_n is a random symmetric matrix of size $n \times n$, whose entries are, up to symmetry, i.i.d. centered and with variance 1;
- A_n is a deterministic symmetric matrix of size $n \times n$ (or random, independent of X_n).

The second one is a multiplicative deformation of a random covariance matrix (or Wishart matrix):

$$S_n = \frac{1}{n} \Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2},$$

where

- X_n is a random matrix of size $n \times m$, $m/n \to \alpha \in (0, \infty)$, whose entries are i.i.d. centered and with variance 1;
- Σ_n is a deterministic symmetric matrix (or random, independent of X_n) having non-negative eigenvalues.

The spectra of these models have been well studied. Let $\mathrm{Spec}(W_n)$ (resp. $\mathrm{Spec}(S_n)$) be the set of eigenvalues of W_n (resp. S_n) counted with multiplicities. The empirical spectral measures of W_n and S_n are:

$$\mu_{W_n} = \frac{1}{n} \sum_{\lambda \in \mathrm{Spec}(W_n)} \delta_\lambda \quad \text{and} \quad \mu_{S_n} = \frac{1}{n} \sum_{\lambda \in \mathrm{Spec}(S_n)} \delta_\lambda.$$

Under mild assumptions, they converge respectively towards probability measures $\mu_{sc} \boxplus \mu_A$ and $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ properly defined in Subsections 4.2.1 and 4.3.1.

In both models, we define an outlier as an eigenvalue that does not lie in the neighborhood of the support of the limiting spectrum. Many works identified necessary and sufficient conditions on the spectrum of A_n (resp. Σ_n) for the appearance of outliers in the spectrum of W_n (resp. S_n). The seminal paper is due to Baik, Ben Arous and Péché [BBAP05], who identified a phase transition for the existence of an outlier in the covariance setting with $\Sigma_n = \mathrm{Diag}(\theta, 1, \dots, 1)$, $\theta \geq 1$. The most recent results can be found in [BBCF17] and we refer to the survey [CDM17] for an extensive bibliography. In all previous approaches, two main techniques were used: a clever identity on determinants (first remarked in [BGN11]), and a precise analysis of the empirical spectral measure.

The goal of this paper is to bring into focus the possible use of the spectral measures in the study of deformed models of random matrices. Let us introduce them. Let θ be an eigenvalue of A_n (or Σ_n) which we consider to be atypical in that it may be responsible for the existence of an outlier. Denote $v_1^{(n)}$ the associated eigenvector. We will call θ a spike of W_n (resp. S_n) and $v_1^{(n)}$ the direction of the spike. The spectral measures in the direction of the spike are respectively defined by:

$$\mu_{(W_n, v_1^{(n)})} := \sum_{\lambda \in \mathrm{Spec}(W_n)} \left| \langle \phi_\lambda, v_1^{(n)} \rangle \right|^2 \delta_\lambda \quad \text{and} \quad \mu_{(S_n, v_1^{(n)})} := \sum_{\lambda \in \mathrm{Spec}(S_n)} \left| \langle \phi_\lambda, v_1^{(n)} \rangle \right|^2 \delta_\lambda,$$

where ϕ_λ is a normalized eigenvector associated to eigenvalue λ. Note that unlike empirical spectral measures, these probability measures contain information on the eigenvectors of W_n and S_n. Following the well-known observation that outliers have associated eigenvectors which are localized in the direction of the spike, their influence should be present in $\mu_{(W_n, v_1^{(n)})}$ and $\mu_{(S_n, v_1^{(n)})}$ at a macroscopic level. Moreover, they can be easily studied as their Stieltjes transforms are given by the generalized entries of the resolvent ($\langle v_1^{(n)}, (W_n - z)^{-1} v_1^{(n)} \rangle$ and $\langle v_1^{(n)}, (S_n - z)^{-1} v_1^{(n)} \rangle$), for which many results already exist. In particular, as stated in Corollaries 4 and 7, $\mu_{(W_n, v_1^{(n)})}$ and $\mu_{(S_n, v_1^{(n)})}$ converge weakly towards deterministic probability measures denoted $\mu_{sc,A,\theta}$ and $\mu_{\alpha,\Sigma,\theta}$. We are going to present two applications of the spectral measures.

The first one recovers a classical result concerning the value of an outlier and the norm of its associated eigenvector projection in the direction of the spike.

The second one is concerned with the behavior of the projection of non-outlier eigenvectors in the direction of the spike. Namely, in the setting of an additive perturbation, if $f_{sc,A}$ and $f_{sc,A,\theta}$ are the respective densities of $\mu_{sc} \boxplus \mu_A$ and $\mu_{sc,A,\theta}$ and if x is in the support of $\mu_{sc} \boxplus \mu_A$, we prove the following convergence in probability

$$\frac{n}{|\{\lambda \in \mathrm{Spec}(W_n),\ |\lambda - x| \leq \varepsilon_n\}|} \sum_{\lambda \in \mathrm{Spec}(W_n),\ |\lambda - x| \leq \varepsilon_n} \left| \left\langle \phi_\lambda, v_1^{(n)} \right\rangle \right|^2 \xrightarrow[n \to +\infty]{} \frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)}, \tag{4.1}$$

for any sequence $n^{-1/2} \ll \varepsilon_n \ll 1$. In other words, the left-hand side, which is an average in the vicinity of x of the square-projections of eigenvectors in the direction of the spike, converges to a deterministic profile. A similar result holds in the covariance setting. These are the content of

Theorems 3 and 5. Our proof is inspired by the work of Benaych-Georges, Enriquez and Michaïl [BGEM18] and uses local laws estimates recently obtained by Knowles and Yin in [KY14]. When θ belongs to the support of the asymptotic spectrum of A_n (resp. Σ_n), Theorems 3 and 5 are *microscopic* confirmations of the results of Allez and Bouchaud [AB14] (in the Wigner setting) and Ledoit and Péché [LP11] (in the Wishart setting), who derived the asymptotic behavior of the overlaps $|\langle \phi_\lambda, v_\gamma \rangle|^2$ by taking the average over eigenvectors ϕ_λ associated to eigenvalues of W_n (resp. S_n) belonging to a *macroscopic* part of Supp $(\mu_{sc} \boxplus \mu_A)$ (resp. Supp $(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$) and the average over eigenvectors v_γ of A_n (resp. Σ_n) belonging to a *macroscopic* part of the asymptotic spectrum of A_n (resp. Σ_n). When θ does not belong to the support of the asymptotic spectrum of A_n (resp. S_n), such a macroscopic average is not available whereas the spectral measure approach still works.

Interestingly, in rank one perturbation cases, that is when A_n (resp. Σ_n) has only one nonzero (resp. different from one) eigenvalue θ, all the computations are explicit. This is because $\mu_{sc} \boxplus \mu_A = \mu_{sc}$ and $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma = \mu_{\mathrm{MP},\alpha}$ are explicit, respectively equal to the semicircle law and to the Marchenko-Pastur law with parameter α:

$$\mu_{sc}(\mathrm{d}x) = \frac{\sqrt{4-x^2}}{2\pi}\mathbf{1}_{|x|\leq 2}\mathrm{d}x$$
$$\mu_{\mathrm{MP},\alpha}(\mathrm{d}x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x}\mathbf{1}_{(a,b)}(x)\mathrm{d}x + \mathbf{1}_{\alpha<1}(1-\alpha)\delta_0(\mathrm{d}x),$$

where $a, b = (1 \pm \sqrt{\alpha})^2$. In particular, we obtain the following formulas for the spectral measures in the direction of the spike:

$$\mu_{sc,\theta}(\mathrm{d}x) = \frac{\sqrt{4-x^2}}{2\pi(\theta^2+1-\theta x)}\mathbf{1}_{|x|\leq 2}\mathrm{d}x + \mathbf{1}_{|\theta|>1}\left(1-\frac{1}{\theta^2}\right)\delta_{\theta+\frac{1}{\theta}}(\mathrm{d}x)$$
$$\mu_{\mathrm{MP},\alpha,\theta}(\mathrm{d}x) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x\left(x(1-\theta)+\theta(\alpha\theta-\alpha+1)\right)}\mathbf{1}_{(a,b)}(x)\mathrm{d}x + c_{\alpha,\theta}\delta_0(\mathrm{d}x) + d_{\alpha,\theta}\mathbf{1}_{|\theta-1|>\frac{1}{\sqrt{\alpha}}}\delta_{x_{\alpha,\theta}}(\mathrm{d}x),$$

where $c_{\alpha,\theta}$, $d_{\alpha,\theta}$ and $x_{\alpha,\theta}$ are explicit constants, see Propositions 3 and 5. These two measures belong to the class of the so-called *free Meixner* laws which appear in a Gaussian context in the work of Lenczewski [Len15].

Hence, by (4.1), the limiting profiles for the averaged square projections of non-outlier eigenvectors are also explicit. We give numerical simulations that agree with our predictions, see Subsections 4.2.2 and 4.3.2. In the covariance setting, Bloemendal, Knowles, Yau and Yin proved that individual square projections of non-outlier eigenvectors that are associated to eigenvalues in the vicinity of the edge (of b) converge towards a chi-squared random variable with given variance (see [BKYY16, Theorem 2.20]). Although it requires an averaging step, our result completes the picture as it is concerned with eigenvectors associated to any fixed location of the bulk of the spectrum. We believe that the convergence still holds for smaller averaging windows and provide numerical simulations supporting this conjecture at the end of Section 4.5.

Let us finally say a few words about previous use of spectral measures in the literature. Benaych-Georges, Enriquez and Michaïl, in [BGEM18], obtained information on the eigenvectors of a diagonal deterministic matrix perturbed by a random symmetric matrix. In a series of works [GR11, GNR16, GNR17, GNR19], Gamboa, Nagel and Rouault studied spectral measures of some classical ensembles of random matrix theory and their connections with sum rules. More related to our setting, in [BMP07], Bai, Miao and Pan studied the spectral measure at the first vector of the canonical basis e_1 in general covariance cases, in the absence of a spike.

We emphasize that our method could apply to other deformed models such as multiplicative perturbation of Wigner matrices or information plus noise matrices, but we chose to restrict our scope so that the present paper remains short and comprehensible.

Notation and organization of the paper. The random matrices that we study are built from an i.i.d. collection of real random variables $(X_{ij}^{(n)})$, $n \geq 1, 1 \leq i, j \leq n$. Let X be a generic random variable with the same law. We suppose that $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$ and that X has moments of all orders. Notice that the complex case could also be treated, replacing each transposed matrix A^T by its transposed-conjugate A^*, and making the hypothesis that $\mathbb{E}[|X|^2] = 1$.

For a complex number $z \in \mathbf{C}$, we will denote by $\Re(z)$ and $\Im(z)$ the real part and imaginary part of z.

For a probability measure ν, we always denote by

$$s_\nu(z) := \int_{\mathbf{R}} \frac{\mathrm{d}\nu(x)}{x - z}$$

its Stieltjes transform that maps the upper half-plane to itself.

In Subsections 4.2.1 and 4.3.1 we give general results concerning additive perturbation of a Wigner matrix and multiplicative perturbation of a Wishart matrix. Subsections 4.2.2 and 4.3.2 provide explicit computations for rank-one deformation cases. The proofs are done in Sections 4.4 and 4.5.

4.2 Additive perturbation of a Wigner matrix

4.2.1 The general framework

In this section we consider for each $n \geq 1$ the following Wigner matrix:

$$X_n := \begin{bmatrix} X_{11}^{(n)} & X_{12}^{(n)} & \cdots & X_{1n}^{(n)} \\ X_{12}^{(n)} & X_{22}^{(n)} & \cdots & X_{2n}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1n}^{(n)} & X_{2n}^{(n)} & \cdots & X_{nn}^{(n)} \end{bmatrix}.$$

We also consider A_n a deterministic matrix (or random matrix independent of X_n) whose eigenvalues are $\gamma_1^{(n)} = \theta, \gamma_2^{(n)}, \ldots, \gamma_n^{(n)}$, with associated eigenvectors $v_1^{(n)}, \ldots, v_n^{(n)}$. We suppose that there exists a probability measure μ_A such that

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \underset{n \to +\infty}{\longrightarrow} \mu_A \tag{4.2}$$

in the sense of weak convergence. Moreover, let us assume that there exists $\Gamma > 0$ such that, for all $n \geq 1$, $\sup_{1 \leq i \leq n} |\gamma_i^{(n)}| \leq \Gamma$, namely that the eigenvalues of the perturbation remain bounded.

We will study following additive perturbation model:

$$W_n := \frac{1}{\sqrt{n}} X_n + A_n.$$

Let $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ be the eigenvalues of W_n and $\phi_1^{(n)}, \ldots, \phi_n^{(n)}$ the associated normalized eigenvectors. Under assumption (4.2), it is known that the empirical spectral measure μ_{W_n} converges to a deterministic probability measure which is the free convolution $\mu_{sc} \boxplus \mu_A$ between the semicircle law $\mu_{sc}(\mathrm{d}x) = (2\pi)^{-1}\sqrt{4 - x^2}\mathbf{1}_{|x| \leq 2}\mathrm{d}x$ and μ_A. Its Stieltjes transform is characterized by

$$s_{\mu_{sc} \boxplus \mu_A}(z) = \int_{\mathbf{R}} \frac{d\mu_A(\lambda)}{\lambda - s_{\mu_{sc} \boxplus \mu_A}(z) - z}. \tag{4.3}$$

This has first been shown by Pastur in [Pas72]. The study of (4.3) provides information on the probability measure $\mu_{sc} \boxplus \mu_A$. In particular, Biane proved that it has a smooth density with respect to the Lebesgue measure, see [Bia97].

The parameter θ is considered as a spike which may create an outlier in the spectrum, that is an eigenvalue that does not lie in $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$. Following the heuristic that an outlier in the spectrum creates a localized eigenvector, we study the spectral measure in the direction of the spike:

$$\mu_{(W_n, v_1^{(n)})} = \sum_{i=1}^{n} |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \delta_{\lambda_i^{(n)}}.$$

The Stieltjes transform of $\mu_{(W_n, v_1^{(n)})}$ is given by $\langle v_1^{(n)}, (W_n - z)^{-1} v_1^{(n)} \rangle$ which is sometimes called a generalized entry of the resolvent and has already been studied in the literature. The most recent result is the local law recently obtained by Knowles and Yin [16]. It consists in a uniform estimation of $\langle v, (W_n - z)^{-1} w \rangle$ for any vectors v and w and for any complex z in a domain of the upper half plane that is allowed to approach the real axis as n tends to infinity. Since it is one ingredient of the proof of the forthcoming Theorem 3, we provide a precise statement.

Let us first introduce some notation. For all $n \geq 1$, writing $z = E + i\eta$, we define:

$$\begin{cases} G_n(z) = (W_n - z)^{-1}, \\ \Pi_n(z) = \left(A_n - z - s_{\mu_{sc} \boxplus \mu_A}(z)\right)^{-1}, \\ \psi_n(z) = \sqrt{\frac{\Im\left(s_{\mu_{sc} \boxplus \mu_A}(z)\right)}{n\eta}} + \frac{1}{n\eta}. \end{cases} \tag{4.4}$$

For all $x \in \mathbf{R}$, $c > 0$ and $\tau > 0$, we also consider:

$$\mathcal{D}_n^{(\tau)}(x, c) := \left\{ z \in \mathbf{C},\ x - c \leq E \leq x + c,\ n^{-1+\tau} \leq \eta \leq \tau^{-1} \right\}. \tag{4.5}$$

The local law we will use consists in a uniform control between the generalized entries of G_n and Π_n in the spectral domain $\mathcal{D}_n^{(\tau)}(x, c)$ whenever the density of $\mu_{sc} \boxplus \mu_A$ is bounded away from 0 on the interval $[x - c, x + c]$.

Theorem A. [KY17, Theorem 12.2] *Let $x \in \mathbf{R}$ and $c > 0$. Suppose that:*

$$\inf_{t \in [x-c, x+c]} \frac{\mathrm{d}(\mu_{sc} \boxplus \mu_A)(t)}{\mathrm{d}t} > 0. \tag{4.6}$$

Then, for any $\tau > 0$, uniformly in all vectors v, w and uniformly in $z \in \mathcal{D}_n^{(\tau)}(x, c)$, for all $\varepsilon > 0$, there exists $D > 0$ such that

$$\mathbb{P}\left(|\langle v, G_n(z) w \rangle - \langle v, \Pi_n(z) w \rangle| \geq n^{\varepsilon} \psi_n(z) ||v||\, ||w|| \right) \leq \frac{1}{n^D}. \tag{4.7}$$

Theorem A is a direct consequence of the work of Knowles and Yin [KY17]. More precisely, Theorem 12.2 of [KY17] provides a local law of the form (4.7) uniformly in a spectral domain $\mathbf{S}_n \subset \mathbf{C}_+$ provided that an entrywise local law (meaning that v and w are vectors of the canonical basis in (4.7)) has been proved in the particular case where A_n is diagonal, uniformly in the same spectral domain $\mathbf{S}_n$. Such a result has indeed been established by Lee, Schnelli, Stetler and Yau in [LSSY16, Theorem 3.3].

Let us comment on hypothesis (4.6). Together with the assumption that the eigenvalues $\{\gamma_i^{(n)}\}_{1\leq i\leq n}$ remain bounded, it implies that there exists a constant $C>0$ such that, uniformly in $z\in\mathcal{D}_n^{(\tau)}(x,c)$, for all $1\leq i\leq n$,

$$\frac{1}{C}\leq\left|\gamma_i^{(n)}-z-s_{\mu_{sc}\boxplus\mu_A}(z)\right|\leq C. \tag{4.8}$$

Equation (4.8) is sometimes referred to as the *stability assumption*. Going back to the proofs of the local laws, it can be checked that, whenever (4.8) is satisfied on a spectral domain $\mathbf{S}_n$, then (4.7) can be proved on $\mathbf{S}_n$. In particular, since (4.8) can hold without assumption (4.6), the local law (4.7) is usually proved on larger domains than $\mathcal{D}_n^{(\tau)}(x,c)$. We choose to state it on $\mathcal{D}_n^{(\tau)}(x,c)$ because we only need this weaker version in the proof of Theorem 3.

The non-local counterpart of Theorem A is the pointwise convergence of $\langle v,(W_n-z)^{-1}w\rangle$ in the domain $\Im(z)>0$. Taking $v=w=v_1^{(n)}$ yields the following Corollary.

Corollary 4. *The spectral measure $\mu_{(W_n,v_1^{(n)})}$ converges in probability towards a deterministic probability measure $\mu_{sc,A,\theta}$ whose Stieltjes transform is given by*

$$s_{\mu_{sc,A,\theta}}(z)=\frac{1}{\theta-s_{\mu_{sc}\boxplus\mu_A}(z)-z}. \tag{4.9}$$

Remark 2. *Equation (4.9) could allow us to retrieve the limit of $\frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((W_n-z)^{-1}g(A_n)\right)\right]$, obtained in [AB14, Lemme 5.1] by Allez and Bouchaud, for any measurable function g. Indeed, this quantity can be rewritten as*

$$\frac{1}{n}\mathbb{E}\left[\sum_{i,j=1}^{n}\frac{|\langle\phi_i^{(n)},v_j^{(n)}\rangle|^2}{\lambda_i^{(n)}-z}g\left(\gamma_j^{(n)}\right)\right],$$

which is nothing but the average of the Stieltjes transforms of the pushforward of the spectral measures of W_n by g. In particular, it converges to a non-degenerate limit only when the support of g is contained in a macroscopic part of $\mathrm{Supp}(\mu_A)$, due to the renormalization by n. When g is non-null on a microscopic part of $\mathrm{Supp}(\mu_A)$, the study of the spectral measures allows us to obtain the limit of $\mathbb{E}\left[\mathrm{Tr}\left((W_n-z)^{-1}g(A_n)\right)\right]$ whereas $\frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((W_n-z)^{-1}g(A_n)\right)\right]$ brings no information as it converges to zero.

Note that such a macroscopic result can be obtained using simpler arguments than the local law of Theorem A. See for example [Cap13, Proposition 6.2] in the case where v and w are vectors of the canonical basis.

We provide two applications of the asymptotic behavior of the spectral measure of W_n in the direction of $v_1^{(n)}$.

The first one is concerned with outliers and the projection in the direction of the spike of their associated eigenvectors and relies on the following observation: unlike the empirical spectral measure which contains information on outliers only at the order $1/n$, the spectral measure in the direction of the spike already contains it at a *macroscopic* order. For all $x\in\mathbf{R}\setminus\mathrm{Supp}(\mu_{sc}\boxplus\mu_A)$, let us introduce

$$w(x):=x+s_{\mu_{sc,A,\theta}}(x).$$

If there exists x such that $w(x)=\theta$, it is easy to deduce the existence of outliers for W_n as explained in the following Corollary. Although it is already-known in random matrix theory, our approach is new and relatively simple.

Corollary 5. *Suppose that there exists $x_\theta \notin \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ such that $w(x_\theta) = \theta$. Then, x_θ is an outlier of W_n. More precisely, set $\delta > 0$ such that $[x_\theta - \delta, x_\theta + \delta] \cap \mathrm{Supp}(\mu_{sc} \boxplus \mu_A) = \emptyset$ and define k_n to be the number of eigenvalues of W_n inside $[x_\theta - \delta, x_\theta + \delta]$. There exists $1 \leq i_n \leq n$ such that these eigenvalues satisfy*

$$x_\theta + \delta \geq \lambda_{i_n+1}^{(n)} \geq \lambda_{i_n+2}^{(n)} \geq \cdots \geq \lambda_{i_n+k_n}^{(n)} \geq x_\theta - \delta.$$

Then, $k_n \geq 1$ for n sufficiently large and:

1. *Both $\lambda_{i_n+1}^{(n)}$ and $\lambda_{i_n+k_n}^{(n)}$ converge in probability towards x_θ;*
2. *$\sum_{p=1}^{k_n} |\langle \phi_{i_n+p}^{(n)}, v_1^{(n)} \rangle|^2$ converges in probability towards $\frac{1}{w'(x_\theta)}$.*

Proof. Let $x_\theta \notin \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ be such that $w(x_\theta) = \theta$. The value of $\mu_{sc,A,\theta}(\{x_\theta\})$ is given by the residue of $s_{\mu_{sc,A,\theta}}$ at x_θ:

$$(x_\theta - z) s_{\mu_{sc,A,\theta}}(z) = \frac{x_\theta - z}{w(x_\theta) - w(z)} \underset{z \to x^+}{\longrightarrow} \frac{1}{w'(x_\theta)} > 0.$$

Since $\mu_{(W_n, v_1^{(n)})}$ converges to $\mu_{sc,A,\theta}$ by Proposition 4, the Corollary is proved. □

In particular, when W_n is known to be a rank-one perturbation of a matrix W_n' whose empirical spectral measure converges to $\mu_{sc} \boxplus \mu_A$ and which contains no outlier, the interlacing property implies that W_n has a unique outlier, namely that $k_n = 1$ in Corollary 5. Therefore, in that case, the unique outlier converges to x_θ and the square projection in the direction of the spike of its associated eigenvector converges to $1/w'(x_\theta)$.

Before stating our the second result, which represents the main novelty of this paper, and is also an illustration of the use of the spectral measure, we need the following observation:

Proposition 2. *$\mu_{sc,A,\theta}$ is absolutely continuous with respect to the Lebesgue measure on $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$.*

Proof. In [Bia97], Biane proved that $\mu_{sc} \boxplus \mu_A$ is absolutely continuous with respect to the Lebesgue measure. Therefore, the inverse formula

$$\frac{\mathrm{d}\mu_{sc,A,\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{t \to 0^+} \Im(s_{\mu_{sc,A,\theta}}(x + it)) \tag{4.10}$$

and Equation (4.9) imply that $\mu_{sc,A,\theta}$ is also absolutely continuous with respect to the Lebesgue measure at any $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$. □

We will denote by $f_{sc,A}$ and $f_{sc,A,\theta}$ the respective densities of $\mu_{sc} \boxplus \mu_A$ and $\mu_{sc,A,\theta}$ on $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ (these are well-defined quantities by Proposition 2). It turns out that the averaged square-projections of the non-outlier eigenvectors associated to eigenvalues in the vicinity of $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ converge to the ratio of these two densities.

Theorem 3. *Let $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ be such that $f_{sc,A}(x) > 0$. Let ε_n be a sequence that satisfies $n^\delta/\sqrt{n} \ll \varepsilon_n \ll 1$ for some $0 < \delta < 1/2$. Then, for every $t > 0$, if $\mathcal{I}_{\varepsilon_n}^{(n)}(x) = \left\{1 \leq i \leq n : |\lambda_i^{(n)} - x| \leq \varepsilon_n\right\}$:*

$$\mathbb{P}\left(\left| \frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left| \langle \phi_i^{(n)}, v_1^{(n)} \rangle \right|^2 - \frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)} \right| > t \right) \underset{n \to +\infty}{\longrightarrow} 0.$$

By taking g the indicator of an interval contained in $\mathrm{Supp}(\mu_A)$ into the statistic introduced in Remark 2, Allez and Bouchaud [AB14] obtained the asymptotic behavior of the overlaps $|\langle \phi_i^{(n)}, v_j^{(n)} \rangle|^2$ after taking average over eigenvectors $\phi_i^{(n)}$'s (resp. $v_j^{(n)}$'s) with associated eigenvalues $\lambda_i^{(n)}$'s belonging to a *macroscopic* proportion of $\mathrm{Supp}(\mu_{sc} \boxplus \mathrm{Supp}(\mu_A))$ (resp. μ_A). When $\theta \in \mathrm{Supp}(\mu_A)$, Theorem 3 confirms their result at a *microscopic* scale. Indeed, denoting respectively a and b the real and imaginary parts of $\frac{1}{\pi} \lim_{t \to 0^+} s_{\mu_{sc} \boxplus \mu_A}(x + it)$, one can rewrite, using the inverse formula (4.10):

$$\frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)} = \frac{1}{(\theta - x - a)^2 + b^2}.$$

When $\theta \notin \mathrm{Supp}(\mu_A)$, the approach of [AB14] provides no information on the overlap because it only gives access to $n^{-1} s_{\mu_{(W_n, v_1^{(n)})}}(z)$ which converges to zero, whereas the spectral measure approach still works.

4.2.2 The rank-one perturbation

In the special case where $\gamma_2^{(n)} = \cdots = \gamma_n^{(n)} = 0$ for all $n \geq 1$, W_n is a rank-one perturbation of a classical Wigner matrix. The limiting spectrum of the perturbation is $\mu_A = \delta_0$ and almost surely, μ_{W_n} weakly converges to the semicircle distribution. In this setting, we provide explicit computations. Proposition 4 has now the more explicit formulation:

Proposition 3. *In probability, $\mu_{(W_n, v_1^{(n)})}$ converges to:*

$$\mu_{sc,\theta}(\mathrm{d}x) := \frac{\sqrt{4 - x^2}}{2\pi(\theta^2 + 1 - \theta x)} \mathbf{1}_{|x| \leq 2} \mathrm{d}x + \mathbf{1}_{|\theta| > 1} \left(1 - \frac{1}{\theta^2}\right) \delta_{\theta + \frac{1}{\theta}}(\mathrm{d}x).$$

Note that W_n is a rank-one perturbation of $n^{-1/2} X_n$. Therefore, since $\lambda_1(n^{-1/2} X_n) \to 2$ and $\lambda_n(n^{-1/2} X_n) \to -2$ in probability (see [FK81]), if $|\theta| > 1$ then W_n has a single outlier whose location is given by the atom of $\mu_{sc,\theta}$ and whose associated eigenvector has a square projection in the direction of the spike given by the mass of this atom.

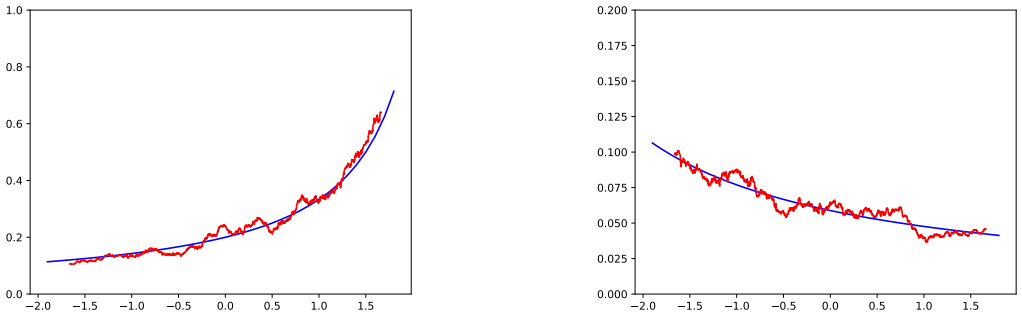
Corollary 6. *The following holds:*

1. *If $\theta > 1$, then, in probability, $\lambda_1(W_n) \underset{n \to +\infty}{\longrightarrow} \theta + \frac{1}{\theta} > 2$ and $|\langle \phi_1^{(n)}, v_1^{(n)} \rangle| \underset{n \to +\infty}{\longrightarrow} \sqrt{1 - \frac{1}{\theta^2}}$.*
2. *If $\theta < -1$, then, in probability, $\lambda_n(W_n) \underset{n \to +\infty}{\longrightarrow} \theta + \frac{1}{\theta} < -2$ and $|\langle \phi_n^{(n)}, v_1^{(n)} \rangle| \underset{n \to +\infty}{\longrightarrow} \sqrt{1 - \frac{1}{\theta^2}}$.*

Finally, the averaged square-projections have also an explicit form, which is just the inverse of a linear function in that case:

Theorem 4. *Let $x \in (-2, 2)$. Let ε_n be a sequence that satisfies $n^\delta / \sqrt{n} \ll \varepsilon_n \ll 1$ for some $\delta > 0$. Then, for every $t > 0$,*

$$\mathbb{P}\left(\left| \frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left| \langle \phi_i^{(n)}, v_1^{(n)} \rangle \right|^2 - \frac{1}{\theta^2 - \theta x + 1} \right| > t \right) \underset{n \to +\infty}{\longrightarrow} 0.$$



(a) (b)

Figure 4.1 – In red: simulations of the average squared-projections around all locations $x \in (-2, 2)$ where we took average over intervals of typical size $n^{0,1}/\sqrt{n}$ for a single matrix $W_n = n^{-1/2}X_n$, where X_n has gaussian entries and is of size 3000×3000. In case (a) $\theta = 2$ and in case (b) $\theta = -4$. In blue: theoretical predictions.

4.3 Multiplicative perturbation of a Wishart matrix

4.3.1 The general framework

Let $m = m(n)$ be a sequence of integers such that $m/n \to \alpha > 0$ as $n \to +\infty$. For all $n \geq 0$, we consider the following random rectangular matrix:

$$X_n := \begin{bmatrix} X_{11}^{(n)} & X_{12}^{(n)} & \cdots & X_{1m}^{(n)} \\ X_{21}^{(n)} & X_{22}^{(n)} & \cdots & X_{2m}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1}^{(n)} & X_{n2}^{(n)} & \cdots & X_{nm}^{(n)} \end{bmatrix}.$$

Let also Σ_n be a general covariance matrix of size $n \times n$, with eigenvalues given by $\gamma_1^{(n)} = \theta, \gamma_2^{(n)}, \ldots, \gamma_n^{(n)}$ and associated eigenvectors $v_1^{(n)}, \ldots, v_n^{(n)}$. We suppose that there exists a probability measure μ_Σ such that

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \underset{n\to+\infty}{\longrightarrow} \mu_\Sigma$$

in the sense of weak convergence. Moreover, let us assume that there exists $\Gamma > 0$ such that for all ≥ 1, $\sup_{1\leq i\leq n} |\gamma_i^{(n)}| \leq \Gamma$. We study the following multiplicative perturbation model:

$$S_n := \frac{1}{n}\Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2}.$$

The matrix S_n can be considered as the sampled covariance matrix of m i.i.d. vectors in $\mathbf{R}^n$ having covariance matrix Σ_n. Let $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ be the eigenvalues of S_n and $\phi_1^{(n)}, \ldots, \phi_n^{(n)}$ the associated normalized eigenvectors.

The empirical spectral distribution of S_n converges to the free product $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ whose Stieltjes transform is characterized by:

$$s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu_\Sigma(t)}{t(\alpha - 1 - z s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(z)) - z}.$$

This is a consequence of the work of Silverstein [Sil95]. A later work of Choï and Silverstein [SC95] proved that $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ is absolutely continuous with respect to the Lebesgue measure at x, for any $x > 0$.

The study of S_n is intimately linked to the study of $\underline{S}_n := \frac{1}{n} X_n^T \Sigma_n X_n$. Indeed, the spectra of S_n and $\underline{S}_n$ only differ by the number of zero eigenvalues. More precisely, denoting μ_{S_n} and $\mu_{\underline{S}_n}$ the respective empirical spectral measures of S_n and $\underline{S}_n$:

$$\mu_{\underline{S}_n} = \frac{1}{n} \sum_{\lambda \in \mathrm{Spec}(S_n)} \delta_\lambda + \frac{m-n}{n} \delta_0.$$

In terms of the Stieltjes transforms, this relation translates into $s_{\mu_{\underline{S}_n}}(z) = s_{\mu_{S_n}}(z) - (\alpha - 1)/z$. Therefore, when n tends to infinity, the empirical spectral measure of $\underline{S}_n$ converges to a deterministic measure whose Stieltjes transform, denoted $\underline{s}$, is given by:

$$\underline{s}(z) = s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z) - \frac{\alpha - 1}{z}. \tag{4.11}$$

Remark 3. (Scaling conventions). *In the literature of random covariance matrices, many authors are using different scaling than this paper. Namely, they study Wishart matrices of the form $B_N = \frac{1}{N} Y_N Y_N^T$ where Y_N is of size $p \times N$ where $p/N \to y > 0$ as N tends to infinity. In this context, the scaling factor is the dimension of the row vectors of Y_N whereas in our case, we scale by the dimension of the column vectors. In order to facilitate the comparison between the two models, let us describe their differences. First, note the following correspondence between the parameters: $n = p$, $m = N$ and $\alpha = 1/y$. Writing $B_N = (p/N)\frac{1}{p} Y_N Y_N^T$, it is easy to see that the empirical spectral measures of the two models satisfy the following equality in law:*

$$\mu_{B_N} = \Lambda_{p/N}\left(\mu_{S_n}\right) = \Lambda_{n/m}\left(\mu_{S_n}\right),$$

where $\Lambda_\xi(\cdot)$ denotes the pushforward by the dilatation $x \mapsto \xi x$. Hence, μ_{B_N} converges to the pushforward of $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ by $\Lambda_{1/\alpha}$. In the particular case where Σ_n is the identity, the limiting measure is given by the pushforward of $\mu_{\mathrm{MP},\alpha}$ which, after computations, is given by

$$\frac{\sqrt{((1+\sqrt{y})^2 - x)(x - (1-\sqrt{y})^2)}}{2\pi y x} \mathbf{1}_{x \in ((1-\sqrt{y})^2, (1+\sqrt{y})^2)} \mathrm{d}x + \left(1 - \frac{1}{y}\right) \delta_0(\mathrm{d}x).$$

We think of θ as a spike, that is an atypical eigenvalue compared to the sequence $\gamma_i^{(n)}$, $2 \leq i \leq n$. We are interested in its influence on the appearance of outliers for S_n. To that purpose, we study the spectral measure in the direction of the spike:

$$\mu_{(S_n, v_1^{(n)})} := \sum_{i=1}^n |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \delta_{\lambda_i^{(n)}}.$$

The Stieltjes transform of $\mu_{(S_n, v_1^{(n)})}$ is given by $\langle v_1^{(n)}, (S_n - z)^{-1} v_1^{(n)} \rangle$ and has already been studied in the literature. As in the Wigner case, the most recent result is the local law obtained by Knowles and Yin [KY17]. It consists in a uniform estimation of $\langle v, (S_n - z)^{-1} w \rangle$ for any vectors v and w and for any complex z in a domain of the upper half plane that is allowed to approach the real axis as n tends to infinity. Since it is an ingredient of the proof of Theorem 5, we provide a precise statement.

Let us first introduce some notation. For all $n \geq 1$, writing $z = E + i\eta$, we define:

$$\begin{cases} G_n(z) = (S_n - z)^{-1}, \\ \Pi_n(z) := -(z(1 + \underline{s}(z)\Sigma_n))^{-1}, \\ \psi_n(z) = \sqrt{\frac{\Im\left(s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z)\right)}{n\eta}} + \frac{1}{n\eta}, \end{cases} \tag{4.12}$$

where we recall that $\underline{s}$ is defined in Equation (4.11). For all $x \in \mathbf{R}$, $c > 0$ and $\tau > 0$, we also consider:

$$\mathcal{D}_n^{(\tau)}(x,c) := \left\{ z \in \mathbf{C},\ x - c \leq E \leq x + c,\ n^{-1+\tau} \leq \eta \leq \tau^{-1} \right\}. \tag{4.13}$$

The local law we will use consists in a uniform control between the generalized entries of G_n and Π_n in the spectral domain $\mathcal{D}_n^{(\tau)}(x,c)$ whenever the density of $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ is bounded away from 0 on the interval $[x-c, x+c]$.

Theorem B. **[KY17, Corollary 3.9]** *Let $x > 0$ and $c > 0$. Suppose that:*

$$\inf_{t \in [x-c, x+c]} \frac{\mathrm{d}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)(t)}{\mathrm{d}t} > 0. \tag{4.14}$$

Then, for any $\tau > 0$, uniformly in all vectors v, w and uniformly for in any $z \in \mathcal{D}_n^{(\tau)}$, for all $\varepsilon > 0$, there exists $D > 0$ such that

$$\mathbb{P}\left(\left| \left\langle v, \Sigma_n^{-1/2} \left(G_n(z) - \Pi_n(z) \right) \Sigma_n^{-1/2} w \right\rangle \right| \geq n^\varepsilon \psi_n(z) ||v||\, ||w|| \right) \leq \frac{1}{n^D}. \tag{4.15}$$

Theorem B is a direct consequence of the work of Knowles and Yin [KY17]. More precisely, they first prove [KY17, Theorem 3.22] that (4.15) holds in the particular case where v and w are vectors of the canonical basis, and when Σ_n is diagonal. Then, using a comparison argument, they extend this result to the general setting (see [KY17, Theorem 3.21]).

It is possible to comment on hypothesis (4.14) as in the Wigner case (see the paragraphs below Theorem A). Together with the assumption that the eigenvalues $\gamma_i^{(n)}$'s remain bounded, it implies that there exists a constant $C > 0$ such that, uniformly in $z \in \mathcal{D}_n^{(\tau)}(x,c)$, for all $1 \leq i \leq n$,

$$\frac{1}{C} \leq \left| 1 + \gamma_i^{(n)} \underline{s}(z) \right| \leq C. \tag{4.16}$$

Equation (4.16) is sometimes referred to as the *stability assumption*. Going back to the proofs of the local laws, it can be checked that, whenever (4.16) is satisfied on a spectral domain $\mathbf{S}_n$, then (4.15) can be proved on $\mathbf{S}_n$. In particular, since (4.16) can hold without assumption (4.14), the local law (4.15) is usually proved on larger domains that $\mathcal{D}_n^{(\tau)}(x,c)$. We choose to state it on $\mathcal{D}_n^{(\tau)}(x,c)$ because we only need this weaker version in the proof of Theorem 5.

The non-local counterpart of Theorem A is the pointwise convergence of $\langle v, (S_n - z)^{-1} w \rangle$ in the domain $\Im(z) > 0$. Taking $v = w = v_1^{(n)}$ yields the following Corollary.

Corollary 7. *In probability, $\mu_{(S_n, v_1^{(n)})}$ converges weakly to a probability measure $\mu_{\alpha,\Sigma,\theta}$ whose Stieltjes transform is given by:*

$$s_{\mu_{\alpha,\Sigma,\theta}}(z) = \frac{1}{\theta(\alpha - 1) - z\theta s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z) - z}. \tag{4.17}$$

Remark 4. *Equation (4.17) could enable us to retrieve the limit of $\frac{1}{n}\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$, obtained in [LP11, Theorem 2] by Ledoit and Péché, for any measurable function g. Indeed, this quantity can be rewritten*

$$\frac{1}{n} \sum_{i,j=1}^{n} \frac{|\langle \phi_i^{(n)}, v_j^{(n)} \rangle|^2}{\lambda_i^{(n)} - z} g\left(\gamma_j^{(n)}\right),$$

which is nothing but the average of the Stieltjes transform of the pushforward of the spectral measures of S_n by g. In particular, it converges to a non-degenerate limit only when the support of g is contained into a macroscopic part of $\mathrm{Supp}(\mu_\Sigma)$, due to the renormalization by n. When g is non-null on a microscopic part of $\mathrm{Supp}(\mu_\Sigma)$, the study of the spectral measures yields the limit of $\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$ whereas $\frac{1}{n}\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$ brings no information as it converges to zero.

Of course, such a macroscopic result can be obtained using more simple arguments than the local law of Theorem B. We refer for example to [Cap13, Proposition 6.2].

In what follows, we provide two applications of the asymptotic behavior of the spectral measure of S_n in the direction of $v_1^{(n)}$.

The first one recovers a classical result on outliers of S_n and the projection in the direction of the spike of their associated eigenvectors. The proof we propose is simple and based on the following observation: unlike the empirical spectral measure which contains information on outliers only at the order $1/n$, the spectral measure in the direction of the spike already contains it at a *macroscopic* order. Before stating our result, let us introduce, for all $x > 0$ such that $x^{-1} \in \mathbf{R} \setminus \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$:

$$F(x) = \alpha x - x - s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(1/x).$$

If there exists $x > 0$ such that $1/F(1/x) = \theta$, we easily obtain the existence of an outlier as explained in the following Corollary.

Corollary 8. *Suppose that there exists $x_{\alpha,\theta} \notin \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$ such that $1/F(1/x_{\alpha,\theta}) = \theta$. Then, $x_{\alpha,\theta}$ is an outlier of S_n. More precisely, set $\delta > 0$ such that $[x_{\alpha,\theta} - \delta, x_{\alpha,\theta} + \delta] \cap \mathrm{Supp}(\mu_{sc} \boxplus \mu_A) = \emptyset$ and define k_n to be the number of eigenvalues of S_n inside $[x_{\alpha,\theta} - \delta, x_{\alpha,\theta} + \delta]$. There exists $1 \leq i_n \leq n$ such that these eigenvalues satisfy*

$$x_{\alpha,\theta} + \delta \geq \lambda_{i_n+1}^{(n)} \geq \lambda_{i_n+2}^{(n)} \geq \cdots \geq \lambda_{i_n+k_n}^{(n)} \geq x_{\alpha,\theta} - \delta.$$

Then, $k_n \geq 1$ for n sufficiently large and:

1. *Both $\lambda_{i_n+1}^{(n)}$ and $\lambda_{i_n+k_n}^{(n)}$ converge in probability towards $x_{\alpha,\theta}$;*
2. *$\sum_{p=1}^{k_n} |\langle \phi_{i_n+p}^{(n)}, v_1^{(n)} \rangle|^2$ converges in probability towards $\frac{x_{\alpha,\theta} F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})}$.*

Proof. Let $x_{\alpha,\theta} > 0$ be such that $x_{\alpha,\theta} \notin \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$ and $1/F(1/x_{\alpha,\theta}) = \theta$. The value of $\mu_{\alpha,\Sigma,\theta}(\{x_{\alpha,\theta}\})$ is given by the residue of $s_{\mu_{\alpha,\Sigma,\theta}}$ at $x_{\alpha,\theta}$:

$$(x_{\alpha,\theta} - z) s_{\mu_{\alpha,\Sigma,\theta}}(z) = \frac{(x_{\alpha,\theta} - z) F(1/x_{\alpha,\theta})}{z(F(1/z) - F(1/x_{\alpha,\theta}))} \underset{z \to x_{\alpha,\theta}^+}{\longrightarrow} \frac{x_{\alpha,\theta} F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})} > 0.$$

Since $\mu_{(S_n, v_1^{(n)})}$ converges to $\mu_{\alpha,\Sigma,\theta}$, the Corollary is easily deduced. □

A particular case is when S_n is a rank-one perturbation of a matrix S_n' which has no outlier and whose empirical spectral measure converges to $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$. In that setting, the interlacing property implies that $k_n = 1$ for all n sufficiently large in Corollary 8, meaning that S_n has only one outlier which converges to $x_{\alpha,\theta}$ and whose associated eigenvector has a square projection in the direction of the spike which converges to $\frac{x_{\alpha,\theta} F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})}$.

Before stating our second result, which is concerned with the projection of non-outlier eigenvectors onto the direction of the spike, we need the following Proposition.

Proposition 4. *$\mu_{\alpha,\Sigma,\theta}$ is absolutely continuous with respect to the Lebesgue measure on $\mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma) \setminus \{0\}$.*

Proof. Let $x \in \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma) \setminus \{0\}$. The work of Choï and Silverstein [SC95] ensures that $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ is absolutely continuous with respect to the Lebesgue measure at x. Then, the inverse formula for the Stieltjes transform:

$$\frac{\mathrm{d}\mu_{sc,A,\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{t \to 0^+} \Im(s_{\mu_{sc,A,\theta}}(x + it)), \tag{4.18}$$

combined with Equation (4.17), implies the Proposition. □

Let $f_{\alpha,\Sigma}$ and $f_{\alpha,\Sigma,\theta}$ be the respective densities of $\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma$ and $\mu_{\alpha,\Sigma,\theta}$ on $\mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$. It turns out that the averaged square projections of the non-outlier eigenvectors associated to eigenvalues in the vicinity of $x \in \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$ converges to the ratio of these two densities.

Theorem 5. *Let $x \in \mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma) \setminus \{0\}$ be such that $f_{\alpha,\Sigma}(x) > 0$. Let ε_n be a sequence that satisfies $n^\delta/\sqrt{n} \ll \varepsilon_n \ll 1$ for some $\delta > 0$. Then, for every $t > 0$, if $\mathcal{I}_{\varepsilon_n}^{(n)}(x) = \left\{1 \leq i \leq n : |\lambda_i^{(n)} - x| \leq \varepsilon_n\right\}$:*

$$\mathbb{P}\left(\left|\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left|\langle \phi_i^{(n)}, v_1^{(n)} \rangle\right|^2 - \frac{f_{\alpha,\Sigma,\theta}(x)}{f_{\alpha,\Sigma}(x)}\right| > t\right) \underset{n \to +\infty}{\longrightarrow} 0.$$

As in the Wigner case, Theorem 5 can be seen as a generalization of a result of [LP11], where the authors obtained the asymptotic behavior of the overlaps $|\langle \phi_i^{(n)}, v_j^{(n)} \rangle|^2$ after taking average over eigenvectors $\phi_i^{(n)}$'s (resp. $v_j^{(n)}$'s) with associated eigenvalues $\lambda_i^{(n)}$'s belonging to a *macroscopic* proportion of $\mathrm{Supp}(\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma)$ (resp. μ_Σ), by taking taking functions of the type $g = \mathbf{1}_{[\gamma,+\infty)}$ in the statistics introduced in Remark 4. Indeed, when $\theta \in \mathrm{Supp}(\mu_\Sigma)$, Theorem 5 is a *microscopic* version of the result of [LP11, Theorem 3]. To obtain their formula it suffices to observe that, if a and b are the real and imaginary parts of $1 - \alpha - z s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z)$, one can rewrite, using Equation (4.18):

$$\frac{f_{\alpha,\Sigma,\theta}(x)}{f_{\alpha,\Sigma}(x)} = \frac{x\theta}{(a\theta - x)^2 + \theta^2 b^2},$$

When $\theta \notin \mathrm{Supp}(\mu_\Sigma)$, the techniques of [LP11] provide no information on the overlaps as it only gives access to $n^{-1} s_{\mu_{(S_n, v_1^{(n)})}}(z)$, which converges to zero, whereas the spectral measure approach still works.

4.3.2 Rank-one perturbation of the Marchenko-Pastur law

In the special case where $\gamma_2^{(n)} = \cdots = \gamma_n^{(n)} = 1$ for all $n \geq 1$, S_n is a rank-one perturbation of a classical Wishart matrix. The limiting spectrum of the perturbation is $\mu_\Sigma = \delta_1$ and almost surely, μ_{S_n} weakly converges to the Marchenko-Pastur law $\mu_{\mathrm{MP},\alpha}$. All previous results have now a more explicit formulation. First, the limit of the spectral measure in the direction of the spike can be identified.

Proposition 5. *In probability, $\mu_{(S_n, v_1^{(n)})}$ converges to:*

$$\mu_{\mathrm{MP},\alpha,\theta}(\mathrm{d}x) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x(x(1-\theta) + \theta(\alpha\theta - \alpha + 1))} \mathbf{1}_{(a,b)}(x)\mathrm{d}x + c_\alpha \mathbf{1}_{\alpha<1} \delta_0(\mathrm{d}x) + d_{\alpha,\theta} \mathbf{1}_{|\theta-1|>\frac{1}{\sqrt{\alpha}}} \delta_{x_{\alpha,\theta}}(\mathrm{d}x),$$

where $c_\alpha = \frac{1-\alpha}{\alpha(\theta-1)+1}$, $d_{\alpha,\theta} = \frac{1-\frac{1}{\alpha(\theta-1)^2}}{1+\frac{1}{\alpha(\theta-1)}}$ and $x_{\alpha,\theta} = \frac{\theta(\alpha\theta-\alpha+1)}{\theta-1}$.

A consequence of [BS98] is that $n^{-1}X_nX_n^T$ has no outlier. Since S_n is a rank one perturbation of this matrix, the discussion following Corollary 8 implies that S_n has a single outlier.

Corollary 9. *The following holds:*

1. *If $\theta > 1 + 1/\sqrt{\alpha}$, then, in probability, $\lambda_1(S_n) \underset{n\to+\infty}{\longrightarrow} x_{\alpha,\theta} > b$ and $|\langle \phi_1^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{d_{\alpha,\theta}}$.*

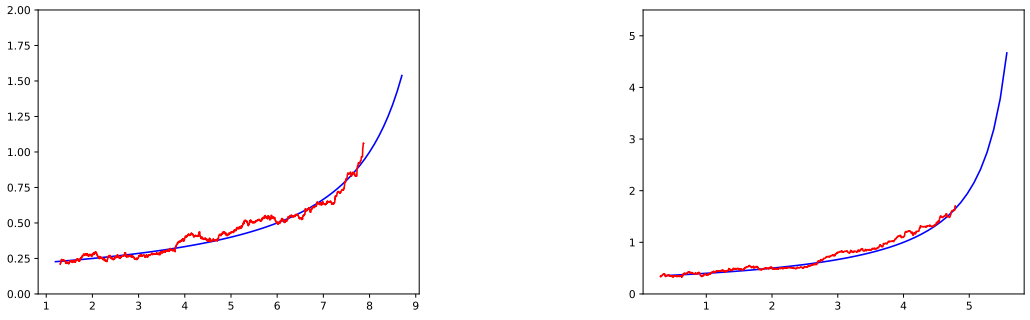
2. *If $\theta < 1 - 1/\sqrt{\alpha}$, then, in probability, $\lambda_n(S_n) \underset{n\to+\infty}{\longrightarrow} x_{\alpha,\theta} < a$ and $|\langle \phi_n^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{d_{\alpha,\theta}}$.*

The ratio of the density of $\mu_{\mathrm{MP},\alpha,\theta}$ and $\mu_{\mathrm{MP},\alpha}$ is explicit and we obtain the following Theorem.

Theorem 6. *Let $x \in (a,b)$. Let ε_n be a sequence that satisfies $n^\delta/\sqrt{n} \ll \varepsilon_n \ll 1$ for some $\delta > 0$. Then, for every $t > 0$,*

$$\mathbb{P}\left(\left|\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)}\left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2 - \frac{\theta}{x(1-\theta)+\theta(\alpha\theta-\alpha+1)}\right| > t\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

When x tends to b, the limiting profile becomes $\theta/(1+\sqrt{\alpha}(1-\theta)^2)$, in accordance with [BKYY16, Theorem 2.20] where the authors obtain a convergence of individual square-projections onto the direction of the spike towards chi-squared random variables with expectation $\theta/(1+\sqrt{\alpha}(1-\theta)^2)$. A natural question would be to study an analogous convergence in law in the bulk of the spectrum (for any $x \in (a,b)$). We do not pursue this issue here.



(a) (b)

Figure 4.2 – In red: simulations of the average squared projections around all locations $x \in (a,b)$ where we took average over interval of typical size $n^{0,1}/\sqrt{n}$, for a single matrix $S_n = n^{-1}\mathrm{Diag}(\sqrt{\theta},1,\dots,1)X_nX_n^T\mathrm{Diag}(\sqrt{\theta},1,\dots,1)$ where X_n is gaussian rectangular of size 2000×8000 ($\alpha = 4$) in case (a) and of size 2000×4000 ($\alpha = 2$) in case (b). In each case $\theta = 2$. In blue: theoretical predictions.

4.4 Identification of the limiting laws in rank-one perturbation cases

In this section we prove Propositions 3 and 5. We use the following branch of the complex square-root:

$$\sqrt{z} = \mathrm{sign}\,(\Im(z))\,\frac{|z|+z}{\sqrt{2(|z|+\Re(z))}}.$$

Proof of Proposition 3. The Stieltjes transform of the semicircle law is given by:

$$s_{\mu_{sc}}(z) = \frac{-z + \sqrt{z^2-4}}{2}.$$

Therefore, using Equation (4.9) and the fact that μ_A is in this case δ_0:

$$\begin{aligned} s_{\mu_{sc,\theta}}(z) &= \frac{2}{2\theta - z - \sqrt{z^2-4}} \\ &= \frac{\sqrt{z^2-4} + 2\theta - z}{2(\theta^2 - \theta z + 1)}. \end{aligned}$$

The absolutely continuous part of $\mu_{sc,\theta}$ is given by

$$\frac{\mathrm{d}\mu_{sc,\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{z \to x^+} \Im(s_{\mu_{sc,\theta}}(z)) = \frac{\sqrt{4-x^2}}{2\pi(\theta^2 + 1 - \theta x)} \mathbf{1}_{|x| \leq 2} \mathrm{d}x.$$

The atom at $\theta + 1/\theta$ is given by the corresponding residue of $s_{\mu_{sc,\theta}}$:

$$-\lim_{z \to (\theta + 1/\theta)^+} (z - \theta - 1/\theta) s_{\mu_{sc,\theta}}(z).$$

By our choice of square-root, $\lim_{z \to \theta + 1/\theta} \sqrt{z^2-4} = |\theta - 1|/\theta$ and one easily deduces:

$$\mu_{sc,\theta}(\{\theta + 1/\theta\}) = \begin{cases} 0 & \text{if } |\theta| \leq 1 \\ 1 - \frac{1}{\theta^2} & \text{if } |\theta| > 1. \end{cases}$$

□

Proof of Proposition 5. Recall the expression of the Stieltjes transform of the Marchenko-Pastur law $\mu_{\mathrm{MP},\alpha} = \mu_{\alpha,1}$:

$$s_{\mu_{\mathrm{MP},\alpha}}(z) = \frac{\alpha - z - 1 + \sqrt{(z-b)(z-a)}}{2z}.$$

Substituting in Equation (4.17), we get

$$\begin{aligned} s_{\mu_{\mathrm{MP},\alpha,\theta}}(z) &= \frac{-2}{2z - 2\theta(\alpha-1) + \theta(\alpha-1) - \theta z + \theta\sqrt{(z-b)(z-a)}} \\ &= \frac{-2\left(2z - \theta(\alpha-1) - \theta z - \theta\sqrt{(z-b)(z-a)}\right)}{\left((2-\theta)z - \theta(\alpha-1)\right)^2 - \theta^2(z-b)(z-a)} \\ &= \frac{\theta\sqrt{(z-b)(z-a)} + z(\theta-2) + \theta(\alpha-1)}{2z\left(z(1-\theta) + \theta(\alpha\theta - \alpha + 1)\right)}. \end{aligned} \tag{4.19}$$

This expression will allow us to obtain an explicit formula for $\mu_{\mathrm{MP},\alpha,\theta}$, through classical inversion results.

The absolutely continuous part of $\mu_{\mathrm{MP},\alpha,\theta}$ is given by:

$$\frac{\mathrm{d}\mu_{\mathrm{MP},\alpha,\theta}}{\mathrm{d}x}(x) = \frac{1}{\pi} \lim_{z \to x^+} \Im(s_{\mu_{\mathrm{MP},\alpha,\theta}}(z)) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x\left(x(1-\theta) + (\alpha\theta - \alpha + 1)\right)} \mathbf{1}_{(a,b)}(x).$$

The atom of $\mu_{\mathrm{MP},\alpha,\theta}$ at zero is given by:

$$-\lim_{\varepsilon \to 0^+} i\varepsilon s_{\mu_{\mathrm{MP},\alpha,\theta}}(i\varepsilon) = -\lim_{\varepsilon \to 0^+} \frac{\sqrt{(i\varepsilon - b)(i\varepsilon - a)} + (\alpha - 1)}{2(\alpha\theta - \alpha + 1)}.$$

By our choice of square-root that preserves the upper-half plane, $\sqrt{(i\varepsilon-b)(i\varepsilon-a)} \to -|ab| = -|\alpha-1|$ as $\varepsilon \to 0^+$. Therefore:

$$\mu_{\mathrm{MP},\alpha,\theta}(\{0\}) = \frac{|\alpha-1|-(\alpha-1)}{2(\alpha\theta-\alpha+1)} = \begin{cases} 0 & \text{if } \alpha \geq 1 \\ \frac{1-\alpha}{\alpha(\theta-1)+1} & \text{if } \alpha < 1. \end{cases}$$

Finally, let us compute the atom at $x_\theta = \frac{\theta(\alpha\theta-\alpha+1)}{\theta-1}$. It is given by:

$$-\lim_{z\to x_\theta^+}(z-x_\theta)s_{\mu_{\mathrm{MP},\alpha,\theta}}(z) = \frac{\sqrt{(x_\theta-b)(x_\theta-a)}+x_\theta(\theta-2)+\theta(\alpha-1)}{2(\theta-1)x_\theta}.$$

We use the following relations which are easily verified.

- $x_\theta - b = \frac{\alpha\left((\theta-1)-\frac{1}{\sqrt{\alpha}}\right)^2}{\theta-1}$,
- $x_\theta - a = \frac{\alpha\left((\theta-1)+\frac{1}{\sqrt{\alpha}}\right)^2}{\theta-1}$.

As for the computation of the atom at zero:

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\sqrt{(x_\theta+i\varepsilon-b)(x_\theta+i\varepsilon-a)} &= \mathrm{sign}(\theta-1)\frac{\alpha}{|\theta-1|}\left|(\theta-1)^2-\frac{1}{\alpha}\right| \\ &= \frac{\alpha}{\theta-1}\left|(\theta-1)^2-\frac{1}{\alpha}\right|.\end{aligned}$$

Besides:

$$\begin{aligned}x_\theta(\theta-2)+\theta(\alpha-1) &= \frac{\theta(\alpha\theta-\alpha+1)(\theta-2)+\theta(\alpha-1)(\theta-1)}{\theta-1} \\ &= \frac{\alpha\theta}{\theta-1}\left((\theta-1)^2-\frac{1}{\alpha}\right).\end{aligned}$$

One obtains:

$$\mu_{\mathrm{MP},\alpha,\theta}(\{x_{\alpha,\theta}\}) = \alpha\theta\frac{\left|(\theta-1)^2-\frac{1}{\alpha}\right|+\left((\theta-1)^2-\frac{1}{\alpha}\right)}{2\theta(\theta-1)(\alpha\theta-\alpha+1)} = \begin{cases} 0 & \text{if } |\theta-1| \leq 1/\sqrt{\alpha} \\ \frac{\alpha\left((\theta-1)^2-\frac{1}{\alpha}\right)}{(\theta-1)(\alpha\theta-\alpha+1)} & \text{if } |\theta-1| > 1/\sqrt{\alpha}. \end{cases}$$

□

4.5 Convergence of the averaged square projections

We only focus on the proof of Theorem 5 concerning the convergence of averaged square-projections into the direction of the spike in the Wishart setting. The proof of Theorem 3, which concerns the Wigner setting, would follow the same reasoning, the only difference being the use of the local law of Theorem A instead of the local law of Theorem B. For the rest of this section, we fix $0 < \delta < 1/2$ and $(\varepsilon_n)_{n\geq 1}$ a sequence of real numbers such that $n^\delta/\sqrt{n} \ll \varepsilon_n \ll 1$. We also fix $x_0 > 0$ such that $\mu_{\alpha,\Sigma,\theta}$ has a positive density at x_0.

Let us explain the heuristic behind Theorem 5. We will denote $I_{\varepsilon_n}(x_0) := [x_0-\varepsilon_n, x_0+\varepsilon_n]$. Recall that $f_{\alpha,\Sigma}$ and $f_{\alpha,\Sigma,\theta}$ are the respective densities of $\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma$ and $\mu_{\alpha,\Sigma,\theta}$ on $\mathrm{Supp}\,(\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma)$. Then,

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n,v_1^{(n)})}(x) \approx \int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{\mathrm{MP},\alpha,\theta}(x) + o_1(1) \approx 2\varepsilon_n f_{\alpha,\Sigma,\theta}(x_0) + o_1(1).$$

On the other hand, if $\mu_{S_n} = n^{-1}\sum_{1\le i\le n}\delta_{\lambda_i^{(n)}}$ denotes the empirical spectral measure of S_n:

$$\begin{aligned}\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n,v_1^{(n)})}(x) \\ &= \sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} |\langle \phi_i, e_1\rangle|^2 = \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} \left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2\right) \times \int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(x) \\ &\approx \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} |\langle \phi_i, e_1\rangle|^2\right) \times \left(2\varepsilon_n f_{\alpha,\Sigma}(x_0) + o_2(1)\right),\end{aligned}$$

where we recall that $\mathcal{I}_{\varepsilon_n}^{(n)}(x_0) = \{1\le i\le n : |\lambda_i^{(n)} - x_0| \le \varepsilon_n\}$. Theorem 5 would be proved if the errors $o_1(1)$ and $o_2(1)$ were explicit and of a smaller order than ε_n. The understanding of these errors is precisely the purpose of the so-called local laws that have been recently developed in random matrix theory. In the Wishart setting, it is given by the local law of Knowles and Yin stated in Theorem B.

The rest of this section makes the above heuristic rigorous. It combines a local law on the Stieltjes transform of $\mu_{(S_n,v_1^{(n)})}$ together with an approximation argument which allows us to estimate a term of the form $\mu_{(S_n,v_1^{(n)})}(I_{\varepsilon_n}(x_0))$. The idea of the latter is to bound the indicator of $I_{\varepsilon_n}(x_0)$ by two smooth analytic functions which we now introduce. Let $(\omega_n)_{n\ge1}$ be a sequence of real numbers such that

$$n^{\delta}/\sqrt{n} \ll \omega_n \ll \varepsilon_n. \tag{4.20}$$

Let Ψ be a smooth decreasing function such that $\Psi(x)\equiv 1$ on $(-\infty,0]$ and $\Psi(x)\equiv 0$ on $[1,+\infty)$. For all $n\ge 1$, we define:

$$\begin{cases}\phi_n^-(x) = \Psi\left(1+\frac{x-x_0-\varepsilon_n+\omega_n}{\omega_n}\right)\Psi\left(1-\frac{x-x_0+\varepsilon_n-\omega_n}{\omega_n}\right), \\ \phi_n^+(x) = \Psi\left(\frac{x-x_0-\varepsilon_n-\omega_n}{\omega_n}\right)\Psi\left(-\frac{x-x_0+\varepsilon_n+\omega_n}{\omega_n}\right).\end{cases} \tag{4.21}$$

With these definitions, it is easy to check the following properties.

1. the support of ϕ_n^- (resp. ϕ_n^+) is included in $[x_0-\varepsilon_n+\omega_n, x_0+\varepsilon-\omega_n]$ (resp. $[x_0-\varepsilon_n-2\omega_n, x_0+\varepsilon_n+2\omega_n]$);
2. ϕ_n^- (resp. ϕ_n^+) is constant equal to 1 on $[x_0-\varepsilon_n+2\omega_n, x_0+\varepsilon_n-2\omega_n]$ (resp. $[x_0-\varepsilon_n-\omega_n, x_0+\varepsilon_n+\omega_n]$);
3. the supports of $(\phi_n^-)'$ and $(\phi_n^-)''$ (resp. $(\phi_n^+)'$ and $(\phi_n^+)''$) are included in $[x_0-\varepsilon_n+\omega_n, x_0-\varepsilon_n+2\omega_n]\cup[x_0+\varepsilon_n-2\omega_n, x_0+\varepsilon_n-\omega_n]$ (resp. $[x_0-\varepsilon_n-2\omega_n, x_0-\varepsilon_n-\omega_n]\cup[x_0+\varepsilon_n+\omega_n, x_0+\varepsilon_n+2\omega_n]$);
4. $||(\phi_n^-)'||_\infty = ||(\phi_n^+)'||_\infty = O(1/\omega_n)$ and $||(\phi_n^-)''||_\infty = ||(\phi_n^+)''||_\infty = O(1/\omega_n^2)$.

By construction,

$$\begin{cases}\int_{\mathbf{R}}\phi_n^-(\lambda)\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda) \le \int_{I_{\varepsilon_n}(x_0)}\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda) \le \int_{\mathbf{R}}\phi_n^+(\lambda)\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda), \\ \int_{\mathbf{R}}\phi_n^-(\lambda)\mathrm{d}\mu_{S_n}(\lambda) \le \int_{I_{\varepsilon_n}(x_0)}\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda) \le \int_{\mathbf{R}}\phi_n^+(\lambda)\mathrm{d}\mu_{S_n}(\lambda).\end{cases} \tag{4.22}$$

The main result of this section is an estimate on both sides of the above inequalities.

Lemma 1. *Let $\epsilon \in \{-,+\}$. There exists $D > 0$ such that, with probability at least $1 - n^{-D}$,*

$$\left| \int_{\mathbf{R}} \phi_n^{\epsilon}(\lambda) \mathrm{d}\left(\mu_{(S_n, v_1^{(n)})} - \mu_{\alpha,\Sigma,\theta} \right)(\lambda) \right| = O(\omega_n) \tag{4.23}$$

and

$$\left| \int_{\mathbf{R}} \phi_n^{\epsilon}(\lambda) \mathrm{d}\left(\mu_{S_n} - \mu_{\mathrm{MP},\alpha} \right)(\lambda) \right| = O(\omega_n). \tag{4.24}$$

Before giving the proof of Lemma 1, let us explain how it leads to Theorem 5.

Proof of Theorem 5. Combining estimates (4.23) and (4.24) with Equation 4.22, for any $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$:

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n, v_1^{(n)})}(x) = 2\varepsilon_n f_{\alpha,\Sigma,\theta}(x_0) + O(\omega_n) \tag{4.25}$$

and

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(x) = 2\varepsilon_n f_{\alpha,\Sigma}(x_0) + O(\omega_n). \tag{4.26}$$

Therefore, since $\varepsilon_n \gg n^{-1/2}$, for all $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$,

$$\begin{aligned} \frac{n}{|\mathcal{I}_{\varepsilon_n}(x_0)|} \sum_{i \in \mathcal{I}_{\varepsilon_n}(x_0)} |\langle \phi_i, e_1 \rangle|^2 &= \frac{\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n, v_1^{(n)})}(\lambda)}{\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(\lambda)} \\ &= \frac{f_{\alpha,\Sigma,\theta}(x_0)}{f_{\alpha,\Sigma}(x_0)} + O\left(\frac{\omega_n}{\varepsilon_n} \right). \end{aligned}$$

This ends the proof of Theorem 5 since $\varepsilon_n \gg \omega_n$. □

We now turn to the proof of Lemma 1. We will use a local law on the Stieltjes transform of $\mu_{(S_n, v_1^{(n)})}$ which can be obtained by taking $v = w = v_1^{(n)}$ in Equation (4.15). Recall the following definitions:

$$\mathcal{D}_n^{(\tau)}(x, c) = \left\{ z \in \mathbf{C},\ x - c \leq E \leq x + c,\ n^{-1+\tau} \leq \eta \leq \tau^{-1} \right\}$$

and

$$\psi_n(z) = \sqrt{\frac{\Im\left(s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}(z) \right)}{n\eta}} + \frac{1}{n\eta}.$$

Since x_0 is such that $f_{\alpha,\Sigma}(x_0) > 0$, there exists $c > 0$ small enough such that Equation (4.14) is satisfied. We fix such a c for the rest of this section. Then, Theorem B translates into the following result.

Theorem C. **[KY17, Corollary 3.9]** *For any $\tau > 0$, uniformly in $z \in \mathcal{D}_n^{(\tau)}(x_0, c)$, for any $\varepsilon > 0$, there exists $D > 0$ such that*

$$\mathbb{P}\left(\left| s_{(S_n, v_1^{(n)})}(z) - s_{\mu_{\alpha,\Sigma,\theta}}(z) \right| \geq n^{\varepsilon} \psi(z) \right) \leq \frac{1}{n^D}. \tag{4.27}$$

In the proof of Lemma 1, we will use the following classical estimate on the error function.

Lemma 2. *For all $x \in \mathbf{R}$ such that $f_{\alpha,\Sigma}(x) > 0$,*

$$|\psi_n(x+iy)| = O\left(\sqrt{\frac{\Im\left(s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(x+iy)\right)}{ny}}\right). \tag{4.28}$$

In particular, for all $1 \leq |y| \leq 2$,

$$|\psi_n(x+iy)| = O\left(n^{-1/2}\right). \tag{4.29}$$

Proof. Let $x \in \mathbf{R}$ be such that $f_{\alpha,\Sigma}(x) > 0$. Then, since

$$f_{\alpha,\Sigma}(x) = \frac{1}{\pi}\lim_{t\to 0^+}\Im\left(s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(x+it)\right),$$

the leading term in the expression of $\psi_n(x+iy)$ is $\sqrt{\frac{\Im\left(s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(x+iy)\right)}{ny}}$, which proves (4.28). This latter term is of order $n^{-1/2}$ whenever y is bounded away from 0, which implies (4.29). □

We are now ready to prove Lemma 1. The strategy is based on the Helffer-Sjöstrand formula, which allows us to translate an estimate of the form (4.27) into an estimate on sufficiently regular functions integrated against $\mu_{(S_n,v_1^{(n)})}$. Although the argument is standard and can be found for the empirical spectral measure of a Wigner matrix in the survey of Benaych-Georges and Knowles [BGK17], we choose to provide the details as it has not been done for the spectral measures. A similar argument is also present in the work of Benaych-Georges, Enriquez and Michaïl [BGEM18].

Proof of Lemma 1. We first prove estimate (4.23) that corresponds to the spectral measure and then explain how to adapt the proof for the second inequality (4.24) which corresponds to the empirical spectral measure. We only focus on the case $\epsilon = +$ because the case $\epsilon = -$ follows from the same argument. In order to lighten notation, we write $\phi_n = \phi_n^+$.

For all $x \in \mathbf{R}$, by the the Helffer-Sjöstrand formula (see [BGK17, Proposition C.1]):

$$\phi_n(x) = \int_{\mathbf{C}} \frac{\overline{\partial}\left(\tilde{\phi}_n(z)\chi(z)\right)}{x-z}\mathrm{d}z, \tag{4.30}$$

where:

- χ is a smooth symmetric cutoff function that equals 1 on $[-1,1]$ and 0 outside $[-2,2]$;
- $\tilde{\phi}_n$ is the quasi-analytic extension of degree 1 of ϕ_n, defined by $\tilde{\phi}_n(x+iy) = \phi_n(x) + iy\phi_n'(x)$;
- $\overline{\partial} = \frac{1}{2}\left(\partial_n + i\partial_y\right)$.

Let us define

$$\hat{\mu}_n := \mu_{(S_n,v_1^{(n)})} - \mu_{\alpha,\Sigma,\theta}$$

and its Stieltjes transform $\hat{s}_n := s_{\mu_{(S_n,v_1^{(n)})}} - s_{\mu_{\alpha,\Sigma,\theta}}$. Equation (4.30) leads to:

$$\begin{aligned}\int_{\mathbf{R}} \phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda) = \quad & \frac{i}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}} \phi_n''(x)y\chi(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y \\ & + \frac{i}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}} \left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y.\end{aligned}$$

Note that the right-hand-side is real so that

$$\int_{\mathbf{R}} \phi_n(\lambda) \mathrm{d}\hat{\mu}_n(\lambda) \leq \quad \frac{-1}{2\pi} \int_{x\in\mathbf{R}} \int_{|y|\leq\omega_n} \phi_n''(x) y\chi(y) \Im\left(\hat{s}_n(x+iy)\right) \mathrm{d}x\mathrm{d}y \tag{4.31}$$

$$+ \frac{-1}{2\pi} \int_{x\in\mathbf{R}} \int_{|y|\geq\omega_n} \phi_n''(x) y\chi(y) \Im\left(\hat{s}_n(x+iy)\right) \mathrm{d}x\mathrm{d}y \tag{4.32}$$

$$+ \left| \frac{1}{2\pi} \int_{x\in\mathbf{R}} \int_{y\in\mathbf{R}} \left(\phi_n(x) + iy\phi_n'(x)\right) \chi'(y) \hat{s}_n(x+iy) \mathrm{d}x\mathrm{d}y \right|. \tag{4.33}$$

We now estimate all of the three terms of the right-hand side. In what follows, we fix $\tau > 0$ and $\varepsilon > 0$ and we argue on the event

$$E_\varepsilon := \left\{ \forall z \in \mathcal{D}_n^{(\tau)}(x_0, c),\ \left| s_{(S_n, v_1^{(n)})}(z) - s_{\mu_{\alpha,\Sigma,\theta}}(z) \right| \leq n^\varepsilon \psi_n(z) \right\}. \tag{4.34}$$

Note that by the local law (4.27), there exists $D > 0$ such that $\mathbb{P}(E) \geq 1 - n^{-D}$.

Estimation of the term (4.31). Recall that, from the definition of ϕ_n, given in Equation (4.21), the support of ϕ_n'' is contained in

$$I_n := [x_0 - \varepsilon_n - 2\omega_n, x_0 - \varepsilon_n - \omega_n] \cup [x_0 + \varepsilon_n + \omega_n, x_0 + \varepsilon_n + 2\omega_n]. \tag{4.35}$$

Moreover, since $\omega_n \ll 1$, $\chi(y) = 1$ when $|y| \leq \omega_n$. Therefore:

$$\begin{aligned} &\left| \frac{-1}{2\pi} \int_{x\in\mathbf{R}} \int_{|y|\leq\omega_n} \phi_n''(x) y\chi(y) \Im\left(\hat{s}_n(x+iy)\right) \mathrm{d}x\mathrm{d}y \right| \\ &\leq \frac{1}{2\pi} \int_{x\in I_n} \int_{0\leq y\leq\omega_n} \left| \phi_n''(x) y \Im\left(\hat{s}_n(x+iy)\right) \right| \mathrm{d}x\mathrm{d}y + \frac{1}{2\pi} \int_{x\in I_n} \int_{-\omega_n\leq y<0} \left| \phi_n''(x) y \Im\left(\hat{s}_n(x+iy)\right) \right| \mathrm{d}x\mathrm{d}y \end{aligned} \tag{4.36}$$

We only treat in details the first term of (4.36) as the second one can be analyzed similarly.

Let $x \in I_n$. The function $y \mapsto y\Im\left(s_{\mu_{(S_n, v_1^{(n)})}}(x+iy)\right)$ is non-decreasing, which implies that:

$$\forall 0 \leq y \leq \omega_n, \quad y\Im\left(\hat{s}_n(x+iy)\right) \leq y\Im\left(s_{\mu_{(S_n, v_1^{(n)})}}(x+iy)\right) \leq \omega_n \Im\left(s_{\mu_{(S_n, v_1^{(n)})}}(x+i\omega_n)\right). \tag{4.37}$$

By assumption, the point $x_0 \in \mathbf{R}$ is such that $f_{\alpha,\Sigma,\theta}(x_0) > 0$. Since $\lim_{t\to 0^+} \Im\left(s_{\mu_{\alpha,\Sigma,\theta}}(x_0 + it)\right) = \pi f_{\alpha,\Sigma,\theta}(x_0)$, this implies that there exists a constant $\tilde{C} > 0$ such that, for large enough n:

$$\forall x \in I_n,\ \forall 0 \leq y \leq \omega_n, \quad \Im\left(s_{\mu_{\alpha,\Sigma,\theta}}(x+iy)\right) \leq \tilde{C}.$$

Therefore, since we are working on the event E_ε introduced in (4.34), uniformly in $x \in I_n$:

$$\Im\left(s_{\mu_{(S_n, v_1^{(n)})}}(x+i\omega_n)\right) \leq n^\varepsilon \psi_n(x+i\omega_n) + \tilde{C}. \tag{4.38}$$

By Equation (4.28) of Lemma 2, $n^\varepsilon \psi_n(x+i\omega_n)$ converges to zero as n tends to infinity. Therefore, by combining Inequalities (4.37) and (4.38), we deduce the existence of some constant $C > 0$ such that:

$$\int_{x\in I_n} \int_{0\leq y\leq\omega_n} \left| \phi_n''(x) y \Im\left(\hat{s}_n(x+iy)\right) \right| \mathrm{d}x\mathrm{d}y \leq C\omega_n \int_{x\in I_n} \int_{0\leq y\leq\omega_n} |\phi_n''(x)| \mathrm{d}x\mathrm{d}y. \tag{4.39}$$

Finally, using that $||\phi_n''||_\infty = O(1/\omega_n^2)$ and $\int_{I_n} \mathrm{d}x = 2\omega_n$ in (4.39) yields:

$$\int_{x\in I_n}\int_{0\le y\le\omega_n} \left|\phi_n''(x)y\Im\left(\hat{s}_n(x+iy)\right)\right| \mathrm{d}x\mathrm{d}y \le O\left(\omega_n\right). \tag{4.40}$$

As already mentioned, the same argument implies that the bound (4.40) also holds if the domain of integration was $I_n \times [-\omega_n, 0)$. Hence, we proved that, on the event E_ε,

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\le\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right| = O(\omega_n). \tag{4.41}$$

Estimation of the term (4.32). We first decompose the term according to the sign of y.

$$\begin{aligned}
&\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\ge\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y \\
&= \frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\le-\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y + \frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\ge\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{4.42}$$

The two terms on the right-hand side of (4.42) can be analyzed in the same way, so that we only focus on the case where $y \ge \omega_n$.

Differentiating with respect to x and y and using that $\partial_x\Im(\hat{s}_n(x+iy)) = -\partial_y\Re(\hat{s}_n(x+iy))$, we obtain:

$$\begin{aligned}
&\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\ge\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y \\
&\quad= \frac{-1}{2\pi}\int_{y\ge\omega_n} y\chi(y)\mathrm{d}y \times (-1)\int_{x\in\mathbf{R}} \phi_n'(x)\partial_x\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x \\
&\quad= \frac{-1}{2\pi}\int_{x\in\mathbf{R}} \phi_n'(x)\mathrm{d}x\left(\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right) - \int_{y>\omega_n}(y\chi'(y)+\chi(y))\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\right) \\
&\quad= \frac{-1}{2\pi}\int_{x\in\mathbf{R}} \phi_n'(x)\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right)\mathrm{d}x && (4.43) \\
&\qquad\qquad + \frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y>\omega_n} \phi_n'(x)y\chi'(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\mathrm{d}x && (4.44) \\
&\qquad\qquad + \frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y>\omega_n} \phi_n'(x)\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\mathrm{d}x. && (4.45)
\end{aligned}$$

The first term (4.43) can be bounded as follows, recalling that ϕ_n' is supported on I_n (defined in Equation (4.35)):

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}} \phi_n'(x)\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right)\mathrm{d}x\right| \le \frac{1}{2\pi}||\phi_n'||_\infty\omega_n\int_{I_n}\left|\Im\left(\hat{s}_n(x+i\omega_n)\right)\right|\mathrm{d}x.$$

By the definition of ϕ_n, it holds that $||\phi_n'||_\infty = O(1/\omega_n)$. Combining this with inequality (4.38), which holds uniformly in $x \in I_n$, we get that:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}} \phi_n'(x)\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right)\mathrm{d}x\right| = O(\omega_n). \tag{4.46}$$

The second term (4.44) can be bounded as follows, using that χ' is supported on $[-2,-1]\cup[1,2]$:

$$\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\ge\omega_n} \phi_n'(x)y\chi'(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\mathrm{d}x\right| \le \frac{1}{2\pi}||\phi_n'||_\infty\int_{x\in I_n}\int_{1\le y\le 2} y\chi'(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y.$$

Since we are working on the event E_ε, $|\Im(\hat{s}_n(x+iy))| \leq n^\varepsilon \psi_n(x+iy)$. Therefore, using Lemma 2:

$$\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\geq\omega_n}\phi_n'(x)y\chi'(y)\Im(\hat{s}_n(x+iy))\,\mathrm{d}y\mathrm{d}x\right| = O\left(\frac{n^\varepsilon}{\sqrt{n}}\right). \tag{4.47}$$

Finally, we bound the last term (4.45). Since ϕ_n' is supported on I_n, we can replace the integral over $\mathbf{R}$ by an integral over I_n. Moreover, using (4.28), we get that:

$$\begin{aligned}\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y>\omega_n}\phi_n'(x)\chi(y)\Im(\hat{s}_n(x+iy))\,\mathrm{d}y\mathrm{d}x\right| &\leq \frac{1}{2\pi}||\phi_n'||_\infty\int_{I_n}\int_{\omega_n}^{2}\frac{n^\varepsilon}{\sqrt{ny}}\mathrm{d}y\mathrm{d}x\\ &= O\left(\frac{n^\varepsilon}{\sqrt{n}}\right).\end{aligned} \tag{4.48}$$

Hence, putting together (4.46), (4.47) and (4.48), we proved that, on the event E_ε:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\geq\omega_n}\phi_n''(x)y\chi(y)\Im(\hat{s}_n(x+iy))\,\mathrm{d}x\mathrm{d}y\right| = O\left(\max\left\{\frac{n^\varepsilon}{n^{1/2}},\omega_n\right\}\right) \tag{4.49}$$

Estimation of the term (4.33). Since χ' is symmetric and supported on $[-2,-1]\cup[1,2]$, there exists some constant $C>0$ such that:

$$\begin{aligned}&\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}}\left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y\right|\\ &\qquad\leq\frac{C}{2\pi}\int_{x\in[x_0-\varepsilon_n-2\omega_n,x_0+\varepsilon_n+2\omega_n]}\int_{y\in[-2,-1]\cup[1,2]}\left(1+||\phi_n'||_\infty\right)|\hat{s}_n(x+iy)|\,\mathrm{d}x\mathrm{d}y.\end{aligned}$$

By Lemma 2, uniformly in $x\in[x_0-\varepsilon_n-2\omega_n,x_0+\varepsilon_n+2\omega_n]$ and $y\in[-2,-1]\cup[1,2]$, $|\hat{s}_n(x+iy)|\leq\frac{n^\varepsilon}{\sqrt{n}}$. Therefore:

$$\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}}\left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y\right| = O\left(\frac{n^\varepsilon}{\sqrt{n}}\right). \tag{4.50}$$

Conclusion Putting together estimates (4.41), (4.49) and (4.50), we proved that, for all $\varepsilon>0$, on the event E_ε:

$$\left|\int_{\mathbf{R}}\phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda)\right| = O\left(\max\left\{\frac{n^\varepsilon}{n^{1/2}},\omega_n\right\}\right) = O(\omega_n). \tag{4.51}$$

Now, recall from (4.20) that $\omega_n \gg n^\delta/\sqrt{n}$. This implies that for all $0<\varepsilon\leq\delta$: $\max\{n^\varepsilon/n^{1/2},\omega_n\} = \omega_n$. Hence, there exists $D>0$ such that with probability at least $1-n^{-D}$:

$$\left|\int_{\mathbf{R}}\phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda)\right| = O(\omega_n).$$

This ends the proof of the first part of Lemma 1.

It remains to explain how to adapt the above argument to obtain the second estimate (4.24). Since it is concerned with the *empirical* spectral measure μ_{S_n}, we need a local law for this quantity, which is an analog of Theorem C in this context, and is a consequence of the work of Knowles and Yin. Recall that $\mathcal{D}_n^{(\tau)}(x_0,c) = \{z\in\mathbf{C},\ x_0-c\leq E\leq x_0+c,\ n^{-1+\tau}\leq\eta\leq\tau^{-1}\}$.

Theorem D. [KY17, Theorem 3.22] *Let $s_{\mu_{S_n}}$ be the Stieltjes transform of μ_{S_n}. Let $\tau>0$. Then, uniformly in $z\in\mathcal{D}_n^{(\tau)}(x_0,c)$, for any $\varepsilon>0$, there exists $D>0$ such that:*

$$\mathbb{P}\left(\left|s_{\mu_{S_n}}(z)-s_{\mu_{\mathrm{MP},\alpha}\boxtimes\mu_\Sigma}(z)\right|\geq n^\varepsilon(n\Im(z))^{-1}\right)\leq\frac{1}{n^D}.$$

As before, we only discuss the case where $\epsilon = +$ and denote $\phi_n = \phi_n^+$. As in the case of the spectral measure, denoting $\tilde{\mu}_n := \mu_{(S_n, v_1^{(n)})} - \mu_{\alpha,\Sigma,\theta}$ and $\tilde{s}_n := s_{\mu_{S_n}} - s_{\mu_{\mathrm{MP},\alpha} \boxtimes \mu_\Sigma}$, the Helffer-Sjöstrand formula leads to:

$$\int_{\mathbf{R}} \phi_n(\lambda) \mathrm{d}\tilde{\mu}_n(\lambda) \leq \frac{-1}{2\pi} \int_{x\in\mathbf{R}} \int_{|y|\leq\omega_n} \phi_n''(x) y \chi(y) \Im\left(\tilde{s}_n(x+iy)\right) \mathrm{d}x\mathrm{d}y \tag{4.52}$$

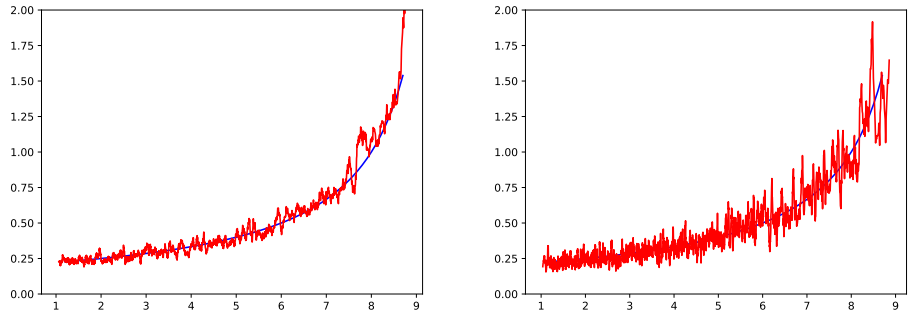
$$+ \frac{-1}{2\pi} \int_{x\in\mathbf{R}} \int_{|y|\geq\omega_n} \phi_n''(x) y \chi(y) \Im\left(\tilde{s}_n(x+iy)\right) \mathrm{d}x\mathrm{d}y \tag{4.53}$$

$$+ \left| \frac{1}{2\pi} \int_{x\in\mathbf{R}} \int_{y\in\mathbf{R}} \left(\phi_n(x) + iy\phi_n'(x)\right) \chi'(y) \tilde{s}_n(x+iy) \mathrm{d}x\mathrm{d}y \right|. \tag{4.54}$$

Now, each term (4.52), (4.53) and (4.54) can be treated in the same way as the previous terms (4.31), (4.32) and (4.33). The only difference is that each occurence of the former error term $\psi_n(z)$ is now replaced by the error term of Theorem C, namely $(n\Im(z))^{-1}$. □

We underline that a weaker version of Theorem 5 (resp. Theorem 3) can be obtained as long as a uniform estimation of $s_{(S_n, v_1^{(n)})}(z) - s_{\mu_{\alpha,\Sigma,\theta}}(z)$ is available for z in a domain of the upper half-plane that is allowed to approach the real axis as n becomes larger. Indeed, if $|s_{(S_n, v_1^{(n)})}(z) - s_{\mu_{\alpha,\Sigma,\theta}}(z)| = O(\varepsilon_n)$, then, the Helffer-Sjöstrand argument that we developed during the proof of Theorem 5 yields a convergence of the averaged-square projection onto the direction of the spike for averaging windows of size ε_n, as long as $\varepsilon_n \gg n^\delta/\sqrt{n}$ for some $\delta > 0$. This limitation corresponds to the optimal rate in the local laws of Knowles and Yin (4.27).

One natural question would be to weaken the assumption on the size ε_n of the averaging window: do our Theorems 4, 6, 3 and 5 hold as soon as $\varepsilon_n = o(1)$? We believe that the answer is positive (see Figure 4.3 for simulations). Morevoever, the results of [BKYY16], which state the convergence in law of properly rescaled *individual* square projection of eigenvectors associated to eigenvalues in the vicinity of the edge, suggest the following natural question: does such a convergence also hold in the bulk of the spectrum?



(a) (b)

Figure 4.3 – In red: simulations of the average squared projections around all locations $x \in ((1-\sqrt{\alpha})^2, (1+\sqrt{\alpha})^2)$ where we took average over interval of typical size $n^{0,3}$ for 4.3a and $n^{0,2}$ for 4.3b, for 10 independent matrices of the form $S_n = n^{-1}\mathrm{Diag}(\sqrt{\theta}, 1, \ldots, 1) X_n X_n^T \mathrm{Diag}(\sqrt{\theta}, 1, \ldots, 1)$ where X_n is Gaussian rectangular of size 2000×8000 ($\alpha = 4$) in case 4.2a and of size 3000×12000 ($\alpha = 4$) in case 4.2b. In each case $\theta = 2$. In blue: theoretical predictions.

Chapter 5

A solvable class of renewal processes

This chapter corresponds to the publication [EN20].

When the distribution of the inter-arrival times of a renewal process is a mixture of geometric laws, we prove that the renewal function of the process is given by the moments of a probability measure which is explicitly related to the mixture distribution. We also present an analogous result in the continuous case when the inter-arrival law is a mixture of exponential laws. We then observe that the above discrete class of renewal processes provides a solvable family of random polymers. Namely, we obtain an exact representation of the partition function of polymers pinned at sites of the aforementioned renewal processes. In the particular case where the mixture measure is a generalized Arcsine law, the computations can be explicitly handled.

5.1 Renewal theory for mixtures of geometric laws

Let K be a probability measure on $\mathbf{N} \cup \{\infty\} = \{1, \dots, \} \cup \{\infty\}$. Let $\eta = (\eta_n)_{n \geq 1}$ be an i.i.d. sequence of random variables with law K. We consider η as inter-arrival times of a renewal process and we define the random variables $\tau_0 = 0$ and for all $n \geq 1$, $\tau_n = \sum_{1 \leq i \leq n} \eta_i$. Finally, we define the renewal process associated to K as the random set $\tau := \{\tau_i;\, i \geq 0\}$. We will denote by $\mathbf{P}$ the law of τ.

Before stating our main result, let us introduce the Stieltjes transform of a positive measure, which turns out to be the key notion in this framework. Let $\mathbf{C}_+ := \{z \in \mathbf{C},\, \Im z > 0\}$. Let μ be a positive measure on $\mathbf{R}$. The Stieltjes transform of μ is the analytic function s_μ, defined from $\mathbf{C}_+$ to $\mathbf{C}_+$ by:

$$s_\mu(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu(x)}{x - z}.$$

Let us mention the useful fact that, for almost every real point $x \in \mathbf{R}$, $s_\mu(x + it)$ converges as $t \to 0^+$ (see (1.2) in [PSZ10]). Moreover, the real and imaginary parts of the limit correspond respectively to the Hilbert transform of μ at x, denoted by $H_\mu(x)$, and to the density f_μ of the absolutely continuous part of μ with respect to the Lebesgue measure (see formula (1.2.8) of Simon [Sim05]). We also remind that the Hilbert transform coincides with the Cauchy principal value $H_\mu(x) = \mathrm{VP} \int \frac{\mathrm{d}\mu(y)}{y-x}$. In short, for almost every $x \in \mathbf{R}$,

$$\lim_{t \to 0^+} s_\mu(x + it) = H_\mu(x) + i\pi f_\mu(x).$$

Theorem 7. *Let μ be a probability measure on $[0,1]$ and suppose that, for all $n \geq 1$,*

$$\begin{cases} K(n) = \int_0^1 (1-x)^{n-1} x \mathrm{d}\mu(x), \\ K(\{\infty\}) = \mu(\{0\}). \end{cases} \tag{5.1}$$

Then, there exists a probability measure ν on $[0,1]$ such that:

$$s_\nu(z) s_\mu(1-z) = \frac{1}{z(1-z)}. \tag{5.2}$$

Moreover, for all $N \geq 0$:

$$\mathbf{P}(N \in \tau) = \int_0^1 x^N \mathrm{d}\nu(x). \tag{5.3}$$

Remark 5. *The mapping from μ to ν is an involution. In particular, if one assumes (5.3), then the inter-arrival distribution satisfies (5.1) with μ defined by*

$$s_\mu(z) = \left[z(1-z) s_\nu(1-z) \right]^{-1}.$$

This presents some interest for the applications. Indeed, one usually has access to the renewal measure $\mathbf{P}(N \in \tau)$ and would like to infer the underlying inter-arrival distribution.

Remark 6. *The family of generalized Arcsine laws with parameters $(1-v, v)$, $v \in (0,1)$, defined by the densities*

$$\frac{\sin(\pi v)}{\pi} x^{-v} (1-x)^{v-1} \mathbf{1}_{x \in [0,1]} \mathrm{d}x, \tag{5.4}$$

are fixed points of the involution $\mu \mapsto \nu$. This can be easily checked from the expression of their Stieltjes transform, which are equal to

$$\frac{1}{1-z} \left(\frac{z}{1-z} \right)^{-v}.$$

We conjecture that these distributions are, in fact, the only fixed points. This family of measures will be further investigated in Section 5.4 in the context of random polymers.

Before proving this Theorem, we state as a Corollary and *without proof* some more explicit formulas for $\mathbf{P}(N \in \tau)$ in two specific but generic cases.

The first one deals with the case where μ is a finite sum of Dirac masses. This case was already considered in the frame of the Fixman-Freire algorithm [FF77]. Indeed, these authors wanted to use an approximation of μ by a finite sum of Dirac masses in order to compute an approximation of $\mathbf{P}(N \in \tau)$. It turns out that they compute this last quantity using a rather heavy recursive scheme and we give here a clear and tractable formula for it.

The second one deals with the cases where μ admits a density with respect to the Lebesgue measure, supported on a finite union of intervals. This includes the previous case considered by Nagaev [Nag15] which assumed among other technical hypothesis that the density of μ is Lipschitz. The formula we give below for the density of ν is quite synthetic and seems to amend the intricate formula of Nagaev.

Corollary 10. *1. When μ is a finite sum of Dirac masses, this is also the case for ν. More precisely, let $0 < x_1 < x_2 < \cdots < x_n < 1$ and $a_1, \ldots, a_n > 0$ with $a_1 + \cdots + a_n = 1$ be such that $\mu = \sum_{1 \leq i \leq n} a_i \delta_{x_i}$. Then,*

$$s_\nu(z) = \frac{1}{z(1-z) \sum_{1 \leq i \leq n} \frac{a_i}{z-(1-x_i)}}$$

and ν is purely atomic and admits exactly $n+1$ atoms which are located in increasing order at $0, y_1, \dots, y_{n-1}, 1$ where y_i is the only root of $x \mapsto \sum_{1\le i\le n} \frac{a_i}{z-(1-x_i)}$ on the interval $(1-x_{n-i+1}, 1-x_{n-i})$. Moreover, the mass $\nu(\{y_i\})$ is given by the residue of the rational function s_ν at y_i. In that case, Equation (5.3) rewrites

$$\mathbf{P}(N \in \tau) = \frac{1}{m_K} + \sum_{1\le i\le n-1} y_i^N \nu(\{y_i\}).$$

2. *When μ is absolutely continuous with respect to the Lebesgue measure with density f_μ supported on a finite union of disjoint intervals $[a_1, b_1] \sqcup [a_2, b_2] \sqcup \cdots \sqcup [a_n, b_n]$ with $0 \le a_1 < b_1 < a_2 < \cdots < a_n < b_n \le 1$, the measure ν is a sum of a pure point measure ν_{pp} and an absolutely continuous one ν_{ac}.*

   The measure ν_{pp} admits exactly one atom on each interval $(1-a_{n-i+1}, 1-b_{n-i})$ located at the point y_i defined as the only root on that interval of the function $x \mapsto \int_0^1 \frac{\mathrm{d}\mu(s)}{s-(1-x)}$. The mass $\nu(\{y_i\})$ is given by the residue of the function s_ν at y_i. Moreover, if $m_K < \infty$, then ν admits an extra atom at 1 with mass $1/m_K$, and if $\int_0^1 \frac{\mathrm{d}\mu(x)}{1-x} < \infty$, then ν admits also an extra atom at 0 with mass $(\int_0^1 \frac{\mathrm{d}\mu(x)}{1-x})^{-1}$. Apart from these points, ν_{pp} admits no other atom.

   The support of the measure ν_{ac} is $[1-b_n, 1-a_n] \sqcup \cdots \sqcup [1-b_1, 1-a_1]$, and its density is given by:

$$\begin{aligned} f_{\nu_{ac}}(x) &= \lim_{t\to 0^+} \frac{1}{\pi} \Im s_\nu(x+it) \\ &= \frac{1}{x(1-x)} \frac{f_\mu(1-x)}{H_\mu(1-x)^2 + \pi^2 f_\mu(1-x)^2}. \end{aligned}$$

   In that case, Equation (5.3) rewrites

$$\mathbf{P}(N \in \tau) = \frac{1}{m_K} + \sum_{1\le i\le n-1} y_i^N \nu(\{y_i\}) + \int_0^1 \frac{(1-x)^{N-1}}{x} \frac{\mathrm{d}\mu(x)}{H_\mu(x)^2 + \pi^2 f_\mu(x)^2}.$$

Proof of Theorem 7. Let $G(z) = \sum_{N\ge 0} \mathbf{P}(N \in \tau) z^N$ be the generating series associated to the sequence $(\mathbf{P}(N \in \tau))_{N\ge 0}$. Then,

$$\begin{aligned} G(z) &= 1 + \mathbf{E}\left[\sum_{k\ge 1} z^{\eta_1+\cdots+\eta_k}\right] \\ &= \frac{1}{1-\mathbf{E}[z^{\tau_1}]}. \end{aligned}$$

Therefore, using (5.1), we deduce that:

$$-\frac{1}{z} G\left(\frac{1}{z}\right) = \frac{1}{-z - \int_0^1 \frac{zx}{1-z-x} \mathrm{d}\mu(x)}. \tag{5.5}$$

The above equality holds for all complex $z \in \mathbf{C}$ such that $1/z$ is inside the disk of convergence of G. It extends analytically to the whole complex upper half-plane $\mathbf{C}_+$ thanks to the right-hand side expression. Now, notice that for all $x \in (0,1)$, the homographic function $z \mapsto \frac{zx}{1-z-x}$ preserves $\mathbf{C}_+$ since it is of determinant $x(1-x) > 0$. Therefore, the function $z \mapsto -\frac{1}{z} G\left(\frac{1}{z}\right)$ preserves $\mathbf{C}_+$ and is a Nevanlinna function. Moreover, $-\frac{1}{z} G\left(\frac{1}{z}\right) \sim -\frac{1}{z}$ as $|z| \to +\infty$. Therefore,

using the characterization of Nevanlinna functions which are Stieltjes transforms of probability measures [Akh65] (page 93), there exists a probability measure ν such that

$$-\frac{1}{z}G\left(\frac{1}{z}\right) = \int_{\mathbf{R}} \frac{\mathrm{d}\nu(x)}{x-z}. \tag{5.6}$$

Identifying the $1/z$ coefficients in Equation (5.6) leads to Equation (5.3).

Once we have got the existence of the measure ν, we want to prove it satisfies Equation (5.2). Indeed, combining (5.5) and (5.6), we get that

$$\forall z \in \mathbf{C}_+, \quad \left(\int_{\mathbf{R}} \frac{\mathrm{d}\nu(x)}{x-z}\right)^{-1} = -z - \int_0^1 \frac{zx}{1-z-x}\mathrm{d}\mu(x). \tag{5.7}$$

Finally, let us notice that since the support of μ, $\mathrm{Supp}(\mu)$, is included in $[0,1]$, its Stieltjes transform s_μ is analytic on $\mathbf{R}\setminus(0,1)$. In turn, by Equation (5.2), the Stieltjes transform of ν is analytic on $\mathbf{R}\setminus(0,1)$, which implies that $\mathrm{Supp}(\nu)\subset[0,1]$. □

Let us comment on assumption (5.1). In the proof of Theorem 7, we crucially rely on it in order to prove that the function

$$z \mapsto -z\mathbf{E}\left[\left(\frac{1}{z}\right)^{\tau_1}\right]$$

preserves the upper half-plane. In the generic case, this property is not satisfied.

Let us finally mention that Theorem 7 allows to easily recover the classical renewal Theorem in our setting.

Proposition 6 (Basic renewal theorem)**.** *Under the setting of Theorem 7,*

$$\mathbf{P}(N \in \tau) \underset{N\to+\infty}{\longrightarrow} \frac{1}{m_K},$$

where $m_K := \sum_{n\geq 1} nK(n)$.

Proof. From Equation (5.3) and the dominated convergence Theorem,

$$\mathbf{P}(N \in \tau) \underset{N\to+\infty}{\longrightarrow} \nu(\{1\}).$$

By the dominated convergence theorem, the mass $\nu(\{1\})$ is equal to the limit of $-(it)\times s_\nu(1+it)$ as $t\to 0^+$. By Equation (5.2), this is equal to

$$\frac{1}{s_\mu(0)} = \left(\int_0^1 \frac{\mathrm{d}\mu(x)}{x}\right)^{-1} = \left(\int_0^1 \sum_{n\geq 1} n(1-x)^{n-1}x\mathrm{d}\mu(x)\right)^{-1} = \frac{1}{m_K}.$$

□

5.2 The continuous counterpart: mixture of exponential laws

In this section, we still consider $\eta = (\eta_n)_{n\geq 1}$ a sequence of i.i.d. positive random, but we now suppose that they admit a density denoted by f_η supported on $\mathbf{R}_+$. We will denote by m_η their mean.

Theorem 8. *Let μ be a probability measure on $[0,+\infty)$ and suppose that, for all $x>0$,*

$$f_\eta(x) = \int_0^{+\infty} s\,\mathrm{e}^{-sx}\,\mathrm{d}\mu(s). \tag{5.8}$$

Define by $H(x)$ the intensity of the renewal process with inter-arrivals $(\eta_i)_{i\geq 1}$.

Then, there exists a positive measure ν on $[0,+\infty)$ such that:

$$(1+s_\nu(z))s_\mu(z) = -\frac{1}{z}. \tag{5.9}$$

Moreover, for all $x>0$:

$$H(x) = \int_0^{+\infty} e^{-xs}\mathrm{d}\nu(s). \tag{5.10}$$

As in Corollary 10, we present, *without proof*, the two special cases where μ is a finite sum of Dirac masses and where μ admits a density with respect to the Lebesgue measure supported on a finite union of intervals. Notice that the first case is another presentation of the content of Chapter 5 of the book of Bladt and Nielsen [BN17], where the authors study the renewal process associated to an absorbed Markov chain which is regularly regenerated.

Corollary 11. 1. *When μ is a finite sum of Dirac masses, this is also the case for ν. More precisely, let $0<x_1<x_2<\cdots<x_n$ and $a_1,\dots,a_n>0$ with $a_1+\cdots+a_n=1$ be such that $\mu=\sum_{1\leq i\leq n} a_i\delta_{x_i}$. Then,*

$$s_\nu(z) = \frac{-1}{z\sum_{1\leq i\leq n}\frac{a_i}{x_i-z}} - 1$$

and ν is purely atomic and admits exactly n atoms which are located in increasing order at $0, y_1,\dots,y_{n-1}$ where y_i is the only root of $x\mapsto\sum_{1\leq i\leq n}\frac{a_i}{x_i-x}$ on the interval (x_i,x_{i+1}). Moreover, the mass $\nu(\{y_i\})$ is given by the residue of the rational function s_ν at y_i. In that case, Equation (5.3) rewrites

$$H(x) = \frac{1}{m_\eta} + \sum_{1\leq i\leq n-1}\mathrm{e}^{-xy_i}\,\nu(\{y_i\}).$$

2. *When μ is absolutely continuous with respect to the Lebesgue measure with density f_μ supported on a finite union of disjoint intervals $[a_1,b_1]\sqcup[a_2,b_2]\sqcup\cdots\sqcup[a_n,b_n]$ with $0\leq a_1<b_1<a_2<\cdots<a_n<b_n$, the measure ν is a sum of a pure point measure ν_{pp} and an absolutely continuous one ν_{ac}.*

   The measure ν_{pp} admits exactly one atom on each interval (b_i,a_{i+1}), $1\leq i\leq n-1$, located at the point y_i defined as the only root on that interval of the function $x\mapsto\int_0^1\frac{\mathrm{d}\mu(s)}{s-x}$ on the interval (b_i,a_{i+1}). The mass $\nu(\{y_i\})$ is given by the residue of the function s_ν at y_i. Moreover, if $m_\eta<+\infty$, then ν admits also an extra atom at 0 with mass $1/m_\eta$. Apart from these points, ν_{pp} admits no other atom.

   The support of the measure ν_{ac} is $[a_1,b_1]\sqcup\cdots\sqcup[a_n,b_n]$, and its density is given by:

$$\begin{aligned} f_{\nu_{ac}}(x) &= \lim_{t\to 0^+}\frac{1}{\pi}\Im s_\nu(x+it) \\ &= \frac{1}{x}\frac{f_\mu(x)}{H_\mu(x)^2+\pi^2 f_\mu(x)^2}. \end{aligned}$$

   In that case, Equation (5.3) rewrites

$$H(x) = \frac{1}{m_\eta} + \sum_{1\leq i\leq n-1}\mathrm{e}^{-xy_i}\,\nu(\{y_i\}) + \int_0^{+\infty}\frac{\mathrm{e}^{-sx}}{s}\frac{\mathrm{d}\mu(s)}{H_\mu(s)^2+\pi^2 f_\mu(s)^2}.$$

Proof of Theorem 8. As usual, we introduce the Laplace transform of H and denote it by $\mathcal{L}$ so that

$$\mathcal{L}(\lambda) = \int_0^{+\infty} \mathrm{e}^{-\lambda x} H(x)\mathrm{d}x.$$

Notice that this expression is finite for all $\lambda \in \mathbf{C}$ with positive real part. Since $H(x) = \sum_{k\geq 1} f_{\eta_1+\cdots+\eta_k}(x)$, we deduce that:

$$\forall \lambda \in \mathbf{C} \text{ s.t. } \Re(\lambda) > 0, \quad \mathcal{L}(\lambda) = \sum_{k\geq 1} \mathbf{E}\left[\mathrm{e}^{-\lambda(\eta_1+\cdots+\eta_k)}\right] = \frac{1}{1-\mathbf{E}\left[\mathrm{e}^{-\lambda\eta_1}\right]} - 1.$$

From Assumption (5.8), we deduce that

$$\mathbf{E}\left[\mathrm{e}^{-\lambda\eta_1}\right] = \int_0^{+\infty} \frac{s}{s+\lambda}\mathrm{d}\mu(s).$$

Hence, for all $\lambda \in \mathbf{C}$ with positive real part,

$$\mathcal{L}(\lambda) = \frac{1}{\int_0^{+\infty} \frac{\lambda}{s+\lambda}\mathrm{d}\mu(s)} - 1 = \frac{1}{\lambda s_\mu(-\lambda)} - 1.$$

Since the Stieltjes transform of μ is defined on $\mathbf{C}\setminus\mathbf{R}_+$, the function $\mathcal{L}$ can be analytically extended to $\mathbf{C}\setminus\mathbf{R}_+$. Since for all $s > 0$, the homography $\lambda \mapsto \lambda/(\lambda+s)$ preserves the lower half-plane, the function $\lambda \mapsto \mathcal{L}(-\lambda)$ preserves the upper half-plane. Moreover, when $|\lambda| \to +\infty$,

$$\mathcal{L}(-\lambda) = \frac{\int_0^{+\infty} \frac{s}{s-\lambda}\mathrm{d}\mu(s)}{\int_0^{+\infty} \frac{-\lambda}{s-\lambda}\mathrm{d}\mu(s)} \sim \frac{-1}{\lambda}\int_0^{+\infty} s\mathrm{d}\mu(s).$$

Therefore, using the characterization of Nevanlinna functions which are Stieltjes transforms of positive measures [Akh65] (page 93), there exists a positive measure ν with total mass $\int_0^{+\infty} s\mathrm{d}\mu(s)$ whenever this integral is finite, such that

$$\mathcal{L}(-\lambda) = s_\nu(\lambda). \tag{5.11}$$

□

5.3 Application to polymers pinned on a defect line

In this section, we first recall the definition and main results about a classical random polymer model associated to a renewal process, as presented in the book of Giacomin [Gia07]. Then, we present our contribution which consists in an explicit integral representation of the partition functions of the model, when the inter-arrival time distribution of the underlying renewal process is a mixture of Geometric laws.

5.3.1 Definition of the model

Let $\beta \in \mathbf{R}$ and $N \geq 1$. The polymer model associated to K with parameter β is defined by the following probability measure $\mathbf{P}_{N,\beta}$ on subsets of $\{0,\ldots,N\}$ whose density with respect to $\mathbf{P}$ is:

$$\frac{\mathrm{d}\mathbf{P}_{N,\beta}}{\mathrm{d}\mathbf{P}}(\tau) := \frac{1}{Z_{N,\beta}} \exp\left(\beta\mathcal{N}_N(\tau)\right)\mathbf{1}_{N\in\tau},$$

where $\mathcal{N}_N(\tau) = |\{1, \dots, N\} \cap \tau|$ and

$$Z_{N,\beta} := E_{\mathbf{P}} \left[\exp\left(\beta \mathcal{N}_N(\tau)\right) \mathbf{1}_{N \in \tau} \right],$$

where $E_{\mathbf{P}}$ stands for the expectation with respect to $\mathbf{P}$. The normalizing constant $Z_{N,\beta}$ is called the partition function and captures much information about the model. For example,

$$\frac{1}{N} \frac{\partial}{\partial \beta} \log Z_{N,\beta} = E_{\mathbf{P}_{N,\beta}} \left[\frac{\mathcal{N}_N(\tau)}{N} \right]$$

is the average time spent at 0 by the polymer. As N tends to infinity, it converges to the derivative of the so-called free energy of the model, which is defined by:

$$F(\beta) := \lim_{N \to +\infty} \frac{1}{N} \log Z_{N,\beta}. \tag{5.12}$$

Hence, $F'(\beta)$ corresponds to the asymptotic fraction of time spent at zero by the polymer. Therefore, we speak about a *delocalized* regime when $F'(\beta) = 0$ and about a *localized* regime when $F'(\beta) > 0$.

It has been shown that there exists a phase transition for this model. More precisely, let $\beta_c := -\log(1 - K(\{\infty\}))$. Then, if $\beta > \beta_c$, the free energy $F(\beta)$ is uniquely determined by

$$\mathbf{E} \left[\exp\left(-F(\beta)\tau_1\right) \right] = \exp\left(-\beta\right). \tag{5.13}$$

Otherwise, if $\beta \leq \beta_c$, the free energy is given by $F(\beta) = 0$. Therefore, the model exhibits a phase transition at $\beta = \beta_c$ from a delocalized regime to a localized regime.

The identification of the free energy is based on a simple rewriting of the partition function that we recall here. First, let us introduce a new family of (sub)-probability measures:

$$\forall n \geq 1, \quad \widetilde{K}_\beta(n) := \begin{cases} \exp(\beta) K(n) \exp(-F(\beta) n) & \text{if } \beta \geq \beta_c, \\ \exp(\beta) K(n) & \text{if } \beta < \beta_c. \end{cases} \tag{5.14}$$

Notice that from the definition of β_c, the measure $\widetilde{K}_\beta$ is a probability measure when $\beta > \beta_c$, whereas it is a sub-probability measure when $\beta < \beta_c$. Let $\widetilde{\mathbf{P}}_\beta$ be the law of the renewal process associated to $\widetilde{K}_\beta$. Then, summing over the inter-arrival times leads to:

$$\begin{aligned} Z_{N,\beta} &= \sum_{n=1}^{N} \sum_{l_1 + \cdots + l_n = N} \prod_{i=1}^{n} \exp(\beta) K(l_i) \\ &= \exp(F(\beta) N) \sum_{n=1}^{N} \sum_{l_1 + \ldots + l_n = N} \prod_{i=1}^{n} \widetilde{K}_\beta(l_i) \\ &= \exp(F(\beta) N) \widetilde{\mathbf{P}}_\beta(N \in \tau). \end{aligned} \tag{5.15}$$

Under the classical assumption that there exists $\alpha > 0$ and a slowly varying function such that

$$K(n) = \frac{L(n)}{n^{1+\alpha}}, \tag{5.16}$$

the probability $\widetilde{\mathbf{P}}_\beta(N \in \tau)$ does not vanish exponentially fast. Therefore, the function F defined in (5.13) is indeed the limit in (5.12), namely the free energy of the model. Moreover, thanks to the asymptotic theory of renewal processes, Equation (5.15) also allows to obtain the asymptotic leading term of $Z_{N,\beta}$ as $N \to +\infty$.

5.3.2 Moment representation of the partition function

When the inter-arrival time distribution of the underlying renewal process is a mixture of Geometric laws, the partition function $Z_{N,\beta}$ is the N-th moment of some measure ν_β which corresponds to an explicit transformation of the mixture distribution.

Theorem 9. *Let μ be a probability measure on $[0,1]$ and suppose that, for all $n \geq 1$,*

$$\begin{aligned} &K(n) = \int_0^1 (1-x)^{n-1} x \mathrm{d}\mu(x), \\ &K(\{\infty\}) = \mu(\{0\}). \end{aligned} \tag{5.17}$$

Then, for all $\beta \in \mathbf{R}$, there exists a probability measure ν_β such that:

$$s_{\nu_\beta}(z)\left(\mathrm{e}^\beta s_\mu(1-z) - \frac{1-\mathrm{e}^\beta}{1-z}\right) = \frac{1}{z(1-z)}. \tag{5.18}$$

Moreover, for all $N \geq 0$:

$$Z_{N,\beta} = \int_{\mathbf{R}} x^N \mathrm{d}\nu_\beta(x). \tag{5.19}$$

Proof. Let $\beta \in \mathbf{R}$ and define, for all complex number in the upper half-plane $z \in \mathbf{C}_+$, $z_\beta = z\exp(F(\beta))$. Thanks to Equation (5.15), the generating function associated to the sequence $(Z_{N,\beta})_{N\geq 0}$ is equal to:

$$\begin{aligned} G(z) :&= \sum_{N\geq 0} Z_{N,\beta} z^N \\ &= 1 + \sum_{N\geq 1} \widetilde{\mathbf{P}}_\beta(N \in \tau) z_\beta^N \\ &= 1 + E_{\widetilde{\mathbf{P}}_\beta}\left[\sum_{N\geq 1} \mathbf{1}_{N\in\tau} z_\beta^N\right]. \end{aligned} \tag{5.20}$$

Let $\left(\widetilde{\eta}_i^{(\beta)}\right)_{i\geq 1}$ be a sequence of i.i.d. random variables with law $\widetilde{K}_\beta$. Then, Equation (5.20) becomes:

$$G(z) = 1 + E_{\widetilde{\mathbf{P}}_\beta}\left[\sum_{k\geq 1} z_\beta^{\widetilde{\eta}_1^{(\beta)}+\cdots+\widetilde{\eta}_k^{(\beta)}}\right] = \frac{1}{1 - E_{\widetilde{\mathbf{P}}_\beta}\left[z_\beta^{\widetilde{\eta}_1^{(\beta)}}\right]}.$$

Finally, from the definition of $\widetilde{K}_\beta$ given in (5.14), we obtain:

$$G(z) = \frac{1}{1 - \exp(\beta) E_{\mathbf{P}}\left[z^{\tau_1}\right]}.$$

Let $S(z) = -\frac{1}{z} G\left(\frac{1}{z}\right)$. Then, using Hypothesis (5.17), it is easy to deduce that:

$$S(z) = \frac{1}{-z - \exp(\beta)\int_0^1 \frac{zx}{1-z-x}\mathrm{d}\mu(x)}. \tag{5.21}$$

Notice that $S(z) \sim -\frac{1}{z}$ as $|z| \to +\infty$. Moreover, since for all $z \in \mathbf{C}_+$ and $x \in (0,1)$, $\frac{zx}{1-z-x} \in \mathbf{C}_+$, the following inclusion holds: $S(\mathbf{C}_+) \subset \mathbf{C}_+$. By [Akh65] (page 93), these two properties imply that there exists a probability measure ν_β such that:

$$S(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\nu_\beta(x)}{x - z},$$

which ends the proof of Theorem 9. □

Remark 7. *Note that by (5.21), the Stieltjes transform of ν_β is also given by*

$$s_{\nu_\beta}(z) = \frac{1}{-z - \exp(\beta)\int_0^1 \frac{zx}{1-z-x}\mathrm{d}\mu(x)}.$$

Observe that the function $y \mapsto \int_0^1 \frac{x}{1-y-x}\mathrm{d}\mu(x)$ is increasing on $(1,+\infty)$ and tends to $-\mu((0,1]) = -\sum_{n\geq 1} K(n) = -(1-K(\{\infty\}))$ as $y \to 1^+$. Therefore, the denominator of s_{ν_β} has a single root x_β on $(1,+\infty)$, which implies that the restriction of ν_β on $(1,+\infty)$ consists in a unique atom at $x_\beta > 1$ if $\beta > -\log(1-K(\{\infty\}))$ and is null otherwise. By Equation (5.19), $F(\beta) = \log(x_\beta)$ if $\beta > -\log(1-K(\{\infty\}))$, and $F(\beta) = 0$ otherwise.

The mass of the atom of ν_β at $\exp(F(\beta))$ is given by the residue of s_{ν_β} at $\exp(F(\beta))$ which happens to be equal to $F'(\beta)$, which is the average time spent at zero by the polymer. In particular, as N tends to infinity,

$$Z_{N,\beta} \sim F'(\beta)\exp(NF(\beta)).$$

This formula was already proved in a general framework, see for example Chapter 2 of [Gia07].

5.3.3 Correlation length

In this section, we are interested, for all $b > 0$, in the renewal process $\tau(b)$ with inter-arrival distribution defined by

$$K_b(n) = \frac{1}{c(b)}K(n)\exp(-nb),$$

and with mean value denoted by m_{K_b}. Namely, we want to compute the rate of convergence in the renewal theorem associated to this process. It is motivated by the article of Giacomin [Gia08], which makes the link between this quantity and the correlation decay for the polymer with law $\mathbf{P}_{N,\beta}$, where $\beta > 0$ corresponds to the only positive solution of the equation $b = F(\beta)$, in the limit $N \to +\infty$.

Proposition 7. *Suppose that μ is a probability measure on $[0,1]$ and that*

$$K(n) = \int_0^1 (1-x)^{n-1}x\mathrm{d}\mu(x),$$
$$K(\{\infty\}) = \mu(\{0\}).$$

Then, for all $b > 0$, there exists a probability measure ν_b on $[0,1]$ such that for all $N \geq 0$,

$$\mathbf{P}(N \in \tau(b)) = \int_0^1 x^N \mathrm{d}\nu_b(x). \tag{5.22}$$

Moreover, if μ has a positive density on a neighborhood of 0 and that

$$\lim_{N\to+\infty} \frac{1}{N}\log\left(\mathbf{P}(N \in \tau(b)) - \frac{1}{m_{K_b}}\right) = -b. \tag{5.23}$$

Remark 8. *Notice that Equation (5.23) is valid for all $b > 0$, whereas the general result of Giacomin [Gia08, Theorem 1.1] establishes this formula only for parameters $b > 0$ up to some positive and implicit value $b_0 > 0$.*

Remark 9. *Suppose that there exists $a \in (0,1)$ such that $\mu([0,a)) = 0$ and such that μ has a positive density at a^+. Then, the limit (5.23) is equal to $-b + \log(1-a)$.*

Proof of Proposition 7. We proceed like in the proof of Theorem 7. Namely, denoting by $G_b(z) = \sum_{N\geq 0} \mathbf{P}(N \in \tau(b))z^N$ the generating series of the renewal probabilities, a direct computation leads to:

$$G_b(z) = \frac{1}{1 - \mathrm{e}^{\beta}\, z\mathbf{E}\left[(z\,\mathrm{e}^{-b})^{\tau_1}\right]}.$$

$$\begin{aligned}
-\frac{1}{z}G_b\left(\frac{1}{z}\right) &= \frac{-1}{-z - \mathrm{e}^{\beta-b} \times (\mathrm{e}^b\, z)\mathbf{E}\left[\left(\frac{1}{z\,\mathrm{e}^b}\right)^{\tau_1}\right]} \\
&= \frac{-1}{-z - \mathrm{e}^{\beta-b} \times \int_0^1 \frac{z\mathrm{e}^b x}{1-z\,\mathrm{e}^b - x}\mathrm{d}\mu(x)} \\
&= \frac{-1}{-z - \mathrm{e}^{\beta} \times \int_0^1 \frac{zx}{1-z\,\mathrm{e}^b - x}\mathrm{d}\mu(x)} \\
&= \frac{1}{-z + \mathrm{e}^{\beta}\, z + z\,\mathrm{e}^{\beta}(1 - z\,\mathrm{e}^b)s_\mu(1 - z\,\mathrm{e}^b)}.
\end{aligned}$$

Using the penultimate equality, one can check that the analytic function $z \mapsto -\frac{1}{z}G_b(\frac{1}{z})$ preserves the upper half-plane. Moreover, $-\frac{1}{z}G_b(\frac{1}{z}) \sim -1/z$ as $|z| \to +\infty$. Therefore, by [Akh65] (page 93), there exists a probability measure ν_b such that

$$s_{\nu_b}(z) = \frac{1}{-z + \mathrm{e}^{\beta}\, z + z\,\mathrm{e}^{\beta}(1 - z\,\mathrm{e}^b)s_\mu(1 - z\,\mathrm{e}^b)}. \tag{5.24}$$

This yields Equation (5.22).

Notice that, by a monotonicity argument, the denominator of s_{ν_b} vanishes only at $z = 0$ and $z = 1$. Moreover, as in Proposition (6), one can prove that $\nu_b(\{1\}) = 1/m_{K_b}$. Finally, the probability measure ν_b has a positive density at x if and only if $\lim_{t\to 0^+} \Im s_{\nu_b}(x + it) > 0$, which, by Equation (5.24), is the case if and only if $\lim_{t\to 0^+} \Im s_\mu(1 - (x + it)\,\mathrm{e}^b) > 0$. Since by assumption, the probability measure μ has a density in a neighborhood of 0^+, we deduce that ν_b has a density in a neighborhood of $(\mathrm{e}^{-b})^-$. Moreover, by Equation (5.24), for all $x \in (\mathrm{e}^{-b}, 1)$:

- $\lim_{t\to 0^+} |s_{\nu_b}(x + it)| < +\infty$,
- $\lim_{t\to 0^+} \Im(s_{\nu_b}(x + it)) = 0$.

From the first fact (see formula (1.2.10) in [Sim05]), we deduce that ν_b has no singular part on $(\mathrm{e}^{-b}, 1)$. On the other hand, the second fact implies that ν_b has no absolutely continuous part either. Hence ν does not charge the interval $(\mathrm{e}^{-b}, 1)$. This yields (5.23).

□

5.4 Computations in the case of generalized Arcsine laws

It turns out that some particular choices of probability measure μ in Theorem 9 yield explicit computations. More precisely, for all $v \in (0, 1)$, let μ_v be the Beta distribution with parameters $(1 - v, v)$, defined by:

$$\mathrm{d}\mu_v(x) := \frac{\sin(\pi v)}{\pi} x^{-v}(1 - x)^{v-1}\mathbf{1}_{x\in[0,1]}\mathrm{d}x.$$

Let K_v be the probability measure on $\mathbf{N}$ associated to μ_v, that is:

$$\forall n \geq 1, \quad K_v(n) = \int_{\mathbf{R}} (1 - x)^{n-1} x \mathrm{d}\mu_v(x) = \frac{\sin(\pi v)}{\pi} \frac{\Gamma(n + v - 1)\Gamma(2 - v)}{\Gamma(n + 1)}.$$

As $n \to +\infty$, $K_v(n) \sim \frac{\sin(\pi v)\Gamma(2-v)}{\pi} \frac{1}{n^{2-v}}$. Hence, the probability measures K_v's satisfy (5.16).

Denoting $Z_{N,\beta,v}$ the partition function of the polymer associated to K_v, Theorem 9 translates into the following result.

Proposition 8. *For all $\beta \in \mathbf{R}$ and $v \in (0,1)$, define:*

$$f_{v,\beta}(x) = \frac{\sin(\pi v)}{\pi x} \frac{\mathrm{e}^{\beta}\, x^{1-v}(1-x)^{1-v}}{(1-\mathrm{e}^{\beta})^2 x^{2(1-v)} - 2\,\mathrm{e}^{\beta}(1-\mathrm{e}^{\beta})\cos(\pi v) x^{1-v}(1-x)^{1-v} + \mathrm{e}^{2\beta}(1-x)^{2(1-v)}}.$$

Moreover, for all $\beta > 0$, define:

$$\gamma_{v,\beta} = (1-\mathrm{e}^{-\beta})^{\frac{1}{1-v}}, \qquad x_{v,\beta} = \frac{1}{1-\gamma_{v,\beta}} \quad \text{and} \quad c_{v,\beta} = \frac{\exp(-\beta)}{1-v} \frac{\gamma_{v,\beta}^{v}}{1-\gamma_{v,\beta}}.$$

For all $\beta \in \mathbf{R}$, let $\nu_{v,\beta}$ be the following probability measure:

$$\mathrm{d}\nu_{v,\beta}(x) = f_{v,\beta}(x)\mathbf{1}_{x\in(0,1)}\mathrm{d}x + \mathbf{1}_{\beta>0} c_{v,\beta} \mathrm{d}\delta_{x_{v,\beta}}(x).$$

Then, for all $N \geq 0$,

$$Z_{N,\beta,v} = \int_{\mathbf{R}} x^N \mathrm{d}\nu_{v,\beta}(x).$$

Remark 10. *Using Remark 7, we deduce that the constant $c_{v,\beta}$ is a positive constant smaller than 1 which is equal to the asymptotic average time spent at 0 as N tends to infinity.*

Proof. Let $v \in (0,1)$ and $\beta \in \mathbf{R}$. Recall that from Equation (5.18), the Stieltjes transform of $\nu_{v,\beta}$ satisfies:

$$\forall z \in \mathbf{C}_+, \quad \frac{1}{s_{\nu_{v,\beta}}(z)} = z(1-z)\left(\mathrm{e}^{\beta} s_{\mu_v}(1-z) - \frac{1-\mathrm{e}^{\beta}}{1-z}\right). \tag{5.25}$$

For all $x \in \mathbf{R}$, let us define $s_{\nu_{v,\beta}}(x) := \lim_{t\to 0^+} s_{\nu_{v,\beta}}(x+it)$ and $s_{\mu_v}(x) := \lim_{t\to 0^+} s_{\mu_v}(x-it)$. Then, Equation (5.25) becomes:

$$\frac{1}{s_{\nu_{v,\beta}}(x)} = x(1-x)\left(\mathrm{e}^{\beta} s_{\mu_v}(1-x) - \frac{1-\mathrm{e}^{\beta}}{1-x}\right). \tag{5.26}$$

We now use the following identity:

$$s_{\mu_v}(x) = \mathrm{VP}\int_{\mathbf{R}} \frac{\mathrm{d}\widetilde{\mu_v}(y)}{y-x} - i\pi \frac{\mathrm{d}\widetilde{\mu_v}(x)}{\mathrm{d}x}(x), \tag{5.27}$$

where the integral in the right-hand side stands for a Cauchy principal value. The latter can be explicitly computed (see [Bat54] page 250). More precisely:

$$\frac{1}{\pi}\mathrm{VP}\int_{\mathbf{R}} \frac{y^{-v}(1-y)^{v-1}}{y-x}\mathrm{d}y = \begin{cases} \frac{1}{\sin(\pi(1-v))}\frac{1}{1-x}\left|\frac{x}{1-x}\right|^{-v} & \text{if } x<0 \text{ or } x>1, \\ -x^{-v}(1-x)^{v-1}\frac{\cos(\pi(1-v))}{\sin(\pi(1-v))} & \text{if } 0<x<1. \end{cases}$$

Using (5.27), we deduce that:

$$s_{\mu_v}(1-x) = \begin{cases} -\frac{1}{x}\left|\frac{1-x}{x}\right|^{-v} & \text{if } x<0 \text{ or } x>1, \\ \cos(\pi v)(1-x)^{-v}x^{v-1} - i\sin(\pi v)x^{-v}(1-x)^{v-1} & \text{if } 0<x<1. \end{cases} \tag{5.28}$$

From (5.28) and (5.26), it is possible to identify the measure $\nu_{v,\beta}$ as explained in the following.

First, the absolutely continuous part of $\nu_{v,\beta}$ is given by $\frac{1}{\pi}\Im s_{\nu_{v,\beta}}(x)$. Therefore, it is supported on the interval $(0,1)$ and given by:

$$\frac{\mathrm{d}\nu_{v,\beta}}{\mathrm{d}x}(x) = \frac{\sin(\pi v)}{\pi x} \frac{\mathrm{e}^{\beta}\, x^{1-v}(1-x)^{1-v}}{(1-\mathrm{e}^{\beta})^2 x^{2(1-v)} - 2\,\mathrm{e}^{\beta}(1-\mathrm{e}^{\beta})\cos(\pi v) x^{1-v}(1-x)^{1-v} + \mathrm{e}^{2\beta}(1-x)^{2(1-v)}}.$$

Now, $\nu_{v,\beta}$ has an atom at $x \in \mathbf{R}$ if $s_{\nu_{v,\beta}}(x) = \infty$. Therefore, the atomic part is contained in $\mathbf{R} \setminus [0,1]$ and $x \in \mathbf{R} \setminus [0,1]$ is an atom of $\nu_{v,\beta}$ if and only if:

$$1 + \mathrm{e}^{\beta}\left(\frac{x}{x-1}\right)^{v-1} - \mathrm{e}^{\beta} = 0 \quad \Longleftrightarrow \quad x = 1 + \frac{(1-\mathrm{e}^{-\beta})^{\frac{1}{1-v}}}{1-(1-\mathrm{e}^{-\beta})^{\frac{1}{1-v}}}.$$

The right-hand side does not belong to $[0,1]$ if and only if $\beta > 0$. Therefore, we have the following dichotomy:

- if $\beta > 0$, the measure $\nu_{v,\beta}$ has an atom at $x_{v,\beta} := 1 + \frac{(1-\mathrm{e}^{-\beta})^{\frac{1}{1-v}}}{1-(1-\mathrm{e}^{-\beta})^{\frac{1}{1-v}}} > 1$;
- if $\beta \leq 0$, the measure $\nu_{v,\beta}$ has no atom.

Suppose that $\beta > 0$. Then, the atom $x_{v,\beta}$ coincides with $\exp(F(\beta))$ and by Remark 7 the mass of $\nu_{v,\beta}$ at $x_{v,\beta}$ is equal to $F'(\beta) = \partial_\beta(x_{v,\beta})/x_{v,\beta}$, which yields the expression of $c_{v,\beta}$. □

Then, a straightforward consequence of Proposition 8 is the following explicit formula for the free energy of the model, defined in (5.12).

Corollary 12. *The following equality holds:*

$$F_\nu(\beta) = \begin{cases} 0 & \text{if } \beta \leq 0, \\ \log\left(\frac{1}{1-\gamma_{v,\beta}}\right) & \text{if } \beta > 0. \end{cases} \tag{5.29}$$

Moreover, when $\beta > 0$, as $N \to +\infty$,

$$Z_{N,\beta,v} \sim \frac{\exp(-\beta)}{1-v} \frac{\gamma_{v,\beta}^v}{1-\gamma_{v,\beta}} \left(\frac{1}{1-\gamma_{v,\beta}}\right)^N. \tag{5.30}$$

When $\beta = 0$,

$$Z_{N,\beta} = \frac{\sin(\pi v)}{\pi} \int_0^1 x^{N-v}(1-x)^{v-1}\mathrm{d}x = \frac{\Gamma(N+1-v)}{\Gamma(1-v)\Gamma(N+1)} \sim \frac{N^{-v}}{\Gamma(1-v)}.$$

5.5 An epilogue on random matrix theory

We end this article by presenting a link between the phase transition of a particular instance of the above pinned model and a famous phase transition in random matrix theory. Since it was the starting point of the present work, we think it justifies its presence here, at least in our view.

When $v = 1/2$, μ_v is the classical Arcsine law:

$$\mathrm{d}\mu_{\frac{1}{2}}(x) = \frac{1}{\pi}\frac{1}{\sqrt{(1-x)x}}\mathbf{1}_{x\in(0,1)}\mathrm{d}x.$$

In that case,

$$\forall n \geq 1, \quad K_{\frac{1}{2}}(n) = \frac{1}{2^{2n}} \frac{1}{2n-1} \binom{2n}{n},$$

which is also the probability that the first return to 0 of the simple random walk is equal to $2n$, see for example [Fel68]. It turns out that in this setting, the phase transition from the delocalized regime to the localized regime for the polymer model corresponds to a famous phase transition in random matrix theory, which we briefly recall in the following.

For all $n \geq 1$, let X_n be a matrix of size $n \times n$ whose entries are i.i.d. random variables, centered and reduced. Let also $\Sigma_n = \mathrm{Diag}(2\,\mathrm{e}^{\beta}, 1, \ldots, 1)$, where $\beta \in \mathbf{R}$. We consider the following random covariance matrix:

$$S_n = \frac{1}{4n} \Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2}.$$

Denoting $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of S_n, it turns out that, in probability, the empirical spectral measure $\frac{1}{n} \sum_{1 \leq i \leq n} \delta_{\lambda_i}$ weakly converges towards the so-called Marchenko-Pastur law with parameter 1, given by the density $(2/\pi)(1-x)^{1/2} x^{-1/2} \mathbf{1}_{0<x<1} \mathrm{d}x$. Moreover, the possible existence of an eigenvalue outside the limiting support $(0,1)$, often called an *outlier*, is the object of the following phase transition. We will denote by ϕ_1 the normalized eigenvector associated to λ_1.

Theorem (Baïk, Ben Arous, Péché phase transition)**.** *Let e_1 be the first vector of the canonical basis. Then, the following converges hold in probability:*

$$\lambda_1 \underset{n \to +\infty}{\longrightarrow} \begin{cases} 1 & \text{if } \beta \leq 0, \\ \frac{\mathrm{e}^{2\beta}}{2\,\mathrm{e}^{\beta}-1} & \text{otherwise,} \end{cases} \qquad \text{and} \qquad |\langle \phi_1, e_1 \rangle|^2 \underset{n \to +\infty}{\longrightarrow} \begin{cases} 0 & \text{if } \beta \leq 0, \\ \frac{2\,\mathrm{e}^{\beta}-2}{\mathrm{e}^{\beta}-1} & \text{otherwise.} \end{cases}$$

This result was first proved by Baïk, Ben Arous and Péché [BBAP05] in a Gaussian setting. Another approach to this problem is to study the *spectral measure in direction e_1* – see [Noi20], defined by

$$\mu_{(S_n, e_1)} := \sum_{i=1}^{n} |\langle \phi_i, e_1 \rangle|^2 \delta_{\lambda_i},$$

where ϕ_i is the normalized eigenvector associated to λ_i. With our notations, it turns out that in probability, $\mu_{(S_n, e_1)}$ weakly converges to $\nu_{\frac{1}{2}, \beta}$, which is given by

$$\mathrm{d}\nu_{\frac{1}{2},\beta}(x) = \frac{\mathrm{e}^{\beta}}{\pi x} \frac{\sqrt{(1-x)x}}{x(1-2\,\mathrm{e}^{\beta}) + \mathrm{e}^{2\beta}} \mathbf{1}_{0<x<1} \mathrm{d}x + \frac{2\,\mathrm{e}^{\beta}-2}{2\,\mathrm{e}^{\beta}-1} 1_{\beta>0} \mathrm{d}\delta_{\frac{\mathrm{e}^{2\beta}}{2\mathrm{e}^{\beta}-1}}(x).$$

In particular, the atomic part of $\nu_{\frac{1}{2},\beta}$ allows to retrieve the convergences of Theorem 5.5. Interestingly, this links the Baïk, Ben Arous and Péché phase transition for the largest eigenvalue of deformed random covariance matrices to the phase transition from the delocalized to the localized regime for the polymer model. In the super-critical regimes, the limit of $\log \lambda_1$ is the free energy of the polymer, and the limit of the square projection of the associated eigenvector is the multiplicative factor in front of the exponential term of the partition function – this can be seen from Equations (5.29) and (5.30).

Chapter 6

Depth First Exploration of a Configuration Model

This chapter corresponds to the prepublication [EFMN19].

We introduce an algorithm that constructs a random uniform graph with prescribed degree sequence together with a depth first exploration of it. In the so-called supercritical regime where the graph contains a giant component, we prove that the rescaled contour process of the Depth First Search Tree has a deterministic limiting profile that we identify. The proof goes through a detailed analysis of the evolution of the empirical degree distribution of unexplored vertices. This evolution is driven by an infinite system of differential equations which has a unique and explicit solution. As a byproduct, we deduce the existence of a macroscopic simple path and get a lower bound on its length.

6.1 Introduction

Historically, the configuration model was introduced by Bender and Canfield [BC78], Bollobás [Bol80] and Wormald [Wor80] as a random multigraph with N vertices and prescribed degree sequence $d_1, \dots, d_N$. It turns out that this model shares a lot of features with the Erdős-Rényi random graph. In particular it exhibits a phase transition for the existence of a unique macroscopic connected component. This phase transition, as well as the size of this so-called giant component, was studied in detail in [MR95, MR98, JL09]. The proof of these results relies on the analysis of a construction algorithm which takes as input a collection of N vertices having respectively $d_1, \dots, d_N$ half-edges coming out of them, and returns as output a random multigraph with degree sequence $d_1, \dots, d_N$, by connecting step by step the half-edges. The way [MR95, MR98, JL09, BR15] connect these half-edges is as follows: at a given step in this algorithm, a uniform half-edge of the growing cluster is connected to a uniform not yet connected half-edge.

In this paper, we introduce a construction algorithm which, in addition to constructing the configuration model, provides an exploration of it. This exploration corresponds to the Depth First Search algorithm which is roughly a nearest neighbor walk on the vertices that greedily tries to go as deep as possible in the graph. The output of the Depth First Search Algorithm is a spanning rooted plane tree for each connected component of the graph, whose height provides a lower bound on the length of the largest simple path in the corresponding component.

A similar exploration (namely, a breadth-first exploration) has been successfuly used by Aldous [Ald97] for the Erdős-Rényi model in the critical window where the connected com-

ponenents are of polynomial size. The structure of the graph in this window was further studied in [ABBG12]. For the configuration model, a similar critical window was also identified and studied. See [HM12, Rio12, BR15, DvdHvLS17].

The purpose of this article is to study this algorithm on a supercritical configuration model and in particular the limiting shape of the contour process of the tree associated to the Depth First exploration of the giant component. Unlike in the previous construction of [MR95, MR98, JL09, BR15], where the authors only studied the evolution of the empirical distribution of the degree of the unexplored vertices, we have to deal with the empirical distribution of the degree of the unexplored vertices in the graph that they induce inside the final graph. The analysis of this evolution is much more delicate and is in fact the heart of our work, this is the content of Theorem 10.

It turns out that a step by step analysis of the construction does not work. Still, it is possible to track, at some ladder times, the evolution of the degrees of the unexplored vertices in the graph they induce. In this time scale, using a generalization of the celebrated differential equations method of Wormald [Wor95] provided in the appendix, we are able to show that the evolution of the empirical degree distribution of the unexplored vertices has a fluid limit which is driven by an infinite system of differential equations. This system as such cannot be handled. We have to introduce a time change which, surprisingly, corresponds to the proportion of explored vertices, in term of the construction algorithm. Another surprise is that the resulting new system of differential equations admits an explicit solution through the generating series they form. In order to apply Wormald's method, we need to establish the uniqueness of this solution. This task, presented in Section 6.6.2, is also intricate and is based on the knowledge of the explicit solution mentioned above.

Combining Theorem 10 with an analysis of the ladder times, we prove that the renormalized contour process of the spanning tree of the Depth First Search algorithm converges to a deterministic profile for which we give an explicit parametric representation. This is the object of Theorem 11. A direct consequence is a lower bound on the length of the longest simple path in a supercritical model, see Corollary 13. To the best of our knowledge, this lower bound seems to be the best available for a generic initial degree distribution. We refer here to the work [FJ87], where the authors establish a lower bound on the longest induced path in a configuration model with bounded degree with a bound that becomes microscopic as the largest degree tends to infinity. We do not believe that our bound is sharp. The question of the length of the longest simple path in a configuration model is actually still open in generic cases. To the best of our knowledge, the only solved cases are d-regular random graphs that are known to be (almost) Hamiltonian [Bol83]. However, a main advantage of our bound is that it is given by an explicit contruction in linear time, which is not the case for the regular graphs setting.

Let us mention that the ingredient of ladder times, used in the proof of Theorem 11, was already present in the context of Erdős-Rényi graphs in [EFM20]. The novelty and core of the present article is the analysis of the empirical degree distribution of the unexplored vertices at the ladder times, which was straightforward in the case of Erdős-Rényi graphs as it is in that case, along the construction, a Binomial distribution with decreasing parameter.

In order to illustrate our results, we provide explicit computations together with simulations in the setting where the initial degree distribution follows respectively a Poisson law (recovering results of [EFM20] in the Erdős-Rényi setting), a Dirac mass at $d \geq 3$ (corresponding to d-regular random graphs) and a Geometric law. We also discuss briefly the heavy tailed case which also falls into the scope of our results.

6.2 Definition of the DFS exploration and main results

6.2.1 The Depth First Search algorithm

Consider a multigraph $G = (V, E)$ whit vertex set $V = \{1, \dots, N\}$. The DFS exploration of G is the following algorithm.

For every step n we consider the following objects, defined by induction.

- A_n, the active vertices, is an ordered list of elements of V.
- S_n, the sleeping vertices, is a subset of V. This subset will never contain a vertex of A_n.
- R_n, the retired vertices, is another subset of V composed of all the vertices that are neither in A_n nor S_n.

At time $n = 0$, choose a vertex v uniformly at random. Set:

$$\begin{cases} A_0 &= (v), \\ S_0 &= V_N \setminus \{v\}, \\ R_0 &= \emptyset. \end{cases}$$

Suppose that A_n, S_n and R_n are constructed. Three cases are possible.

1. If $A_n = \emptyset$, the algorithm has just finished exploring a connected component of G. In that case, we pick a vertex v_{n+1} uniformly at random inside S_n and set:
$$\begin{cases} A_{n+1} &= (v_{n+1}), \\ S_{n+1} &= S_n \setminus \{v_{n+1}\}, \\ R_{n+1} &= R_n. \end{cases}$$
2. If $A_n \neq \emptyset$ and if its last element u has a neighbor in S_n, the DFS goes to the smallest neighbor of u, say v, and we set:
$$\begin{cases} A_{n+1} &= A_n + v \\ S_{n+1} &= S_n \setminus \{v\}, \\ R_{n+1} &= R_n. \end{cases}$$
3. If $A_n \neq \emptyset$ and if its last element u has no neighbor in S_n, the DFS backtracks and we set:
$$\begin{cases} A_{n+1} &= A_n - u, \\ S_{n+1} &= S_n, \\ R_{n+1} &= R_n \cup \{u\}. \end{cases}$$

This algorithm explores the whole graph and provides a spanning tree of each connected component. In Section 6.4, we will provide an algorithm that construct simultaneously a random graph and a DFS on it.

The algorithm finishes after $2N$ steps. For every $0 \leq n \leq 2N$, we set $X_n := |A_n|$. This walk is called the *contour process* associated to the spanning forest of the DFS. In words, it is a ± 1 walk that starts at $X_0 = 0$, stays nonnegative and ends at $X_{2N} = 0$, which increases by 1 each time the DFS moves on (corresponding to point 1. or 2.) and decreases by one each time the DFS backtracks (corresponding to point 3.). Notice that $X_n = 0$ when the process starts the exploration of a new connected component. Therefore, each excursion of (X_n) corresponds to a connected component of G.

6.2.2 The Configuration model

We now turn to the definition of the configuration model.

Definition 2. *Let $\mathbf{d} = (d_1, \dots, d_N) \in \mathbb{Z}_+^N$. Let $\mathscr{C}(\mathbf{d})$ be a random multigraph whose law is uniform among all multigraphs with degree sequence $\mathbf{d}$ if $d_1 + \cdots d_N$ is even, and $(d_1, d_2, \dots, d_N + 1)$ otherwise.*

We will study sequence of configuration models whose associated sequence of empirical degree distribution converges to a given probability measure.

Definition 3. *Let $\boldsymbol{\pi}$ be a probability distribution on $\mathbb{Z}_+$. For every $N \geq 1$, let $\mathbf{d}^{(N)} = (d_1^{(N)}, \dots, d_N^{(N)}) \in \mathbb{Z}_+^N$. We say that $(\mathscr{C}(\mathbf{d}^{(N)}))_{N\geq 1}$ has asymptotic degree distribution $\boldsymbol{\pi}$ if*

$$\forall k \geq 0, \quad \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{d_i^{(N)}=k\}} \underset{N\to+\infty}{\longrightarrow} \boldsymbol{\pi}(\{k\}).$$

As observed in [MR95], the configuration model exhibits a phase transition for the existence of a unique macroscopic connected component. In this article, we will restrict our attention to supercritical configuration models, that is where this giant component exists.

Definition 4. *Let $\boldsymbol{\pi}$ be a probability distribution on $\mathbb{Z}_+$ such that $\sum_{k\geq 0} \boldsymbol{\pi}(\{k\})k^2 < \infty$ and denote by f_π its generating function. Let $\hat{\boldsymbol{\pi}}$ be the probability distribution having generating function*

$$\widehat{f}_\pi(s) := f_{\hat{\pi}}(s) = \frac{f'_\pi(s)}{f'_\pi(1)}.$$

We say that $\boldsymbol{\pi}$ is supercritical if $\widehat{f_\pi}'(1) > 1$. Notice that, denoting by D_π a random variable with law $\boldsymbol{\pi}$, it is equivalent to

$$\frac{\mathbb{E}[D_\pi(D_\pi - 1)]}{\mathbb{E}[D_\pi]} > 1.$$

In that case we define ρ_π to be the smallest positive solution of the equation

$$1 - \rho_\pi = \widehat{f}_\pi(1 - \rho_\pi).$$

Finally, we set

$$\xi_\pi := 1 - f_\pi(1 - \rho_\pi).$$

The number ρ_π is the probability that a Galton-Watson tree with distribution $\widehat{\boldsymbol{\pi}}$ is infinite, whereas the number ξ_π is the survival probability of a tree where the root has degree distribution $\boldsymbol{\pi}$ and individuals of the next generations have a number of children distributed according to $\widehat{\boldsymbol{\pi}}$. In this article, we study sequence of configuration models $\mathscr{C}(\mathbf{d}^{(N)})$ whose asymptotic degree distribution is a supercritical probability measure $\boldsymbol{\pi}$.

As shown in [MR95, MR98, JL09, BR15], in this context, denoting by $\mathrm{C}_1^{(N)}, \mathrm{C}_2^{(N)}, \dots$ the sequence of connected components of $\mathcal{C}(\mathbf{d}^{(N)})$ ordered by decreasing number of vertices,

- $\frac{|\mathrm{C}_1^{(N)}|}{N} \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} \rho_\pi$,
- $|\mathrm{C}_2^{(N)}| = \mathcal{O}_{\mathbb{P}}(\log N)$ i.e. $\lim\limits_{\kappa\to+\infty} \limsup\limits_{N\to+\infty} \mathbb{P}(|\mathrm{C}_2^{(N)}| \geq \kappa \log N) = 0$.

We finally make the two following technical assumptions:

- The following convergence holds:

$$\lim_{N\to+\infty} \frac{{d_1^{(N)}}^2 + \cdots + {d_N^{(N)}}^2}{N} = \sum_{k\geq 0} k^2 \pi(\{k\}). \tag{A1}$$

- There exists $\gamma > 2$ such that:

$$\max\left\{d_1^{(N)}, \ldots, d_N^{(N)}\right\} \leq N^{1/\gamma}. \tag{A2}$$

6.2.3 Main results

We now state our first result. Define $\alpha \geq 0$ and consider the graph induced by the sleeping vertices after having explored $\lfloor \alpha N \rfloor$ vertices when performing the DFS algorithm on a configuration model. It is clear that this induced graph is also a configuration model. The purpose of the following theorem is to identify its asymptotic degree distribution. It turns out this distribution only depends on α and on the initial degree distribution $\boldsymbol{\pi}$.

Theorem 10. *Let $\boldsymbol{\pi}$ be a probability measure on $\mathbb{Z}_+$ with generating series f and let $(\mathscr{C}(\mathbf{d}^{(N)}))_{N\geq 1}$ be a configuration model with supercritical asymptotic degree distribution $\boldsymbol{\pi}$. Assume* **(A1)** *and* **(A2)**.

Let α_c be the smallest positive solution of the equation

$$\frac{f_\pi''\left(f_\pi^{-1}(1-\alpha)\right)}{f_\pi'(1)} = 1.$$

For every $\alpha \in [0, \alpha_c]$, let $\boldsymbol{\pi}_\alpha$ be the probability distribution on $\mathbb{Z}_+$ with generating series

$$g(\alpha, s) = \frac{1}{1-\alpha} f_\pi\left(f_\pi^{-1}(1-\alpha) - (1-s)\frac{f_\pi'(f_\pi^{-1}(1-\alpha))}{f_\pi'(1)}\right),$$

and write $\tau^{(N)}(\alpha) = \inf\{k \geq 1 : |S_k^{(N)}| \leq (1-\alpha)N\}$. Then, conditionally on their degree sequence, the graphs induced by the vertices of $S^{(N)}_{\tau^{(N)}(\alpha)}$ inside $\mathscr{C}(\mathbf{d}^{(N)})$ have the law of configuration models with asymptotic degree distribution $\boldsymbol{\pi}_\alpha$.

Remark 11. *We consider α up to some constant α_c, which corresponds to the time when so many vertices have been visited that the remaining graph of sleeping vertices is subcritical.*

The second result concerns the asymptotic of the contour process $X_n = |A_n|$ of the plane forest constructed by the DFS on a configuration model.

Theorem 11. *Under the assumptions of Theorem 10, the following limit holds in probability for the topology of uniform convergence:*

$$\forall t \in [0,2], \quad \lim_{N\to\infty} \frac{X_{\lceil tN \rceil}}{N} = h(t),$$

where the function h is continuous on $[0,2]$, null on the interval $[2\xi_\pi, 2]$ and defined below on the interval $[0, 2\xi_\pi]$.

There exists an implicit function $\alpha(\rho)$ defined on $[0, \rho_\pi]$ such that $1-\rho = \widehat{g}(\alpha(\rho), 1-\rho)$ where, for any $\alpha \in [0, \alpha_c]$, the function $s \mapsto \widehat{g}(\alpha, s)$ is the size-biased version of $s \mapsto g(\alpha, s)$ defined in Theorem 10,

namely $\widehat{g}(\alpha, s) = \partial_s g(\alpha, s)/\partial_s g(\alpha, 1)$. *The graph* $(t, h(t))_{t\in[0,2\xi_\pi]}$ *can be divided into a first increasing part and a second decreasing part. These parts are respectively parametrized for* $\rho \in [0, \rho_\pi]$ *by :*

$$\begin{cases} x^{\uparrow}(\rho) & := (2-\rho)\,\alpha(\rho) - \int_\rho^{\rho_\pi} \alpha(u)\mathrm{d}u, \\ y^{\uparrow}(\rho) & := \rho\,\alpha(\rho) + \int_\rho^{\rho_\pi} \alpha(u)\mathrm{d}u, \end{cases}$$

for the increasing part and

$$\begin{cases} x^{\downarrow}(\rho) := x^{\uparrow}(\rho) + 2\,(1-\alpha(\rho))\left(1 - g\big(\alpha(\rho), 1-\rho\big)\right), \\ y^{\downarrow}(\rho) := y^{\uparrow}(\rho), \end{cases}$$

for the decreasing part.

A direct consequence of this result in the following.

Corollary 13. *Let* $\mathcal{H}_N$ *be the length of the longest simple path in a configuration model of size* N *with asymptotic distribution* $\boldsymbol{\pi}$ *satisfying hypothesis of Theorem 10. Then, with the notation of Theorem 11,*

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\frac{\mathcal{H}_N}{N} \geq y^{\uparrow}(0) - \varepsilon = \int_0^{\rho_\pi} \alpha(u)\mathrm{d}u - \varepsilon\right) \underset{N\to+\infty}{\longrightarrow} 1.$$

Remark 12. *Note that the formulas in Theorems 10 and 11 have a meaning when* $\boldsymbol{\pi}$ *has a first moment. Therefore, it is natural to expect that the restriction on the tail of* $\boldsymbol{\pi}$ *is only technical.*

6.3 Examples

In this section we provide explicit formulations of Theorems 10 and 11 for particular choices of the initial probability distribution $\boldsymbol{\pi}$.

6.3.1 Poisson distribution

Since the Erdős-Rényi model on N vertices with probability of connection c/N is contiguous to the configuration model on N vertices with sequence of degree $D_1^{(N)}, \ldots, D_N^{(N)}$ that are i.i.d. with Poisson law of parameter c, we can recover the result of Enriquez, Faraud and Ménard [EFM20]. Indeed, in the Erdős-Rényi case, after having explored a proportion α of vertices, the graph induced by the unexplored vertices is an Erdős-Rényi random graph with $(1-\alpha)N$ vertices and parameter c/N, hence its asymptotic degree distribution is Poisson with parameter $(1-\alpha)c$. This is in accordance with our Theorem 10 since in that case, denoting $f(s) = \exp(c(s-1))$ the generating series of the Poisson law with parameter c,

$$\begin{aligned} g(\alpha, s) &= \frac{1}{1-\alpha} f\left(f^{-1}(1-\alpha) - (1-s)\frac{f'(f^{-1}(1-\alpha))}{f'(1)}\right) \\ &= \frac{1}{1-\alpha}\exp\left(c\left(f^{-1}(1-\alpha) - (1-s)\frac{f'(f^{-1}(1-\alpha))}{f'(1)} - 1\right)\right) \\ &= \frac{1}{1-\alpha}\exp\left(c\left(1 + \frac{\log(1-\alpha)}{c} - (1-s)\frac{cf(f^{-1}(1-\alpha))}{c} - 1\right)\right) \\ &= \frac{1}{1-\alpha}\exp\left(c\left(1 + \frac{\log(1-\alpha)}{c} - (1-s)(1-\alpha) - 1\right)\right) \\ &= \exp\left(c(1-\alpha)(s-1)\right). \end{aligned}$$

Using the formulas of Theorem 11, we obtain the same equations as in [EFM20] for the limiting profile of the DFS spanning tree.

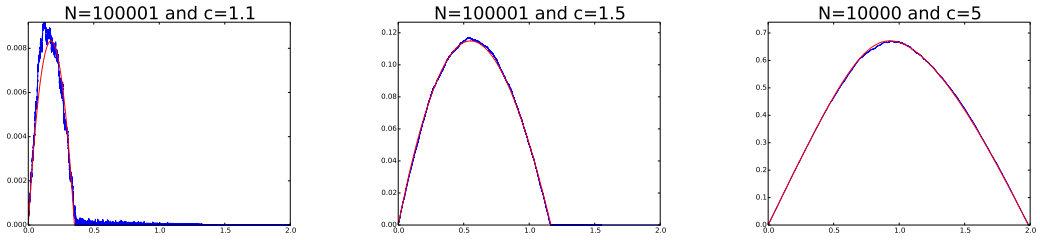


Figure 6.1 – Simulations of $(X_{\lceil tN \rceil}/N)_{t\in[0,2]}$ (blue) and the limiting shape (red) for various values of N and c. Notice that when c is close to 1, we have to take N very large for the walk to be close to its limit.

6.3.2 d-Regular and Binomial distributions

Let $d \geq 3$. Since the results of Theorem 10 and 11 hold with probability tending to 1, we can obtain results on d-regular uniform random graphs by applying them to the contiguous model which consists in choosing $\boldsymbol{\pi} = \delta_d$. By Theorem 10, the degree distribution $\boldsymbol{\pi}_\alpha$ has generating function

$$\begin{aligned} g(\alpha, s) &= \frac{1}{1-\alpha}\left((1-\alpha)^{1/d} - (1-s)\frac{d(1-\alpha)^{(d-1)/d}}{d}\right)^d \\ &= \left(1 + (s-1)(1-\alpha)^{\frac{d-2}{d}}\right)^d. \end{aligned} \tag{6.1}$$

Hence, $\boldsymbol{\pi}_\alpha$ is a binomial distribution $\mathrm{Bin}\left(d, (1-\alpha)^{\frac{d-2}{d}}\right)$. From (6.1), we get $\hat{g}(\alpha, s) = (1 + (s-1)(1-\alpha)^{(d-2)/d})^{d-1}$. Solving the equation $1-\rho = \hat{g}(\alpha, 1-\rho)$ in α gives:

$$\alpha(\rho) = 1 - \left(\frac{1-(1-\rho)^{\frac{1}{d-1}}}{\rho}\right)^{\frac{d}{d-2}}.$$

From this, we deduce a parametrization of the limiting profile in terms of hypergeometric functions. In particular, the height of the limiting DFS spanning tree is given by

$$H_{\max}(d) = 1 - \int_0^1 \left(\frac{1-x^{\frac{1}{d-1}}}{1-x}\right)^{\frac{d}{d-2}} \mathrm{d}x.$$

When $\boldsymbol{\pi}$ has a binomial distribution with parameters d and p, $\boldsymbol{\pi}_\alpha$ is also a binomial distribution.

$$\boldsymbol{\pi}_\alpha = \mathrm{Bin}\left(d, p(1-\alpha)^{\frac{d-2}{d}}\right).$$

6.3.3 Geometric distribution

Let $p > 0$ and suppose that the initial distribution $\boldsymbol{\pi}$ is a geometric distribution starting at 0 with parameter p. The generating series of $\boldsymbol{\pi}$ is $f(s) = \frac{p}{1-(1-p)s}$. We assume $p < 2/3$ so that the configuration model with asymptotic degree distribution $\boldsymbol{\pi}$ has a giant component. Then, by Theorem 10, the distribution $\boldsymbol{\pi}_\alpha$ has generating series

$$g(\alpha, s) = \frac{p(\alpha)}{1-(1-p(\alpha))s},$$

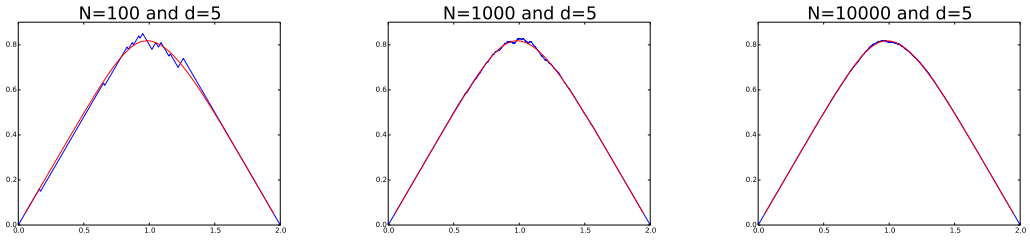


Figure 6.2 – Simulations of $(X_{\lceil tN \rceil}/N)_{t\in[0,2]}$ (blue) and the limiting shape (red) for 5-regular graphs of various sizes.

where $p(\alpha) = \frac{p}{p+(1-p)(1-\alpha)^3}$. Hence, $\boldsymbol{\pi}_\alpha$ is a geometric distribution that starts at 0 with parameter $p(\alpha)$. The generating series of $\hat{\boldsymbol{\pi}}_\alpha$ is $\hat{g}(\alpha, s) = \left(\frac{p(\alpha)}{1-(1-p(\alpha))s}\right)^2$. Therefore, the solution in α of $1-\rho = \hat{g}(\alpha, 1-\rho)$ is

$$\alpha(\rho) = 1 - \left(\frac{p}{1-p}\right)^{1/3} \left(\frac{1}{1-\rho+\sqrt{1-\rho}}\right)^{1/3}.$$

In particular, the height of the limiting DFS spanning tree is given by:

$$H_{\max}(p) = \rho_{\boldsymbol{\pi}} - \left(\frac{p}{1-p}\right)^{1/3} \int_0^{\rho_{\boldsymbol{\pi}}} \left(\frac{1}{x+\sqrt{x}}\right)^{1/3} \mathrm{d}x,$$

where $\rho_{\boldsymbol{\pi}}$ is given by:

$$\rho_{\boldsymbol{\pi}} = \frac{1}{2}\left(\frac{1-3p}{1-p} + \sqrt{\frac{1+3p}{1-p}}\right).$$

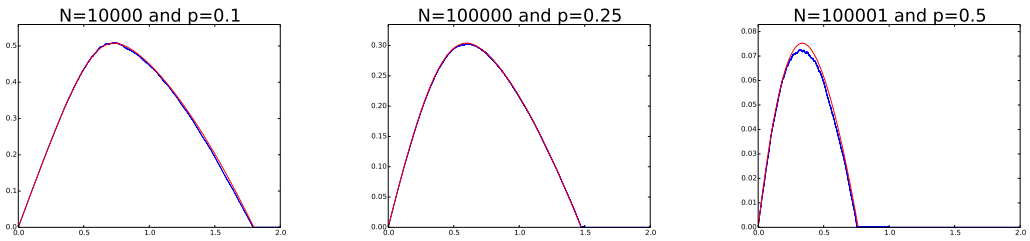


Figure 6.3 – Simulations of $(X_{\lceil tN \rceil}/N)_{t\in[0,2]}$ (blue) and the limiting shape (red) for random graphs with geometric degrees with various perimeters.

6.3.4 Heavy tailed distribution

When $\boldsymbol{\pi}$ is a power law distribution of parameter $\gamma > 2$, that is when $\boldsymbol{\pi}(\{k, k+1, \ldots\}) \sim C/k^\gamma$ for a constant C, only the first $\lfloor \gamma \rfloor$ moments of $\boldsymbol{\pi}$ are finite. Let $\alpha \in (0, \alpha_c)$. Then, for all $n \geq 0$,

the n-th factorial moment of π_α is equal to

$$\pi_\alpha(x^n) = \frac{\partial^n}{\partial s^n}\bigg|_{s=1} g(\alpha, s)$$
$$= \left(\frac{f'_\pi\left(f_\pi^{-1}(1-\alpha)\right)}{f'_\pi(1)}\right)^n \frac{f^{(n)}\left(f^{-1}(1-\alpha)\right)}{1-\alpha}.$$

Therefore, after visiting a proportion εN of the vertices in the DFS, the asymptotic distribution of the degrees of the graph induced by the unexplored vertices is not a power law and has moments of all orders. This remarkable phenomenon could be explained by the fact that vertices of high degree are visited in a microscopic time. We believe that a precise study of this case could be of independent interest.

6.4 Constructing while exploring

Let $(\mathbf{d}^{(N)})_{N\geq 1}$ be a sequence of degree sequences of increasing length satisfying the assumptions of Theorem 10. For a fixed $N \geq 1$, we use the sequence $\mathbf{d}^{(N)} = (d_1^{(N)}, \dots, d_N^{(N)})$ to construct a configuration model $\mathscr{C}(\mathbf{d}^{(N)})$ with vertex set $\mathrm{V}_N = \{1, \dots, N\}$. More precisely, we simultaneously build the graph and its DFS exploration. This will be done in a similar way as for the DFS defined in Section 6.2.1, while revealing as little information about the unexplored part of the graph as possible. For every step n we consider the following objects, defined by induction.

- A_n, the active vertices, is an ordered list of pairs $(v, \mathbf{m}_v)$ where v is a vertex of V_N and $\mathbf{m}_v$ is the list of vertices corresponding to the vertices that will be **m**atched to v during the rest of the exploration.
- S_n, the sleeping vertices, is a subset of V_N. This subset will never contain a vertex of A_n.
- R_n, the retired vertices, is another subset of V_N composed of all the vertices that are neither in A_n nor S_n.

At time $n = 0$, choose a vertex v uniformly at random and pair each of its $d_v^{(N)}$ half edges to a half edge of the graph. This gives an unordered set of vertices that will be matched to v at some point of the exploration. We denote by $\mathbf{m}_v$ this set with a uniform order. Set:

$$\begin{cases} A_0 &= ((v, \mathbf{m}_v)), \\ S_0 &= \mathrm{V}_N \setminus \{v\}, \\ R_0 &= \emptyset. \end{cases}$$

Suppose that A_n, S_n and R_n have already been constructed. Three cases are possible.

1. If $A_n = \emptyset$, the algorithm has just finished exploring and building a connected component of $\mathscr{C}(\boldsymbol{d}^{(N)})$. In that case, we pick a vertex v_{n+1} uniformly at random from S_n and we pair each of its $d_{v_{n+1}}^{(N)}$ half edges to a uniform half edge belonging to a vertex of S_n. We denote by $\mathbf{m}_{v_{n+1}}$ the set of these paired vertices which are different from v_{n+1} (corresponding to loops in the graph), ordered uniformly and set:

$$\begin{cases} A_{n+1} &= (v_{n+1}, \mathbf{m}_{v_{n+1}}), \\ S_{n+1} &= S_n \setminus \{v_{n+1}\}, \\ R_{n+1} &= R_n. \end{cases}$$

2. If $A_n \neq \emptyset$ and if its last element $(u, \mathbf{m}_u)$ is such that $\mathbf{m}_u = \emptyset$, the DFS backtracks and we set:
$$\begin{cases} A_{n+1} &= A_n - (u, \mathbf{m}_u), \\ S_{n+1} &= S_n, \\ R_{n+1} &= R_n \cup \{u\}. \end{cases}$$

3. If $A_n \neq \emptyset$ and if its last element $(u, \mathbf{m}_u)$ is such that $\mathbf{m}_u \neq \emptyset$, the algorithm goes to the first vertex of $\mathbf{m}_u$, say v_{n+1}. By construction, this vertex always belongs to S_n. We first update A_n into A'_n by removing each occurrence of v_{n+1} in the lists $\mathbf{m}_x$ for $x \in A_n$. The half edges of v_{n+1} that have not been matched up to now are uniformly matched with half edges of S_n that have not yet been matched. We order the set of corresponding vertices that v_{n+1} itself uniformly and denote $\mathbf{m}_{v_{n+1}}$ this list. We finally set
$$\begin{cases} A_{n+1} &= A'_n + (v_{n+1}, \mathbf{m}_{v_{n+1}}) \\ S_{n+1} &= S_n \setminus \{v_{n+1}\}, \\ R_{n+1} &= R_n. \end{cases}$$

Since each matching of half-edges in the algorithm is uniform, it indeed constructs a random graph $\mathscr{C}(\boldsymbol{d}^{(N)})$. Moreover, as advertised at the end of Section 6.2.1, this algorithm simultaneously constructs the DFS on this random graph as each of the three cases are in correspondence to the same three cases in the definition of the DFS given in Section 6.2.1.

From this construction, it is clear that for every n, the graph induced by S_n in the whole graph is a configuration model. Moreover, for each vertex v of S_n, its degree in this induced graph is given by its initial degree $d_v^{(N)}$ minus the number of times that v appears in the lists $\mathbf{m}_x$ for $x \in A_n$.

In order to prove Theorem 1, we will first analyse the part of the algorithm corresponding to the increasing part of the limiting profile. This has the same law as the increasing part of the process $(X_n)_{0 \leq n \leq 2N}$. During this first phase, at each time, the graph induced by the sleeping vertices, which we will call the remaining graph, is a supercritical configuration model. We will see in Section 6.4.1 that there is a sequence of random times where the DFS discovers a vertex belonging to what will turn out to be the giant component of the remaining graph. We will call these times ladder times and study in detail the law of the remaining graph at these times in Section 6.4.2.

6.4.1 Ladder times

Fix $\delta \in (0,1)$. Let $T_0 = 0$ and define, for $k \in \{0, \dots, K\}$,
$$T_{k+1} := \min \left\{ i > T_k,\ X_i = k+1 \text{ and } \forall i \leq j \leq i + N^\delta,\ X_j \geq k+1 \right\},$$
where K is the last index for which this definition makes sense (i.e. the set for which the min is taken is not empty). Of course, this sequence of times will only be useful to analyse the DFS on $\mathscr{C}(\mathbf{d}^{(N)})$ when K is of macroscopic order, which is indeed the case with high probability under the assumptions of Theorem 10.

For all $k \in \{0, \dots, K\}$, let $\mathscr{S}_k$ be the graph induced by the vertices of $S_{T_k - 1}$ in the graph constructed by the algorithm of the previous section. We also denote by v_k the last vertex of A_{T_k}.

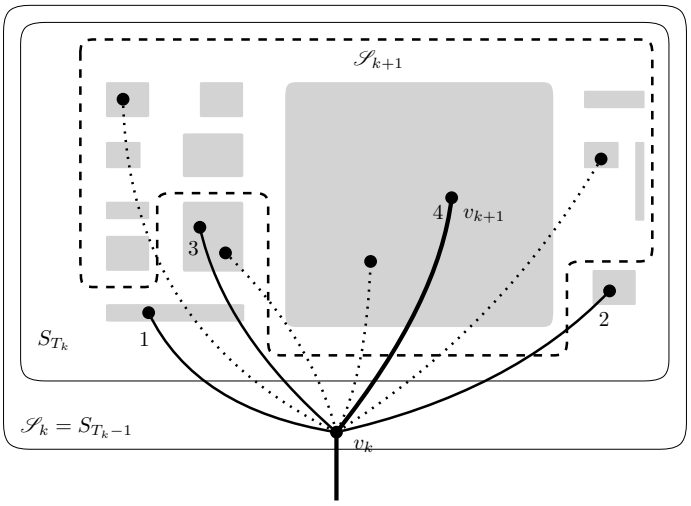


Figure 6.4 – Structure of the remaining graph at a ladder time. The first half edges of v_k are numbered according to their matching order during the construction. Here, the last matched half edge is in bold and connect v_k to v_{k+1}. The remaing half edges of v_k are represented with dotted line and matched to unexplored vertices.

The graphs $\mathscr{S}_k$ and S_{T_k} have the same vertex set except for v_k which belongs to $\mathscr{S}_k$ but not to S_{T_k}. See Figure 6.4 for an illustration of these definitions. We chose to emphasize $\mathscr{S}_k$ because the structural changes between two such consecutive graphs will be easier to track.

Fix $k < K$. From the definition of the times T_k and T_{k+1}, we can deduce that v_{k+1} and v_k are neighbors in $\mathscr{S}_k$. Between the times $n = T_k$ and $n = T_{k+1}$ the process $X_n = |A_n|$ stays above k and is equal to k at time $T_{k+1} - 1$. Each excursion of X_n strictly above k between T_k and $T_{k+1} - 1$ corresponds to the exploration of a different connected component of $\mathscr{S}_k \setminus \{v_k\}$ and we have

$$T_{k+1} - T_k = 1 + 2 \times (\text{ number of vertices in } \mathscr{S}_k \setminus \mathscr{S}_{k+1} - 1).$$

In addition, the definition of the ladder times implies that these connected components have sizes smaller than N^δ.

For every $n \in \{0, \dots, 2N\}$, let $D_n^{(N)}$ be the degree of a uniform vertex in the graph induced by S_n. For every $\varepsilon > 0$, we define

$$n_\varepsilon = n_\varepsilon^{(N)} = \sup \left\{ n \in [\![0, 2N]\!] : \forall m \in [\![0, n]\!], \frac{\mathbb{E}[D_n^{(N)}(D_n^{(N)} - 1)]}{\mathbb{E}[D_n^{(N)}]} > 1 + \varepsilon \right\}.$$

For $n < n_\varepsilon$, the subgraphs induced by S_n are all supercritical. For $0 < \delta < 1/2$, let $\mathbf{G}_\varepsilon = \mathbf{G}_\varepsilon^{(N)}(\delta)$ be the event that, for all $n < n_\varepsilon$,

- there is at least one connected component with size greater than $N^{1-\delta}$ in the graph induced by S_n;
- there is no connected component of size between N^δ and $N^{1-\delta}$ in the graph induced by S_n.

Under the assumptions of Theorem 10 we have, for every $\lambda > 0$,

$$\mathbb{P}(\mathbf{G}_\varepsilon) = 1 - \mathcal{O}(N^{-\lambda}). \tag{6.2}$$

The event $\mathbf{G}_\varepsilon$ will be instrumental in the analysis of the DFS and the times T_k because, on this event, if $T_k < n_\varepsilon$, then the graph $S_{T_k} = \mathscr{S}_k \setminus \{v_k\}$ has a connected component of size larger than $N^{1-\delta}$ and, in $\mathscr{S}_k$, the vertex v_k has a neighbor in this giant component. Indeed, if every neighbor of v_k in S_{T_k} belonged to a small component, the size of the connected component of v_k in $\mathscr{S}_k$ would be at most $N^{1/\gamma}\, N^\delta \ll N^{1-\delta}$. On the other hand, we know that this component has size larger than N^δ meaning that, on $\mathbf{G}_\varepsilon$, it is in fact larger than $N^{1-\delta}$ leading to a contradiction. By induction, this means that on $\mathbf{G}_\varepsilon$ and if $T_k < n_\varepsilon$, then $k < K$.

Let us finally set

$$K_\varepsilon := \sup\{k \in [\![0, K]\!],\ T_k < n_\varepsilon\},$$

and note that, thanks to (6.2), $K_\varepsilon < K$ with probability $1 - \mathcal{O}(N^{-\lambda})$.

6.4.2 Analysis of the graphs $\mathscr{S}_k$

Let $N_i(k)$ be the number of vertices of degree i in $\mathscr{S}_k$. The graph $\mathscr{S}_k$ has the law of a configuration model with vertex degrees given by the sequence $(N_i(k))_{i\geq 0}$. Recalling that $\mathbf{m}_{v_k}$ denotes the list of neighbors of of v_k in $\mathscr{S}_k$ (self-loops not included), the evolution of N_i is given by:

$$N_i(k+1) - N_i(k) = - V_i(\mathscr{S}_k \setminus \mathscr{S}_{k+1}) \tag{6.3}$$

$$+ \sum_{v \in \mathbf{m}_{v_k} \cap \mathscr{S}_{k+1}} \left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(v)=i} + \mathbf{1}_{\deg_{\mathscr{S}_k}(v)=i+1} \right), \tag{6.4}$$

where $V_i(S)$ stands for the number of vertices with degree i in the graph S and, if H is a subgraph of S, $S \setminus H$ is the subgraph of S induced by its vertices that do not belong to H. Indeed, the first contribution corresponds to the complete removal of vertices belonging to $\mathscr{S}_k$ but not $\mathscr{S}_{k+1}$. The second contribution corresponds to edges of $\mathscr{S}_k$ connecting v_k to vertices of $\mathscr{S}_{k+1}$. Figure 6.4 gives an illustration of this situation. In this figure, the contribution (6.3) comes from the connected components of the vertices attaches to the half edges of v_k numbered 1, 2 and 3. The contribution (6.4) comes from v_{k+1} and the vertices matched to dotted half edges.

A fundamental step in understanding the behaviour of the exploration process is to identify the asymptotic behaviour of the variables T_k and $N_i(k)$ for large N. This is the object of Theorem 12. To state this, we first introduce some technical notation.

Let $(z_i)_{i\geq 0} \in \mathbb{R}^{\mathbb{Z}_+}$ be such that $\sum_{i\geq 0} z_i \leq 1$ and $\sum_{k\geq 0} i z_i < \infty$. For any $i \geq 0$ let $\hat{z}_i = (i+1)z_i / \sum_j j z_j$ and define:

$$\begin{cases} g_{(z_i)_{i\geq 0}}(s) &= \sum_{i\geq 0} \frac{z_i}{\sum_{l\geq 0} z_l} s^i \\ \hat{g}_{(z_i)_{i\geq 0}}(s) &= \sum_{i\geq 0} \hat{z}_i s^i = \frac{g'_{(z_i)_{i\geq 0}}(s)}{g'_{(z_i)_{i\geq 0}}(1)} \end{cases} \tag{6.5}$$

respectively the generating series associated to $(z_k)_{k\geq 0}$ and its sized-biased version. Let also $\rho_{(z_i)_{i\geq 0}}$ be the largest solution in $[0,1]$ of

$$1 - s = \hat{g}_{(z_i)_{i\geq 0}}(1-s). \tag{6.6}$$

Remark 13. *Since $\hat{g}$ is the generating function of a probability distribution on the integers, it is convex on $[0,1]$. Therefore, Equation (6.6) has a positive solution in $(0,1]$ if and only if $\hat{g}'(1) > 1$, which is equivalent to $\frac{\sum_{l\geq 1}(l-1)lz_l}{\sum_{l\geq 1} l z_l} > 1$.*

We also define the following functions:

$$f(z_0, z_1, \ldots) = \frac{2 - \rho_{(z_i)_{i\geq 0}}}{\rho_{(z_i)_{i\geq 0}}} \tag{6.7}$$

$$f_i(z_0, z_1, \ldots) = -\frac{1}{\rho_{(z_j)_{j\geq 0}}} \frac{iz_i}{\sum_{j\geq 0} jz_j} + \frac{1}{\rho_{(z_j)_{j\geq 0}}} \left(1 - \frac{\sum_{j\geq 0}(j-1)jz_j}{\sum_{n\geq 0} jz_j}\right) \left(\frac{iz_i}{\sum_{j\geq 0} jz_j} - \frac{(i+1)z_{i+1}}{\sum_{j\geq 0} jz_j}\right). \tag{6.8}$$

The asymptotic behaviour of the variables T_k and $N_i(k)$ will be driven by the solution of an infinite system of differential equations whose uniqueness and existence is provided by the following lemma, whose proof is postponed to Section 6.6.2.

Lemma 3. *Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_i)_{i\geq 0} \in [0,1]^{\mathbb{N}}$ such that $\sum_{i\geq 0} \boldsymbol{\pi}_i = 1$. Then, the following system of differential equations has a unique solution which is well defined on $[0, t_{\max})$ for some $t_{\max} > 0$:*

$$\begin{cases} \frac{\mathrm{d}z_i}{\mathrm{d}t} &= f_i(z_0, z_1, \ldots); \\ z_i(0) &= \boldsymbol{\pi}_i. \end{cases} \tag{S}$$

We are now ready to state the main result of this section.

Theorem 12. *With high probability, for all $k \leq K_\varepsilon$:*

$$T_k = Nz\left(\frac{k}{N}\right) + o(N)$$
$$N_i(k) = Nz_i\left(\frac{k}{N}\right) + o(N),$$

where $(z_0, z_1, \ldots)$ is the unique solution of (S) and z is the unique solution of $\frac{\mathrm{d}z}{\mathrm{d}t} = f(z_0, z_1, \ldots)$ with initial condition given by $z(0) = 0$.

Proof. Our main tool to prove this result is Corollary 14, which is stated and proved in the Appendix. This corollary is a version of a result of Wormald [Wor95] tailored for our purpose. To apply this result we need to check the following two points:

1. There exists $0 < \beta < 1/2$ such that with high probability for all $k \leq K_\varepsilon$,
$$|T_{k+1} - T_k| \leq N^\beta \text{ and for all } k \geq 0,\ |N_i(k+1) - N_i(k)| \leq N^\beta.$$
2. We denote by $(\mathcal{F}_k)_{k\geq 0}$ the canonical filtration associated to the sequence $((N_i(k))_{i\geq 0})_{k\geq 0}$. There exists $\lambda > 0$ such that for every $k \leq K_\varepsilon$,
$$\mathbb{E}[T_{k+1} - T_k \,|\, \mathcal{F}_k] = f\left(\frac{N_0(k)}{N}, \frac{N_1(k)}{N}, \ldots\right) + O\left(N^{-\lambda}\right),$$
$$\mathbb{E}[N_i(k+1) - N_i(k) \,|\, \mathcal{F}_k] = f_i\left(\frac{N_0(k)}{N}, \frac{N_1(k)}{N}, \ldots\right) + O\left(N^{-\lambda}\right).$$

The first point is a consequence of Equation (6.2) with $\delta < 1/2 - 1/\gamma$. Indeed on the event $\mathbf{G}_\varepsilon$ the vertices v_k have degree at most $N^{1/\gamma}$ and therefore $T_{k+1} - T_k \leq 1 + 2N^{1/\gamma}N^\delta \ll N^\beta$ for some $\beta < 1/2$. Since $|N_i(k+1) - N_i(k)| \leq (T_{k+1} - T_k)/2$ the second inequality is trivial.

In order to establish the second point, we need to analyse the structure of $\mathscr{S}_k$ and the contributions (6.3) and (6.4). To this end, we will study the random variable $\mathfrak{e}_k$ that counts the

number of excursions strictly above k of the walker (X_n) coding the DFS between the times T_k and $T_{k+1}-1$ (in Figure 6.4, $\mathfrak{e}_k = 3$). In particular, the expectation of $\mathfrak{e}_k$ conditionally on $\mathcal{F}_k$ is well defined on the event $\mathbf{G}_\varepsilon$.

If we disconnect the edges joining the $\mathfrak{e}_k$ first children of v_k in the tree constructed by the DFS, the remaining connected components in $\mathscr{S}_k$ of these children have size smaller than N^δ. This motivates the following notation:

- for every $i \geq 0$, let $\mathbf{Ext}_i^k$ (resp. $\mathbf{Surv}_i^k$) be the set of half-edges $e \in \mathscr{S}_k$ connected to a vertex w of degree i (in $\mathscr{S}_k$) such that the the connected component of w after removing this half-edge has size smaller than N^δ (resp. larger than N^δ);
- let $\mathbf{Ext}^k$ (resp. $\mathbf{Surv}^k$) be the set of half-edges $e \in \mathscr{S}_k$ connected to a vertex w such that the connected component of w after removing this half-edge has size smaller than N^δ (resp. larger than N^δ). Note that $\mathbf{Ext}^k = \sqcup_{j\geq 0}\mathbf{Ext}_j^k$ and $\mathbf{Surv}^k = \sqcup_{j\geq 0}\mathbf{Surv}_j^k$.

Recall that on $\mathbf{G}_\varepsilon$, for all $k \leq K_\varepsilon$, v_k has a neighbor in $\mathscr{S}_k$ that belongs to a connected component of $\mathscr{S}_k$ with more than N^δ vertices. This means that for every such k, with probability $1-\mathcal{O}(N^{-1-\lambda})$, the random variable $\mathfrak{e}_k$ is the number of half edges of $\mathbf{Ext}^k$ attached to v_k before attaching a half edge of $\mathbf{Surv}^k$ during the DFS. In order to compute its expectation, we first condition on $\{\deg_{\mathscr{S}_k}(v_k) = d\}$, with $d > 0$ fixed.

Notice that, conditional on the event $\{\deg_{\mathscr{S}_k}(v_k) = d\} \cap \{\mathfrak{e}_k < \deg_{\mathscr{S}_k}(v_k)\}$, the law of $(\mathscr{S}_k, v_k)$ is the law of a rooted configuration model $\mathbf{C}^d_{\mathbf{N}(k)}$ with root degree d and degree sequence $\mathbf{N}(k) := (N_i(k))_{i\geq 0}$, conditioned on the root having one of its half-edges paired to an element of $\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})$. We define the new random variable $\tilde{\mathfrak{e}}_k$ as the number of half edges of the root paired to an element of $\mathbf{Ext}(\mathbf{C}^d_{\mathbf{N}(k)})$ before pairing a half edge to an element of $\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})$ when doing successive uniform matching in the configuration model (with the convention $\tilde{\mathfrak{e}}_k = d$ if the root has no half-edged paired to an element of $\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})$). We have the following equality for all j:

$$\mathbb{P}\big(\mathfrak{e}_k = j \,\big|\, \mathcal{F}_k \text{ and } \deg_{\mathscr{S}_k}(v_k) = d\big) = \mathbb{P}\big(\tilde{\mathfrak{e}}_k = j \,\big|\, \tilde{\mathfrak{e}}_k < d\big) + \mathcal{O}(N^{-1-\lambda}).$$

Let

$$\tilde{\rho}_k := \frac{|\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} = 1 - \frac{|\mathbf{Ext}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|},$$

the proportion of half-edges in $\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})$ (resp. $\mathbf{Ext}(\mathbf{C}^d_{\mathbf{N}(k)})$). This proportion is close to a constant ρ_k that we now define with the help of additional notation. Recalling (6.5), let

$$p_i = p_i(k) = \tfrac{N_i(k)}{\sum_{j\geq 0} N_j(k)}, \qquad g_k = g_{(p_j)_{j\geq 0}},$$
$$\hat{p}_i = \hat{p}_i(k) = \tfrac{(i+1)p_{i+1}(k)}{\sum_{j\geq 0} j p_j(k)}, \qquad \hat{g}_k = \hat{g}_{(p_j)_{j\geq 0}} = g_{(\hat{p}_j)_{j\geq 0}},$$

and let $\rho_k = \rho_{(p_j(k))_{j\geq 0}}$ be the largest solution in $[0,1]$ of $1-s = \hat{g}_k(1-s)$. We have the following lemma, whose proof is postponed to Section 6.6.1.

Lemma 4. *For all $0 \leq k \leq K_\varepsilon$, there exists $\lambda > 0$ and $\eta > 0$ such that, conditionally on $\mathcal{F}_k$, uniformly in k,*

$$\begin{cases} \mathbb{P}\left(\left|\frac{|\mathbf{Ext}_i(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - \frac{ip_i}{g_k'(1)}(1-\rho_k)^{i-1}\right| \geq N^{-\lambda}\right) = \mathcal{O}\left(N^{-1-\lambda}\right), \\ \mathbb{P}\left(\left|\frac{|\mathbf{Surv}_i(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - \frac{ip_i}{g_k'(1)}(1-(1-\rho_k)^{i-1})\right| \geq N^{-\lambda}\right) = \mathcal{O}\left(N^{-1-\lambda}\right), \\ \mathbb{P}\left(\left|\frac{|\mathbf{Ext}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - (1-\rho_k)\right| \geq N^{-\lambda}\right) = \mathcal{O}\left(N^{-1-\lambda}\right), \\ \mathbb{P}\left(\left|\frac{|\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - \rho_k\right| \geq N^{-\lambda}\right) = \mathcal{O}\left(N^{-1-\lambda}\right). \end{cases}$$

Using this lemma, we obtain:

$$\mathbb{P}\left(\mathfrak{e}_k = j \,\middle|\, \mathcal{F}_k \text{ and } \deg_{\mathscr{S}_k}(v_k) = d\right) = \frac{\mathbb{P}\left(\{\tilde{\mathfrak{e}}_k = j\} \cap \{\tilde{\mathfrak{e}}_k < d\}\} \cap \{|\tilde{\rho}_k - \rho_k| \leq \mathcal{O}(N^{-\lambda})\}\right)}{\mathbb{P}\left(\tilde{\mathfrak{e}}_k < d \cap \{|\tilde{\rho}_k - \rho_k| \leq \mathcal{O}(N^{-\lambda})\}\right)} + \mathcal{O}\left(N^{-1-\lambda}\right). \tag{6.9}$$

Fix $j < d$. To estimate the probabilities in (6.9), we successively match the half edges $c_1, \dots, c_{j+1}$ of the root uniformly among the half edges of $\mathbf{C}^d_{\mathbf{N}(k)}$. Notice that if none of these half edges are matched together, this is equivalent to an urn model without replacement. At each of these steps, the proportion of available half edges of $\mathbf{Ext}(\mathbf{C}^d_{\mathbf{N}(k)})$ diminishes and is therefore between $1 - \widetilde{\rho}_k - \frac{d}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|}$ and $1 - \widetilde{\rho}_k$. Recalling that $|E(\mathbf{C}^d_{\mathbf{N}(k)})|$ is uniformly of order N, we can write for every $j < d$

$$\begin{aligned}
&\frac{\left(1 - \rho_k - C\frac{d}{N} + \mathcal{O}\left(N^{-\lambda}\right)\right)^j \left(\rho_k + \mathcal{O}\left(N^{-\lambda}\right)\right)}{1 - \left(1 - \rho_k - C\frac{d}{N} + \mathcal{O}\left(N^{-\lambda}\right)\right)^d} + \mathcal{O}\left(N^{-1-\lambda}\right) \\
&\qquad \leq \mathbb{P}\left(\mathfrak{e}_k = j \,\middle|\, \mathcal{F}_k \text{ and } \deg_{\mathscr{S}_k}(v_k) = d\right) \\
&\qquad \leq \frac{\left(1 - \rho_k + \mathcal{O}\left(N^{-\lambda}\right)\right)^j \left(\rho_k + C\frac{d}{N} + \mathcal{O}\left(N^{-\lambda}\right)\right)}{1 - \left(1 - \rho_k + \mathcal{O}\left(N^{-\lambda}\right)\right)^d} + \mathcal{O}\left(N^{-1-\lambda}\right).
\end{aligned}$$

where C is a constant and the error terms $\mathcal{O}(N^{-\lambda})$ are the same everywhere and uniform in d. This easily translates into

$$\mathbb{P}\left(\mathfrak{e}_k = j \,\middle|\, \mathcal{F}_k \text{ and } \deg(v_k) = d\right) = \frac{(1-\rho_k)^j \rho_k}{1 - (1-\rho_k)^d}\left(1 + \mathcal{O}\left(d^2 N^{-1} + dN^{-\lambda}\right)\right)\mathbf{1}_{\{j<d\}} + \mathcal{O}\left(N^{-1-\lambda}\right)$$

where, once again, the error terms are uniform. We can now compute the conditional expectation of $\mathfrak{e}_k$:

$$\begin{aligned}
&\mathbb{E}\left[\mathfrak{e}_k \middle| \mathcal{F}_k, \deg_{\mathscr{S}_k}(v_k) = d\right] \\
&\quad = \frac{1-\rho_k}{\rho_k\left(1 - (1-\rho_k)^d\right)}\left(-d\rho_k(1-\rho_k)^{d-1} + 1 - (1-\rho_k)^d\right)\left(1 + \mathcal{O}\left(d^2N^{-1} + dN^{-\lambda}\right)\right) \\
&\qquad\qquad + \mathcal{O}(N^{-\lambda}),
\end{aligned}$$

where the last error term comes from the fact the $\mathfrak{e}_k$ is smaller that $\mathcal{O}(N)$ by definition.

To finally compute the expectation of $\mathfrak{e}_k$, we want to sum the above equality with respect to the law of $\deg_{\mathscr{S}_k}(v_k)$. By construction, in $\mathscr{S}_{k-1}$, the vertex v_k is attached to v_{k-1} by a half edge of $\mathbf{Surv}^{k-1}$ chosen uniformly. Therefore, by Lemma 4, the law of the degree of v_k in $\mathscr{S}_k$ is given by

$$\mathbb{P}(\deg_{\mathscr{S}_k}(v_k) = d \,|\, \mathcal{F}_k) = \frac{(d+1)p_{d+1}(k-1)}{\rho_{k-1}\mathcal{G}'_{k-1}(1)}\left(1 - (1-\rho_{k-1})^d\right)(1 + \mathcal{O}(N^{-\lambda})),$$

where the error term is uniform in d and k. We can replace $k-1$ by k in the above probabilities at the cost of a factor $1 + \mathcal{O}(N^{-\lambda})$ which is uniform in k and d. Indeed, on $\mathbf{G}_\varepsilon$, the difference between $\mathscr{S}_{k-1}$ and $\mathscr{S}_k$ consists of at most $N^{1/\gamma}$ components of size at most N^δ and we have

$p_d(k-1) = p_d(k)\left(1+\mathcal{O}(N^{1/\gamma+\delta-1})\right)$ uniformly in k and d. The difference between ρ_{k-1} and ρ_k is then of the same order by a Taylor expansion. Therefore

$$\mathbb{P}(\deg_{\mathscr{S}_k}(v_k) = d \,|\, \mathcal{F}_k) = \frac{(d+1)p_{d+1}(k)}{\rho_k g_k'(1)}\left(1-(1-\rho_k)^d\right)(1+\mathcal{O}(N^{-\lambda})), \tag{6.10}$$

and we get:

$$\begin{aligned}
&\mathbb{E}\left[\mathfrak{e}_k|\mathcal{F}_k\right]\\
&= \frac{(1-\rho_k)}{g_k'(1)\rho_k^2}\sum_{d\geq 0}(d+1)p_{d+1}(k)\left(-d\rho_k(1-\rho_k)^{d-1}+1-(1-\rho_k)^d\right)\left(1+\mathcal{O}\left(d^2N^{-1}+dN^{-\lambda}\right)\right)\\
&\qquad\qquad + \mathcal{O}(N^{-\lambda})\\
&= \frac{(1-\rho_n)}{g_n'(1)\rho_n^2}\left(g_n'(1)-\rho_n g_n''(1-\rho_n)-g_n'(1-\rho_n)\right)+\mathcal{O}(N^{\frac{1}{\gamma}-1})\cdot\mathcal{O}\left(\sum_{d\geq 0}d^2p_d(k)\right)+\mathcal{O}(N^{-\lambda}).
\end{aligned}$$

Notice that the error $\mathcal{O}(N^{-\lambda})$ is uniform in k and d. Let us prove that $\sum_{d\geq 0}d^2p_d(k)$ is of order 1. First note that it is of the same order as $\frac{1}{N}\sum_{d\geq 0}d^2N_d(k)$, where we recall that $N_d(k)$ is the number of vertices of degree d in $\mathscr{S}_k$. Indeed the number of vertices of $\mathscr{S}_k$ is of order N. Denoting by $N_{\geq d}(k)$ the number of vertices of degree larger than d in $\mathscr{S}_k$, it holds that $N_{\geq d}(k)\geq N_{\geq d}(k+1)$ from the definition of the algorithm. This monotonicity implies that

$$\frac{1}{N}\sum_{d\geq 0}d^2N_d(k) \leq \sum_{d\geq 0}d^2\frac{N_d(0)}{N},$$

where the right-hand side converges to a finite limit by assumption (**A2**). Therefore

$$\begin{aligned}
\mathbb{E}\left[\mathfrak{e}_k|\mathcal{F}_k\right] &= \frac{(1-\rho_k)}{g_k'(1)\rho_k^2}\left(g_k'(1)-\rho_k g_n''(1-\rho_k)-g_k'(1-\rho_k)\right)+\mathcal{O}(N^{-\lambda})\\
&= \frac{1-\rho_k}{\rho_k}\left(1-\hat{g}_k'(1-\rho_k)\right)+\mathcal{O}(N^{-\lambda}),
\end{aligned} \tag{6.11}$$

where we used $1-\rho_k = \hat{g}_k(1-\rho_k) = g_k'(1-\rho_k)/g_k'(1)$.

Now that we know more about the random variable $\mathfrak{e}_k$, we can study in more depth the time difference between two consecutive ladder times.

With high probability, the first $\mathfrak{e}_k$ neighbours of v_k in the tree constructed by the DFS all belong to distinct connected components of $\mathscr{S}_k\setminus\{v_k\}$. We denote these components by $W^{(1)},\dots,W^{(\mathfrak{e}_k)}$. Notice that by Lemma 4, for all $i\geq 0$, the ratio $|\mathbf{Ext}_i^k|/|\mathbf{Ext}^k|$ concentrates around $ip_i(k)(1-\rho_k)^{i-1}/g_k'(1)$. Therefore, conditionally on $\mathfrak{e}_k$, with probability $1-\mathcal{O}(N^{-\lambda})$, the size of these components can be coupled with the size of $\mathfrak{e}_k$ i.i.d. Galton-Watson trees independent of $\mathfrak{e}_k$ and whose reproduction laws have generating series given by $\tilde{g}_k(s) := \hat{g}_k((1-\rho_k)s)/(1-\rho_k)$. Therefore, the expected size of a component is given by:

$$\mathbb{E}\left[\left|W^{(1)}\right|\,\middle|\,\mathcal{F}_k\right] = \frac{1}{1-\tilde{g}_k'(1)}+\mathcal{O}(N^{-\lambda}) = \frac{1}{1-\hat{g}_k'(1-\rho_k)}+\mathcal{O}(N^{-\lambda}),$$

and we obtain, using Equation (6.11):

$$\begin{aligned}\mathbb{E}\left[T_{k+1}-T_k\,\middle|\,\mathcal{F}_k\right] &= 1+2\times\mathbb{E}\left[\sum_{p=1}^{\mathfrak{e}_k}\left|W^{(i)}\right|\,\middle|\,\mathcal{F}_k\right]\\ &= 1+2\left(\frac{1-\rho_k}{\rho_k}\left(1-\hat{g}_k'(1-\rho_k)\right)+\mathcal{O}(N^{-\lambda})\right)\left(\frac{1}{1-\hat{g}_k'(1-\rho_k)}+\mathcal{O}(N^{-\lambda})\right)\\ &= \frac{2-\rho_k}{\rho_k}+\mathcal{O}\left(N^{-\lambda}\right)\\ &= f\left(\frac{N_0(k)}{N},\frac{N_1(k)}{N},\dots\right)+\mathcal{O}(N^{-\lambda})\end{aligned}\tag{6.12}$$

which is the desired result for the evolution of (T_k).

We now turn to the evolution of the $(N_i(k))$ which follows from the analysis of the expectation of the terms (6.3) and (6.4). The term (6.3) accounts for the vertices of degree i in the graph $\mathscr{S}_k\setminus\mathscr{S}_{k+1}$. Among these vertices, the vertex v_k has a special role because it is conditioned to be matched to an element of $\mathbf{Surv}^k$. Therefore, we write

$$V_i(\mathscr{S}_k\setminus\mathscr{S}_{k+1})=\mathbf{1}_{\{\deg_{\mathscr{S}_k}(v_k)=i\}}+\sum_{j=1}^{\mathfrak{e}_k}\sum_{v\in W^{(j)}}\mathbf{1}_{\{\deg_{\mathscr{S}_k}(v)=i\}}.$$

We first compute the expectation of the sum in the right hand side of the previous equation. The connected components $W^{(1)},\dots,W^{(\mathfrak{e}_k)}$ are well approximated by independent Galton-Watson trees with offspring distribution given by $\hat{g}_n$, conditioned on extinction. Let C_i be the number of individuals that have $i-1$ children in such a tree. These individuals all have degree i in $\mathscr{S}_k$ and contribute to the sum. The quantity C_i satisfies the following recursion established by summing over the possible number of children of the root:

$$\mathbb{E}[C_i]=\mathbb{E}\left[\sum_{l\geq 0}\hat{p}_l(1-\rho_k)^l\left(lC_i+\delta_{l=i-1}\right)\right]=\mathbb{E}[C_i]\hat{g}_k'(1-\rho_k)+\hat{p}_{i-1}(1-\rho_k)^{i-1},$$

which leads to

$$\mathbb{E}[C_i]=\frac{\hat{p}_{i-1}(1-\rho_k)^{i-1}}{1-\hat{g}_k'(1-\rho_k)}.\tag{6.13}$$

Therefore, multiplying (6.11) and (6.13), we obtain

$$\mathbb{E}\left[V_i(\mathscr{S}_k\setminus\mathscr{S}_{k+1})\,\middle|\,\mathcal{F}_k\right]=\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_k)=i\,\middle|\,\mathcal{F}_k\right)+\frac{\hat{p}_{i-1}}{\rho_k}(1-\rho_k)^{i-1}+\mathcal{O}\left(N^{-\lambda}\right).\tag{6.14}$$

Note that the sum over i of these terms gives the total number of vertices in the connected components associated to the first $\mathfrak{e}_k$ children of v_k: $(1-\rho_n)/\rho_n+o(1)$. This is in agreement with Equation (6.12).

For the last term (6.4), we use the fact that, with probability $1-\mathcal{O}(N^{-\lambda})$, the elements of $\mathbf{m}_{v_k}$ that belong to $\mathscr{S}_{k+1}$ are distinct. One of these elements is v_{k+1} and has a special role, while all the others correspond to a uniform matching to a half edge of a vertex of $\mathscr{S}_{k+1}\setminus\{v_{k+1}\}$ and therefore have degree i with probability $\hat{p}_{i-1}$. Note that there are $\deg_{\mathscr{S}_k}(v_k)-\mathfrak{e}_k-1$ terms in

the sum (6.4) when excluding v_{k+1}. We have:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{v\in \mathbf{m}_{v_k}\cap\mathscr{S}_{k+1}}\left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(v)=i}+\mathbf{1}_{\deg_{\mathscr{S}_k}(v)=i+1}\right)\right]\\
&\quad=-\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i\,\middle|\,\mathcal{F}_k\right)+\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i+1\,\middle|\,\mathcal{F}_k\right)\\
&\qquad+\mathbb{E}\left[\deg(v_k)-\mathfrak{e}_k-1\,\middle|\,\mathcal{F}_k\right](-\hat{p}_{i-1}+\hat{p}_i)+\mathcal{O}(N^{-\lambda})\\
&\quad=-\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i\,\middle|\,\mathcal{F}_k\right)+\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i+1\,\middle|\,\mathcal{F}_k\right)\\
&\qquad+\frac{1}{\rho_k}\Bigg[\left(\hat{g}_k'(1)-(1-\rho_k)\hat{g}_k'(1-\rho_k)-(1-\rho_k)(1-\hat{g}_k'(1-\rho_k))-\rho_k\right)\\
&\qquad\qquad\times(-\hat{p}_{i-1}+\hat{p}_i)\Bigg]+\mathcal{O}(N^{-\lambda})\\
&\quad=-\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i\,\middle|\,\mathcal{F}_k\right)+\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i+1\,\middle|\,\mathcal{F}_k\right)\\
&\qquad+\frac{1}{\rho_k}\left(1-\hat{g}_k'(1)\right)(\hat{p}_{i-1}-\hat{p}_i)+\mathcal{O}(N^{-\lambda}).
\end{aligned}\tag{6.15}$$

Hence, summing (6.14) and (6.15), we obtain the total contribution of (6.3) and (6.4):

$$\begin{aligned}
\mathbb{E}\left[N_i(k+1)-N_i(k)\,\middle|\,\mathcal{F}_k\right]&=-\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_k)=i\,\middle|\,\mathcal{F}_k\right)-\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i\,\middle|\,\mathcal{F}_k\right)\\
&\quad+\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i+1\,\middle|\,\mathcal{F}_k\right)\\
&\quad-\frac{\hat{p}_{i-1}}{\rho_k}(1-\rho_k)^{i-1}+\frac{1}{\rho_k}\left(1-\hat{g}_k'(1)\right)(\hat{p}_{i-1}-\hat{p}_i)+\mathcal{O}(N^{-\lambda}).
\end{aligned}$$

Recall that the conditional law of $\deg_{\mathscr{S}}(v_k)$ is given by equation (6.10). Similar arguments to those used to compute it lead to

$$\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_{k+1})=i\,\middle|\,\mathcal{F}_k\right)=\mathbb{P}\left(\deg_{\mathscr{S}_k}(v_k)=i-1\,\middle|\,\mathcal{F}_k\right)+\mathcal{O}(N^{-\lambda}).$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}\left[N_i(k+1)-N_i(k)\,\middle|\,\mathcal{F}_k\right]&=-\frac{\hat{p}_{i-1}}{\rho_k}\left(1-(1-\rho_k)^{i-1}\right)-\frac{\hat{p}_{i-1}}{\rho_k}(1-\rho_k)^{i-1}\\
&\quad+\frac{1}{\rho_k}\left(1-\hat{g}_k'(1)\right)(\hat{p}_{i-1}-\hat{p}_i)+\mathcal{O}(N^{-\lambda})\\
&=-\frac{\hat{p}_{i-1}}{\rho_k}+\frac{1}{\rho_k}\left(1-\hat{g}_k'(1)\right)(\hat{p}_{i-1}-\hat{p}_i)+\mathcal{O}(N^{-\lambda})\\
&=f_i\left(\frac{N_0(k)}{N},\frac{N_1(k)}{N},\dots\right)+O\left(N^{-\lambda}\right).
\end{aligned}$$

This ends the proof of Theorem 12. □

6.5 Proofs of the main results

We now turn to the proofs of Theorems 10 and 11. We will use the following general fact about contour processes of trees, which can be easily proved by induction on n.

$$\forall n\geq 0,\quad \text{number of vertices explored by the DFS by time n}=\frac{n+X_n}{2}.\tag{6.16}$$

6.5.1 Proof of Theorem 10

The time variable in Theorem 10 is the proportion of vertices explored by the DFS whereas in Theorem 12 it is the index of the ladder times T_k. Therefore, to prove Theorem 10, a first step is to study the asymptotic proportion of vertices explored by time T_k. By Equation (6.16), for all $N \geq 1$ and all $1 \leq k \leq K_\varepsilon$, this proportion is given by $\omega(T_k) := \frac{k+T_k}{2N}$. Therefore, by Theorem 12, this proportion satisfies

$$\omega(T_k) = \widetilde{z}\left(\frac{k}{N}\right) + o(1), \qquad \text{with} \quad \widetilde{z}(t) = \frac{1}{2}\left(t + z\left(t\right)\right). \tag{6.17}$$

Fix $0 \leq \alpha < \alpha_c$ and recall the definition of $\tau^{(N)}(\alpha)$ given in Theorem 10. At time $T_{N\widetilde{z}^{-1}(\alpha)}$, by Equation (6.17), the number of explored vertices is $\alpha N + o(N)$. Therefore $\tau^{(N)}(\alpha) = T_{N\widetilde{z}^{-1}(\alpha)} + o(1)$. Hence, for all $i \geq 0$,

$$\begin{aligned} N_i(\tau^{(N)}(\alpha)) &= N_i\left(T_{N\widetilde{z}^{-1}(\alpha)} + o(1)\right) \\ &= Nz_i\left(\widetilde{z}^{-1}(\alpha)\right) + o(N). \end{aligned}$$

It is easy to check that the sequence of functions $(z_i \circ \widetilde{z}^{-1})_{i\geq 0}$ is solution of the system (S') of Lemma 7 below. The generating function $g(\alpha, s)$ of Theorem 10 is given by

$$g(\alpha, s) = \frac{1}{1-\alpha}\sum_{i\geq 0} z_i \circ \widetilde{z}^{-1}(\alpha)s^i,$$

which is the desired result by Equation (6.26) and Proposition 9.

6.5.2 Proof of Theorem 11

Let $N \geq 1$. By definition, for all $1 \leq k \leq K_\varepsilon$, the contour process of the tree constructed by the DFS algorithm at time T_k is located at point (T_k, k). Furthermore, by Theorem 12,

$$(T_k, k) = N\left(z\left(\frac{k}{N}\right) + o(1), \frac{k}{N}\right).$$

Note that $|T_{k+1} - T_k| = o(N)$ and that, between two consecutive T_k's, the contour process cannot fluctuate by more than $o(N)$. Hence, after normalization by N, the limiting contour process converges to the curve $(z(t), t)$ where t ranges from 0 to $t_{\max} = \sup\{t > 0,\, z'(t) < +\infty\}$. Recall that by the definition of z in Theorem 12 and Equation (6.7), $z'(t) = (2 - \rho_{(z_i(t))_{i\geq 0}})/\rho_{(z_i(t))_{i\geq 0}}$. Hence, if we parametrize $(z(t), t)$ in terms of $\rho = \rho_{(z_i(t))_{i\geq 0}}$, the curve can be written $(x(\rho), y(\rho))$ where the functions x and y satisfy

$$\frac{x'(\rho)}{y'(\rho)} = \frac{2-\rho}{\rho}.$$

Note that when t ranges from 0 to $t_{\max}$, the parameter ρ decreases from ρ_π to 0. In order to get a second equation connecting x' and y', we go back to the discrete process and observe that, by Equation (6.16), the number of explored vertices at time T_k is equal to $(k + T_k)/2$. Using the notation of Theorem 10, let $\hat{g}(\alpha, \cdot)$ be the size-biased version of $g(\alpha, \cdot)$. For all $\rho \in (0, \rho_\pi]$, let $\alpha(\rho)$ be the unique solution of $1 - \rho = \hat{g}(\alpha(\rho), 1 - \rho)$. After renormalizing by N, we get that:

$$\frac{x(\rho) + y(\rho)}{2} = \alpha(\rho).$$

This yields the following system of equations:

$$\begin{cases} \frac{x'(\rho)}{y'(\rho)} = \frac{2-\rho}{\rho} \\ \frac{x'(\rho)+y'(\rho)}{2} = \alpha'(\rho). \end{cases}$$

Therefore,

$$\begin{cases} x'(\rho) = (2-\rho)\alpha'(\rho) \\ y'(\rho) = \rho\alpha'(\rho). \end{cases}$$

Integrating by parts, this gives the formulas for $x^{\uparrow}$ and $y^{\uparrow}$ in Theorem 11. Fix $\rho \in (0, \rho_{\pi}]$. Then, the asymptotic profile of the decreasing phase of the DFS is obtained by translating horizontally each point $(x^{\uparrow}(\rho), y^{\uparrow}(\rho))$ of the ascending phase to the right by twice the asymptotic proportion of the giant component of the remaining graph of parameter ρ, which is $2(1-g(\alpha(\rho), 1-\rho))$. Indeed, the time it takes to the DFS to return at a given height k attained during the ascending phase corresponds to the time of exploration of the giant component of the unexplored graph at time T_k. The latter is given by twice the number of vertices of the giant component which is equal to $2(1-g_k(1-\rho_k))$.

6.6 Technical lemmas

6.6.1 Asymptotic densities in a configuration model

In this section we establish Lemma 4. The proofs of each of the four estimates follow the same scheme, therefore we only focus on the proof the last one, namely that there exists $\lambda > 0$ such that:

$$\mathbb{P}\left(\left| \frac{|\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - \rho_k \right| \geq N^{-\lambda} \right) = \mathcal{O}\left(N^{-1-\lambda}\right).$$

First, notice that for the values of k that we consider and under our assumptions (**A1**) and (**A2**), the number of edges and vertices of the graphs $\mathbf{C}^d_{\mathbf{N}(k)}$ are all of order N. Therefore, it is enough to prove the following bound:

$$\mathbb{P}\left(\left| \frac{|\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})|}{2|E(\mathbf{C}^d_{\mathbf{N}(k)})|} - \rho_k \right| \geq |E(\mathbf{C}^d_{\mathbf{N}(k)})|^{-\lambda} \right) = \mathcal{O}\left(|E(\mathbf{C}^d_{\mathbf{N}(k)})|^{-1-\lambda}\right).$$

This is a direct consequence of the two following Lemmas. The first one is a general concentration result for the configuration model.

Lemma 5. *Fix $\gamma > 2$ and $n \geq 1$. Let $\mathbf{d} = (d_1, \ldots, d_n)$ be such that $\max\{d_1, \ldots, d_n\} \leq n^{1/\gamma}$. Fix also $\delta \in (0, 1/2)$ and recall that, for a graph G, $\mathbf{Surv}(G)$ denotes the set of half edges of G attached to a vertex v such that the connected component of v after removing this half edge has at least n^{δ} vertices. Let $m = \sum_i d_i$ the number of half edges of a configuration graph $\mathcal{C}(\mathbf{d})$, then, for any $\delta' \geq \delta$ one has*

$$\mathbb{P}\left(\left| \frac{|\mathbf{Surv}(\mathcal{C}(\mathbf{d}))|}{m} - \frac{\mathbb{E}\left(|\mathbf{Surv}(\mathcal{C}(\mathbf{d}))|\right)}{m} \right| \geq \frac{n^{\delta'+\frac{1}{\gamma}}}{2\sqrt{m}} \right) \leq C \exp\left(-Cn^{2(\delta'-\delta)}\right).$$

The second Lemma consists in an estimation of the expectation of $|\mathbf{Surv}(\mathcal{C}(\mathbf{d}^{(n)}))|$ for a sequence of configuration models that satisfy the assumptions of Theorem 10.

Lemma 6. *Let $(\mathcal{C}(\mathbf{d}^{(n)}))_{n\geq 1}$ be a sequence of configuration models with asymptotic degree distribution $\boldsymbol{\pi}$. We suppose that $\boldsymbol{\pi}$ is supercritical in the sense of Definition 4 and that the sequence $\mathbf{d}^{(n)}$ satisfies assumption (**A1**) and (**A2**).*

For all $n \geq 1$, let g_n be the generating series associated to the empirical distribution of the degree sequence $\mathbf{d}^{(n)}$. Let ρ_n be the smallest positive solution of the equation $\hat{g}_n(1-x) = 1-x$. Then, for n sufficiently large:

$$\frac{\mathbb{E}\left[|\mathbf{Surv}(\mathcal{C}(\mathbf{d}^{(n)}))|\right]}{2g_n'(1)} = \rho_n + \mathcal{O}\left(n^{2\delta+\frac{1}{\gamma}-1}\right).$$

Proof of Lemma 5. In order to prove Lemma 5, it is sufficient to check that the function $\mathbf{Surv}(\cdot)$ is Lipschitz in the following sense. We say that two configuration models are related by a switching if they differ by exactly two pairs of matched half-edges (see Figure 6.5). Then, we claim that $\mathbf{Surv}(\cdot)$ is such that, for any two graphs G_1 and G_2 differing by a switching:

$$||\mathbf{Surv}(G_1)| - |\mathbf{Surv}(G_2)|| \leq 8n^{\delta+\frac{1}{\gamma}}. \tag{6.18}$$

Using a result of Bollobás and Riordan [BR15, Lemma 8], this regularity implies the following concentration inequality:

$$\mathbb{P}\left(\left||\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})| - \mathbb{E}[|\mathbf{Surv}(\mathbf{C}^d_{\mathbf{N}(k)})|]\right| \geq t\right) \leq 2\exp\left(\frac{-t^2}{Cn^{2\delta+\frac{2}{\gamma}}m}\right), \tag{6.19}$$

Figure 6.5 – Switching two edges in a graph.

By taking $t = n^{\delta'+\frac{1}{\gamma}}m^{\frac{1}{2}}$ in (6.19), we obtain Lemma 5.

It remains to prove inequality (6.18). To pass from G_1 to G_2, one has to delete two edges in G_1 and then add two other edges. Therefore, it suffices to study the effect of adding an edge e on a graph G having maximal degree $n^{1/\gamma}$. Indeed, the effect of deleting an edge f of a graph H is equal to the effect of adding the edge f to the graph $H \setminus \{f\}$.

Let u and v be the extremities of e. Let us define two partial orders associated respectively to u and v among the half-edges of $\mathbf{Ext}(G) = \mathbf{Surv}(G)^c$. We say that:

- $e_1 \preceq_u e_2$ if all the paths connecting e_2 to u contain e_1,
- $e_1 \preceq_v e_2$ if all the paths connecting e_2 to v contain e_1.

Let f_u (resp. f_v) be a maximal element for the partial order $\preceq_u$ (resp. $\preceq_v$), and denote by $\mathscr{C}_{f_u}$ (resp. $\mathscr{C}_{f_v}$) the connected component of the extremity of f_u (resp. f_v) after the removal of f_u (resp. f_v) in G. Then, by maximality, the set of extremities of half-edges that change their status from $\mathbf{Ext}(G)$ to $\mathbf{Surv}(G)$ after adding e is included in $\mathscr{C}_{f_u} \cup \mathscr{C}_{f_v}$. See Figure 6.6 for an illustration. Since f_u (resp. f_v) was in $\mathbf{Ext}(G)$, the number of vertices in $\mathscr{C}_{f_u}$ (resp. $\mathscr{C}_{f_v}$) is at most n^{δ}. Since the maximal degree of a vertex in G is $n^{1/\gamma}$, we deduce that:

$$||\mathbf{Surv}_n(G)| - |\mathbf{Surv}_n(G \cup e)|| \leq 2n^{\delta+\frac{1}{\gamma}}.$$

This implies (6.18) and Lemma 5. □

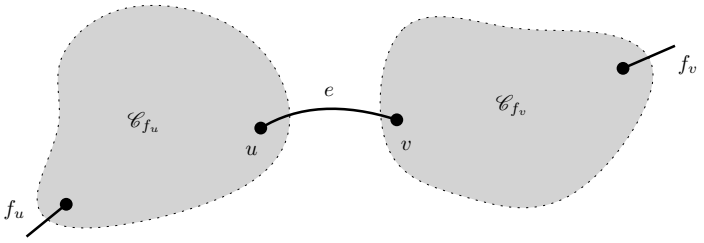


Figure 6.6 – Effect of the edge e.

Proof of Lemma 6. Fix $n \geq 1$. Let e be a uniformly chosen half-edge in $\mathcal{C}(\mathbf{d}^{(n)})$ and let v be the extremity of e. We denote $\mathscr{C}_v$ the connected component of v inside $\mathcal{C}(\mathbf{d}^{(n)}))$ after removing e. Then, since $\mathbb{E}\left[|\mathbf{Surv}(\mathcal{C}(\mathbf{d}^{(n)}))|\right] = 2g_n'(1)\mathbb{P}\left(e \in \mathbf{Surv}(\mathcal{C}(\mathbf{d}^{(n)}))\right)$, it is sufficient to prove that

$$\mathbb{P}\left(|\mathscr{C}_v| \geq n^\delta\right) = \rho_n + \mathcal{O}\left(n^{2\delta+\frac{2}{\gamma}-1}\right). \tag{6.20}$$

Let $(d_i^{\uparrow})_{1\leq i\leq n}$ and $(d_i^{\downarrow})_{1\leq i\leq n}$ respectively denote the increasing and decreasing reordering of the degree sequence $(d_i)_{1\leq i\leq n}$:

$$d_1^{\uparrow} \leq \cdots \leq d_n^{\uparrow} \qquad \text{and} \qquad d_1^{\downarrow} \geq \cdots \geq d_n^{\downarrow}.$$

In order to prove (6.20), we will use a coupling argument. More precisely, we first introduce two Galton-Watson trees:

- $\mathscr{T}^-$ with reproduction law: $q_i^- := \frac{(i+1)|\{j\geq\lceil n^\delta\rceil, d_j^{\downarrow}=i+1\}|}{\sum_{j\geq\lceil n^\delta\rceil}(j+1)d_j^{\downarrow}}$,
- $\mathscr{T}^+$ with reproduction law: $q_i^+ := \frac{(i+1)|\{j\geq\lceil n^\delta\rceil, d_j^{\uparrow}=i+1\}|}{\sum_{j\geq\lceil n^\delta\rceil}(j+1)d_j^{\uparrow}}$.

We also let E be the event where, in the $\lfloor n^\delta \rfloor$ first steps of the exploration of $\mathscr{C}_v$, no loop is discovered. Then, the following inequalities hold:

$$(1-\mathbb{P}(E))\,\mathbb{P}(|\mathscr{T}^-| \geq n^\delta) \leq \mathbb{P}\left(|\mathscr{C}_v| \geq n^\delta\right) \leq \mathbb{P}(|\mathscr{T}^+| \geq n^\delta). \tag{6.21}$$

Now, we prove that:

$$\begin{cases} \mathbb{P}(|\mathscr{T}^-| \geq n^\delta) = \rho_n + \mathcal{O}(n^{\delta+\frac{1}{\gamma}-1}), \\ \mathbb{P}(|\mathscr{T}^+| \geq n^\delta) = \rho_n + \mathcal{O}(n^{\delta+\frac{1}{\gamma}-1}). \end{cases} \tag{6.22}$$

Since the proofs of these two bounds are similar, we only focus on the second one. Let $g_n^+(s) = \sum_{k\geq 0} q_k^+ s^k$ be the generating series of $(q_k^+)_{k\geq 0}$. Let ρ_n^+ be the smallest positive solution of $g_n^+(1-x) = 1-x$. Then:

$$\begin{aligned} \mathbb{P}(|\mathscr{T}^+| \geq n^\delta) &= \mathbb{P}(|\mathscr{T}^+| = +\infty) + \mathbb{P}(n^\delta \leq |\mathscr{T}^+| < +\infty) \\ &= \rho_n^+ + o\left(\frac{1}{n}\right). \end{aligned} \tag{6.23}$$

The difference between ρ_n^+ and ρ_n can be written as follows:

$$\begin{aligned} \rho_n^+ - \rho_n &= g_n^+(1-\rho_n^+) - g_n(1-\rho_n) \\ &= g_n(1-\rho_n^+) - g_n(1-\rho_n) + g_n^+(1-\rho_n^+) - g_n(1-\rho_n^+) \\ &= g_n'(1-\rho_n)(\rho_n - \rho_n^+) + o\left(\rho_n^+ - \rho_n\right) + g_n^+(1-\rho_n^+) - g_n(1-\rho_n^+), \end{aligned} \tag{6.24}$$

where in the last equality, we used a Taylor expansion. From the definition of $(q_k^+)_{k\geq 0}$, for all $k \geq 0$, it holds that:

$$q_k^+ = p_k + \mathcal{O}\left(\frac{n^{\delta+\frac{1}{\gamma}}}{n}\right),$$

where the error term is uniform in k. In particular, this implies that $g_n^+(1-\rho_n^+) - g_n(1-\rho_n^+)$ is of order $n^{\delta+\frac{1}{\gamma}-1}$. Inserting this into (6.24), we get

$$\left(1 - g_n'(1-\rho_n) + o(1)\right)\left(\rho_n^+ - \rho_n\right) = \mathcal{O}\left(n^{\delta+\frac{1}{\gamma}-1}\right).$$

By the assumptions of Lemma 6, ρ_n converges to the fixed point of g_π, which is bounded away from 0. Therefore, for large enough n , $g_n'(1-\rho_n)$ is bounded away from 1. Hence

$$|\rho_n^+ - \rho_n| = \mathcal{O}\left(n^{\delta+\frac{1}{\gamma}-1}\right).$$

Together with (6.23), this implies (6.22).

It remains to estimate the probability of the event E. During the first $\lfloor n^\delta \rfloor$ steps of the exploration of $\mathscr{C}_v$, the number of half-edges of the explored cluster is at most $n^\delta \times n^{1/\gamma}$. Hence, the probability of creating a loop at each of these steps is of order $n^{\delta+\frac{1}{\gamma}-1}$. Therefore, by the union bound:

$$\mathbb{P}(E) = \mathcal{O}\left(n^{2\delta+\frac{1}{\gamma}-1}\right). \tag{6.25}$$

Gathering (6.21), (6.22) and (6.25), we get (6.20) and therefore Lemma 6. □

6.6.2 An infinite system of differential equations

The aim of this section is to prove Lemma 3. In the following, we fix a probability distribution $\boldsymbol{\pi} = (\pi_i)_{i\geq 0}$ which is supercritical in the sense of Definition 4.

First, we prove that the problem can be reduced to the study of another system of differential equations. Recall that, given a sequence $(\zeta_i)_{i\geq 0} \in \mathbb{R}^{\mathbb{Z}_+}$ such that $\sum_{i\geq 0}\zeta_i \leq 1$, the implicit quantity $\rho_{(\zeta_i)_{i\geq 0}}$ is defined through Equations (6.5) and (6.6).

Lemma 7. *If the following system has a unique solution well defined on some maximal interval $[0, t'_{\max})$ for some $t'_{\max} > 0$:*

$$\begin{cases} \frac{\mathrm{d}\zeta_i}{\mathrm{d}t} &= -\frac{i\zeta_i}{\sum_{j\geq 0} j\zeta_j} + \frac{1}{\sum_{j\geq 0} j\zeta_j}\left(1 - \frac{\sum_{j\geq 0}(j-1)j\zeta_j}{\sum_{n\geq 0} j\zeta_j}\right)\left(i\zeta_i - (i+1)\zeta_{i+1}\right) \\ \zeta_i(0) &= \boldsymbol{\pi}_i, \end{cases} \tag{S'}$$

then the system (S) has a unique solution well defined on a maximal interval $[0, t_{\max})$ for some $t_{\max} > 0$.

Proof. Suppose that (S′) has a unique solution $(\zeta_i)_{i\geq 0}$. Let ϕ be the unique function defined by

$$\begin{cases} \phi'(t)\rho_{(\zeta_i(t))_{i\geq 0}} = 1, \\ \phi(0) = 0. \end{cases}$$

Then, for all $i \geq 0$, $(\zeta_i \circ \phi)'(t) = \frac{1}{\rho_{(\zeta_i(t))_{i\geq 0}}} \times \rho_{(\zeta_i(t))_{i\geq 0}} f_i(\zeta_0(t), \zeta_1(t), \ldots) = f_i(\zeta_0(t), \zeta_1(t), \ldots)$ which proves that $(\zeta_i \circ \phi)_{i\geq 0}$ is a solution of the system (S).

Let $(z_i)_{i\geq 0}$ be a solution of (S). Then, for all $t \geq 0$ where it is well defined,

$$\sum_{i\geq 0} z_i(t) = 1 - \int_0^t \frac{1}{\rho_{(z_i(t))_{i\geq 0}}} \mathrm{d}u =: 1 - \psi(t).$$

Then, $(z_i \circ \psi^{-1})_{i\geq 0}$ is a solution of (S′). Therefore, since $(\zeta_i)_{i\geq 0}$ is unique, $(z_i \circ \psi^{-1} \circ \phi)_{i\geq 0} = (\zeta_i \circ \phi)_{i\geq 0}$ is also solution of (S). In particular, this implies that

$$\frac{-1}{\rho_{(z_i\circ\psi^{-1}\circ\phi(t))_{i\geq 0}}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i\geq 0} z_i \circ \psi^{-1} \circ \phi\right)(t) = \left(\psi^{-1}\circ\phi\right)'(t) \times \frac{-1}{\rho_{(z_i\circ\psi^{-1}\circ\phi(t))_{i\geq 0}}}.$$

Therefore, $\psi = \phi$ only depends on $(\zeta_i)_{i\geq 0}$, yielding the uniqueness of the solution. □

We now exhibit a solution of (S′). Let $f_\pi(s) = \sum_{i\geq 0} \pi_i s^i$ be the generating series associated to π. Define $t'_{\max}$ to be the unique root between 0 and $1 - \pi_0$ of the equation

$$\frac{f''_\pi\left(f_\pi^{-1}(1-t)\right)}{f'_\pi(1)} = 1.$$

For all $0 \leq t \leq t'_{\max}$ and $0 \leq s \leq 1$, let

$$f(t,s) := f_\pi\left(f_\pi^{-1}(1-t) - (1-s)\frac{f'_\pi(f_\pi^{-1}(1-t))}{f'_\pi(1)}\right). \tag{6.26}$$

Note that this restriction to the interval $[0, t'_{\max})$ will play a crucial role in the analytic proof of the uniqueness of the solution. Moreover, from a probabilistic point of view, it corresponds to the range of times where $\frac{1}{1-t}f_\pi(t,s)$ is the generating series of a supercritical probability law.

Proposition 9. *For all $0 \leq t \leq t'_{\max}$ and $i \geq 0$, let $\zeta_i(t) := [s^i]f(t,s)$ be the coefficient of s^i in $f(t,s)$. Then $(\zeta_i)_{i\geq 0}$ is a solution of (S′).*

Proof. It can be easily verified that $f(t,s)$ satisfies the following equation:

$$\frac{\partial f}{\partial t}(t,s) = \frac{\frac{\partial f}{\partial s}(t,s)}{\frac{\partial f}{\partial s}(t,1)}\left((1-s)\frac{\frac{\partial^2 f}{\partial s^2}(t,1)}{\frac{\partial f}{\partial s}(t,1)} - 1\right).$$

By extracting the coefficient of s^i we get that

$$\frac{\mathrm{d}\zeta_i}{\mathrm{d}t} = -\frac{i\zeta_i}{\sum_{j\geq 0} j\zeta_j} + \frac{1}{\sum_{j\geq 0} j\zeta_j}\left(1 - \frac{\sum_{j\geq 0}(j-1)j\zeta_j}{\sum_{n\geq 0} j\zeta_j}\right)(i\zeta_i - (i+1)\zeta_{i+1}),$$

which ends the proof the proposition. □

It remains to prove the uniqueness of the solution that we found. Let $(\zeta_i)_{i\geq 0}$ be a solution of (S′). We will prove that $\sum_{i\geq 0}\zeta_i(t)s^i = f(t,s)$, which implies that for all $i \geq 0$, the function ζ_i is the coefficient of s^i in $f(t,s)$.

Remark 14. *Notice that when π has bounded support, we only have to deal with a finite number of differential equations and the uniqueness follows merely from the Cauchy-Lipschitz Theorem.*

We introduce the following quantities:

$$E(t) := \sum_{i\geq 0} i\zeta_i(t) \qquad \text{and} \qquad Z(t) := \int_0^t \left(\frac{E'}{2\sqrt{E}} + \frac{1}{\sqrt{E}}\right)(u)\mathrm{d}u.$$

Lemma 8. *For all $0 \leq t \leq t'_{\max}$:*

1. $\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i\geq 0}\zeta_i\right)(t) = -1$;

2. $E'(t) = -2\frac{\sum_{i\geq 1} i(i-1)\zeta_i(t)}{E(t)}$.

In particular, $\sum_{i\geq 0}\zeta_i(t) = 1-t$.

Proof. The first point is obtained by summing the equations of (S′). Let us prove the second point. We omit the argument t for brevity.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i\geq 1} i\zeta_i\right) &= \frac{-1}{\left(\sum_{i\geq 1} i\zeta_i\right)^2}\left[\left(\sum_{i\geq 1} i\zeta_i\right)\left(\sum_{i\geq 1} i^2\zeta_i\right)\right.\\ &\qquad \left.-\left(\sum_{i\geq 1} i\zeta_i - \sum_{i\geq 1} i(i-1)\zeta_i\right)\left(\sum_{i\geq 1} i^2\zeta_i - \sum_{i\geq 1} i(i-1)\zeta_i\right)\right]\\ &= \frac{-1}{\sum_{i\geq 1} i\zeta_i}\left(2\sum_{i\geq 0} i^2\zeta_i - 2\sum_{i\geq 0} i\zeta_i\right)\\ &= -2\frac{\sum_{i\geq 0} i(i-1)\zeta_i}{\sum_{i\geq 0} i\zeta_i}.\end{aligned}$$

□

By Lemma 8, the system (S′) can be rewritten:

$$\frac{\mathrm{d}\zeta_i}{\mathrm{d}t} = \frac{i}{2}\frac{E'}{E}\zeta_i - (i+1)\zeta_{i+1}\left(\frac{E'}{2E} + \frac{1}{E}\right).$$

We are going to compare ζ_i with a truncated version of it $\zeta_i^{(\Delta)}$ defined below, the idea being to prove that ζ_i is arbitrarily close to $\zeta_i^{(\Delta)}$ and that $\zeta_i^{(\Delta)}$ converges to the coefficient of s^i in $f(t,s)$, as $\Delta \to +\infty$.

Let $\varepsilon > 0$ and let

$$\Delta := \Delta(\varepsilon) = \left\lfloor \sqrt{\frac{\sum_{k\geq 0}\boldsymbol{\pi}(\{k\})k^2}{\varepsilon}} \right\rfloor. \tag{6.27}$$

Note that, by Markov's inequality, $\sum_{i\geq\Delta}\boldsymbol{\pi}(\{i\}) \leq \varepsilon$. Let $(\zeta_i^{(\Delta)})_{0\leq i\leq\Delta}$ be *the* solution of the following *finite* system of differential equations:

$$\begin{cases} \frac{\mathrm{d}\zeta_\Delta^{(\Delta)}}{\mathrm{d}t} &= \frac{i}{2}\frac{E'}{E}\zeta_\Delta^{(\Delta)};\\ \frac{\mathrm{d}\zeta_i^{(\Delta)}}{\mathrm{d}t} &= \frac{i}{2}\frac{E'}{E}\zeta_i^{(\Delta)} - (i+1)\zeta_{i+1}^{(\Delta)}\left(\frac{E'}{2E} + \frac{1}{E}\right);\\ \zeta_i^{(\Delta)}(0) &= \boldsymbol{\pi}_i. \end{cases}$$

It turns out that the generating function of the $\zeta_i^{(\Delta)}$ has a simple expression in term of the functions E and Z.

Lemma 9. *Let* $f_\Delta(s) := \sum_{0\leq i\leq\Delta}\boldsymbol{\pi}_i s^i$ *be the truncated version of* $f_{\boldsymbol{\pi}}$. *Then, for all* $0 \leq t \leq t'_{\max}$ *and* $0 \leq s \leq 1$*:*

$$\sum_{0\leq i\leq\Delta}\zeta_i^{(\Delta)}(t)s^i = f_\Delta\left(\frac{s\sqrt{E(t)} - Z(t)}{\sqrt{E(0)}}\right). \tag{6.28}$$

Moreover, the initial solution is close to its truncated version.

Lemma 10. *For all* $0 \le t \le t'_{\max}$ *and all* $0 \le i \le \Delta$, $\zeta_i^{(\Delta)}(t) \le \zeta_i(t) \le \zeta_i^{(\Delta)} + 2\varepsilon$.

We postpone the proofs of Lemmas 9 and 10 to the end of this section and explain now how they lead to the uniqueness part of Lemma 3.

By Lemma 10,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}\right)(t) = -\frac{\sum_{0\le i\le\Delta} i\zeta_i^{(\Delta)}(t)}{\sum_{i\ge 0} i\zeta_i(t)} \ge -1.$$

By our choice of Δ, $\sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}(0) = \sum_{0\le i\le\Delta}\pi_i \ge 1-\varepsilon$. Therefore:

$$1 - t - \varepsilon \le \sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}(t) \le \sum_{0\le i\le\Delta}\zeta_i(t) = 1 - t.$$

Evaluating (6.28) at $s = 1$ gives:

$$1 - t - \varepsilon \le f_\Delta\left(\frac{\sqrt{E(t)} - Z(t)}{\sqrt{E(0)}}\right) \le 1 - t.$$

Recalling the definition of Δ in (6.27) and letting ε converge to zero, we get that

$$f_\pi\left(\frac{\sqrt{E(t)} - Z(t)}{\sqrt{E(0)}}\right) = 1 - t.$$

We now take the inverse of f_π and differentiate in t to obtain

$$\begin{cases} \sqrt{\frac{E(t)}{E(0)}} &= \frac{f'_\pi(f_\pi^{-1}(1-t))}{f'_\pi(1)}, \\ \frac{Z(t)}{\sqrt{E(0)}} &= \frac{f'_\pi(f_\pi^{-1}(1-t))}{f'_\pi(1)} - f_\pi^{-1}(1-t). \end{cases}$$

By re-inserting in (6.28), we have proved that, for all $t \ge 0$ and $0 \le s \le 1$,

$$\sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}(t)s^i = f_\Delta\left(f_\pi^{-1}(1-t) - (1-s)\frac{f'_\pi(f_\pi^{-1}(1-t)}{f'_\pi(1)}\right).$$

It is now easy to conclude since, by Lemma 10 and our choice of Δ:

$$\begin{aligned} f(t,s) = \lim_{\varepsilon\to 0}\sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}(t)s^i \le \sum_{i\ge 0}\zeta_i(t)s^i &= \sum_{0\le i\le\Delta}\zeta_i(t)s^i + \sum_{i>\Delta}\zeta_i(t)s^i \\ &\le \lim_{\varepsilon\to 0}\left(\sum_{0\le i\le\Delta}\zeta_i^{(\Delta)}(t)s^i + 2\Delta\varepsilon + \varepsilon\right) = f(t,s). \end{aligned}$$

This ends the proof of Lemma 3. We now turn to the proofs of Lemmas 9 and 10.

Proof of Lemma 9. We first prove by an induction from Δ to 0 that for all $0 \le i \le \Delta$,

$$\zeta_i^{(\Delta)}(t) = \sum_{k=i}^{\Delta} c_k \binom{n}{k}(-Z)^{k-i}E^{i/2},$$

where $c_k = \pi_k E(0)^{-k/2}$.

The initialization is straightforward because the function $\zeta_\Delta^{(\Delta)}(t) = c_\Delta E(t)^{\Delta/2}$ is indeed the solution of $y' = \frac{\Delta}{2}\frac{E'}{E}y$ with initial condition $y(0) = \pi_\Delta$.

Suppose that the property holds at $i+1$ for some $0 \leq i \leq \Delta - 1$. Since

$$\frac{\mathrm{d}\zeta_i^{(\Delta)}}{\mathrm{d}t} = \frac{i}{2}\frac{E'}{E}\zeta_i^{(\Delta)} - (i+1)\zeta_{i+1}^{(\Delta)}\left(\frac{E'}{2E} + \frac{1}{E}\right)$$

and $\zeta_i^{(\Delta)}(0) = \pi_i$,

$$\zeta_i^{(\Delta)}(t) = \pi_i \left(\frac{E(t)}{E(0)}\right)^{i/2} - \int_0^t (i+1)\zeta_{i+1}^{(\Delta)}(u)\left(\frac{E'(u)}{2E(u)} + \frac{1}{E(u)}\right)\left(\frac{E(t)}{E(u)}\right)^{-i/2}\mathrm{d}u.$$

We now use the induction hypothesis to obtain:

$$\begin{aligned}
\zeta_i^{(\Delta)}(t) &= \pi_i \left(\frac{E(t)}{E(0)}\right)^{i/2} - \int_0^t (i+1)\sum_{k=i+1}^{\Delta} c_k \binom{k}{k-i-1}(-Z(u))^{k-i-1} \\
&\qquad \times E^{(i+1)/2}(u)\left(\frac{E'(u)}{2E(u)} + \frac{1}{E(u)}\right)\left(\frac{E(t)}{E(u)}\right)^{i/2}\mathrm{d}u \\
&= \pi_i \left(\frac{E(t)}{E(0)}\right)^{i/2} - E(t)^{i/2}\int_0^t (i+1)\sum_{k=i+1}^{\Delta} c_k \binom{k}{k-i-1}(-Z(u))^{k-i-1} \\
&\qquad \times \left(\frac{E'(u)}{2\sqrt{E(u)}} + \frac{1}{\sqrt{E(u)}}\right)\mathrm{d}u \\
&= \pi_i \left(\frac{E(t)}{E(0)}\right)^{i/2} - E(t)^{i/2}\int_0^t (i+1)\sum_{k=i+1}^{\Delta} c_k \binom{k}{k-i-1}(-Z(u))^{k-i-1}Z'(u)\mathrm{d}u \\
&= c_i E(t)^{i/2} - \sum_{k=i+1}^{\Delta} c_k \frac{i+1}{k-i}\binom{k}{k-i-1}(-Z(t))^{k-i}E(t)^{i/2} \\
&= \sum_{k=i}^{\Delta} c_k \binom{k}{k-i}(-Z(t))^{k-i}E(t)^{i/2}.
\end{aligned}$$

This ends the proof by induction. It is now easy to conclude:

$$\begin{aligned}
\sum_{i=0}^{\Delta}\zeta_i^{(\Delta)}(t)s^i &= \sum_{i=0}^{\Delta}\sum_{k=i}^{\Delta} c_k \binom{k}{k-i}(-Z(t))^{k-i}E(t)^{i/2}s^i \\
&= \sum_{k=0}^{\Delta}\pi_k E(0)^{-k/2}\sum_{i=0}^{k}\binom{k}{k-i}(-Z(t))^{k-i}\left(s\sqrt{E(t)}\right)^i \\
&= \sum_{k=0}^{\Delta}\pi_k\left(\frac{s\sqrt{E(t)} - Z(t)}{\sqrt{E(0)}}\right)^k \\
&= f_\Delta\left(\frac{s\sqrt{E(t)} - Z(t)}{\sqrt{E(0)}}\right).
\end{aligned}$$

□

Proof of Lemma 10. We first prove the lower bound by an induction from Δ to 0. It is important to notice that for all $0 \leq t \leq t'_{\max}$,

$$-\left(\frac{E'}{2E} + \frac{1}{E}\right) = \frac{1}{\sum_{j\geq 0} j\zeta_j}\left(\frac{\sum_{j\geq 0}(j-1)j\zeta_j}{\sum_{n\geq 0} j\zeta_j} - 1\right) \geq 0.$$

Therefore, the lower bound holds for $i = \Delta$ since

$$\frac{\mathrm{d}}{\mathrm{d}t}\zeta_\Delta^{(\Delta)} = \frac{\Delta}{2}\frac{E'}{E}\zeta_\Delta^{(\Delta)} \leq \frac{\Delta}{2}\frac{E'}{E}\zeta_\Delta^{(\Delta)} - (\Delta+1)\zeta_{\Delta+1}\left(\frac{E'}{2E} + \frac{1}{E}\right).$$

Indeed, by Gronwall's Lemma, this implies that $\zeta_\Delta^{(\Delta)}$ is upper-bounded by the solution of the differential equation $y' = \frac{\Delta}{2}\frac{E'}{E}y - (\Delta+1)\zeta_{\Delta+1}\left(\frac{E'}{2E} + \frac{1}{E}\right)$, which is nothing but ζ_Δ.

Assume the lower bound holds for an index $1 \leq i \leq \Delta$. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\zeta_{i-1}^{(\Delta)} &= \frac{i-1}{2}\frac{E'}{E}\zeta_{i-1}^{(\Delta)} - i\zeta_i^{(\Delta)}\left(\frac{E'}{2E} + \frac{1}{E}\right) \\ &\leq \frac{i-1}{2}\frac{E'}{E}\zeta_{i-1}^{(\Delta)} - i\zeta_i\left(\frac{E'}{2E} + \frac{1}{E}\right).\end{aligned}$$

Again, by Gronwall's Lemma, it implies that $\zeta_{i-1}^{(\Delta)} \leq \zeta_{i-1}$.

The proof of the upper bound can be obtained by contradiction. Indeed, suppose that there exists a time $t \in (0, t'_{\max})$ and an index $0 \leq i \leq \Delta$ such that $\zeta_i(t) > \zeta_i^{(\Delta)}(t) + 2\varepsilon$. Then, using the lower bound we have just obtained,

$$\begin{aligned}1 - t = \sum_{i\geq 0}\zeta_i(t) &\geq \sum_{0\leq i\leq\Delta}\zeta_i(t) \\ &> \sum_{0\leq i\leq\Delta}\zeta_i^{(\Delta)}(t) + 2\varepsilon \\ &\geq 1 - \varepsilon - t + 2\varepsilon = 1 - t + \varepsilon.\end{aligned}$$

□

6.7 Appendix : a theorem by Wormald

In [Wor95], Wormald provided general conditions for a *finite* number of sequences of discrete stochastic processes to converge to a fluid limit after a proper rescaling. We first recall his result in the case of a single sequence and provide a proof containing some improved bounds which will allow us to generalize Wormald's result to a polynomial number of sequences of stochastic processes in Corollary 14 below.

Theorem 13. *For all $N \geq 1$, let $Y(i) = Y^{(N)}(i)$ be a Markov chain with respect to a filtration $\{\mathcal{F}_i\}_{i\geq 1}$. Suppose that*

- *$Y(0)/N$ converges to $z(0)$ in probability;*
- *$|Y(i+1) - Y(i)| \leq N^\beta$;*
- *$\mathbb{E}\left[Y(i+1) - Y(i) \,\middle|\, \mathcal{F}_i\right] = f\left(\frac{i}{N}, \frac{Y(i)}{N}\right) + O\left(N^{-\lambda}\right)$,*

where $0 < \beta < 1/2$, $\lambda > 0$ and f is a Lipschitz function. Then, the differential equation

$$z'(t) = f(t, z(t)),$$

has a unique solution z with given initial condition $z(0)$ and $Y(\lfloor tN \rfloor)/N$ converges in probability towards z for the topology of uniform convergence.

Proof. By regularity of the solution of the differential equation with respect to the initial condition, it suffices to treat the case where $Y(0)/N \equiv z(0)$. Let $1 < \varepsilon < \frac{1-\beta}{\beta}$ which exists by our hypothesis on β. Let $w = N^{(1+\varepsilon)\beta}$ and fix $\alpha \in \left(\frac{1+\varepsilon}{2}\beta, \varepsilon\beta\right)$. Let also $0 < \lambda' < \lambda$. We prove by induction the following property for all $0 \leq i \leq N/w$:

$$\mathbb{P}\left(|Y(iw) - z(iw/N)N| > i\left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'} + N^{2(1+\varepsilon)\beta-2\lambda}\right)\right) \leq 2i\exp\left(-\frac{N^{2\alpha-(1+\varepsilon)\beta}}{2}\right). \quad (6.29)$$

Note that the lower bound in the probability tends to zero after dividing by N for all $i \leq N/w$ and that the probability tends to zero by our hypothesis.

The initialization is satisfied by the choice we made for $Y(0)$.

Suppose that the property is verified for $0 \leq i \leq N/w - 1$. Rewrite

$$\begin{aligned} Y((i+1)w) - z((i+1)w/N)N = \quad & Y((i+1)w) - Y(iw) - wf(iw/N, Y(iw)/N) && (6.30)\\ & + Y(iw) - z(iw/N)N && (6.31)\\ & + z((i+1)w/N)N - z(iw/N)N - wf(iw/N, Y(iw)/N). && (6.32) \end{aligned}$$

By our induction hypothesis, the second term can be bounded as in (6.29). We now claim that it suffices to establish the following inequality:

$$\mathbb{P}\left(\left(Y((i+1)w) - Y(iw) - wf(iw/N, Y(iw)/N)\right) > \left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'}\right)\right) \leq 2\exp\left(-\frac{N^{2\alpha-(1+\varepsilon)\beta}}{2}\right). \quad (6.33)$$

Indeed, using inequalities (6.29) and (6.33) and the fact that $|(6.32)|$ is bounded by $(N/w)^2 = O(N^{2(1+\varepsilon)-2\lambda})$ by a Taylor expansion, we would obtain that:

$$\begin{aligned} \mathbb{P}\Big(|Y((i+1)w) - z((i+1)w/N)N| > (i+1)\left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'} + N^{2(1+\varepsilon)\beta-2\lambda}\right)\Big) & \\ \leq \mathbb{P}\left(|Y(iw) - z(iw/N)N| > i\left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'} + N^{2(1+\varepsilon)\beta-2\lambda}\right)\right) & \\ + \mathbb{P}\left(|Y((i+1)w) - Y(iw) - wf(iw/N, Y(iw)/N)| > \left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'}\right)\right) & \\ \leq 2(i+1)\exp\left(-\frac{N^{2\alpha-(1+\varepsilon)\beta}}{2}\right). & \end{aligned}$$

Let us prove (6.33). By the trend assumption, namely point 3. in Theorem 13, there exists a function $g(N)$ such that $g(N) = O(N^{-\lambda})$ and such that the process

$$\{Y(iw+k) - Y(iw) - kf(iw/N, Y(iw)/N) - kg(N)\}_{1\leq k\leq w}$$

is a supermartingale with increments bounded by N^β. Using Azuma-Hoeffding inequality with $k = w$, this implies that:

$$\mathbb{P}\left(\left(Y((i+1)w) - Y(iw) - wf(iw/N, Y(iw)/N)\right) > \left(N^{\alpha+\beta} + wg(N)\right)\right) \leq \exp\left(-\frac{1}{2w}\frac{N^{2\alpha+2\beta}}{N^{2\beta}}\right).$$

Since $\lambda' < \lambda$ and since $wg(N) = O(N^{(1+\varepsilon)\beta)-\lambda})$, we have proved that:

$$\mathbb{P}\left(\left(Y\left((i+1)w\right) - Y(iw) - wf(iw/N, Y(iw)/N)\right) > \left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'}\right) \right) \leq \exp\left(-\frac{N^{2\alpha-(1+\varepsilon)\beta}}{2}\right).$$

Using a similar argument, one can obtain the same bound on the probability that $\left(Y\left((i+1)w\right) - Y(iw) - wf(iw/N, Y(iw)/N)\right) < \left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'}\right)$. Therefore:

$$\mathbb{P}\left(\left(Y\left((i+1)w\right) - Y(iw) - wf(iw/N, Y(iw)/N)\right) > \left(N^{\alpha+\beta} + N^{(1+\varepsilon)\beta-\lambda'}\right) \right) \leq 2\exp\left(-\frac{N^{2\alpha-(1+\varepsilon)\beta}}{2}\right),$$

which is exactly (6.33). This ends the proof of Theorem 13. □

The main point here is the exponential bound (6.29). By a union bound, this enables us to state a version of Wormald's result for a polynomial number of Markov Chains driven by an infinite number of differential equations, which is what is needed for our work.

Corollary 14. *Let $a > 0$. For all $N \geq 1$ and all $1 \leq k \leq N^a$, let $Y_k(i) = Y_k^{(N)}(i)$ be a Markov chain with respect to a filtration $\{\mathcal{F}_i\}_{i\geq 1}$. Suppose that, for all $k \geq 1$, there exists a function f_k such that:*

- *$Y_k(0)/N$ converges to $z_k(0)$ in probability;*
- *$|Y_k(i+1) - Y_k(i)| \leq N^\beta$;*
- *$\mathbb{E}\left[Y_k(i+1) - Y_k(i) \,\middle|\, \mathcal{F}_i\right] = f_k\left(\frac{i}{N}, \frac{(Y_k(i))_{1\leq k\leq N^a}}{N}\right) + O\left(N^{-\lambda}\right)$,*

where $0 < \beta < 1/2$, $\lambda > 0$. Suppose that the following infinite system of differential equations with initial conditions $(z_k(0))_{k\geq 1}$ has a unique solution $(z_k)_{k\geq 1}$:

$$\forall k \geq 1, \quad z_k'(t) = f_k(t, (z_k(t))_{k\geq 1}).$$

Then, for all $k \geq 1$, $Y_k(\lfloor tN \rfloor)/N$ converges in probability towards z_k for the topology of uniform convergence.

Chapter 7

Long Induced Paths in a Configuration Model

This chapter corresponds to the ongoing work [EFMN20].

In an article published in 1987 in Combinatorica [FJ87], Frieze and Jackson established a lower bound on the length of the longest induced path (and cycle) in a sparse random graph. Their bound is obtained through a rough analysis of a greedy algorithm. In the present work, we provide a sharp asymptotic for the length of the induced path constructed by their algorithm. To this end, we introduce an alternative algorithm that builds the same induced path and whose analysis falls into the framework of a previous work by the authors on depth-first exploration of a configuration model [EFMN19]. We also analyze an extension of our algorithm that mixes depth-first and breadth-first explorations and generates m-induced paths.

7.1 Introduction

In this paper, we are interested in the existence of long induced cycles and long induced paths in random (multi-)graphs. An induced path of length $k \geq 1$ in a (multi-)graph $G = (V, E)$ is a sequence $v_1, \dots, v_k$ of distinct vertices of the graph such that, for any $1 \leq i, j \leq k$, v_i and v_j are neighbors in the graph G (*i.e.* $\{v_i, v_j\} \in E$) if and only if $|i - j| = 1$. Similarly, an induced cycle is a cycle of distinct vertices of the graph such that two non-consecutive vertices are not linked by an edge. Induced cycles are often also called *holes* of the graph. Our main result, Theorem 14, is a lower bound on the length of induced paths and cycles for random graphs constructed by the configuration model in the supercritical sparse regime. This length is linear in the size of the graph.

The configuration model was introduced by Bollobás in 1980 [Bol80] and can be defined as follows: Let $N \geq 1$ be an integer and let $d_1, \dots, d_N \in \mathbb{Z}_+$ be such that $d_1 + \cdots + d_N$ is even. We interpret d_i as a number of half-edges attached to vertex i. Then, the configuration model $\mathscr{C}((d_i)_{1 \leq i \leq N})$ associated to the sequence $(d_i)_{1 \leq i \leq N}$ is the random multigraph with vertex set $\{1, \dots, N\}$ obtained by a uniform matching of these half-edges. If $d_1 + \cdots + d_N$ is odd, we change d_N into $d_N + 1$ and do the same construction.

We are going to study sequences of configuration models whose degree sequences $\mathbf{d}^{(N)}$ satisfy classical hypothesis. The first assumption corresponds to the sparsity of the graphs:

- For every $N \geq 1$, $\mathbf{d}^{(N)} = (d_1^{(N)}, \dots, d_N^{(N)}) \in \mathbb{Z}_+^N$, and there exists $\boldsymbol{\pi}$ a probability measure

on $\mathbb{Z}_+$ with finite second moment such that

$$\forall k \geq 0, \quad \frac{1}{N}\sum_{j=1}^{N} \mathbf{1}_{d_j^{(N)}=k} \underset{N\to+\infty}{\longrightarrow} \boldsymbol{\pi}(\{k\}). \tag{A1}$$

Such graphs are known to exhibit a phase transition for the existence of a unique macroscopic connected component depending on the properties of the limiting distribution $\boldsymbol{\pi}$, see [MR95, MR98, JL09]. Our second assumption is that our configuration models are supercritical, meaning that they have a unique giant component, so that long induced paths and cycles can have linear length. To state this assumption, we need the following notation: for every $s \in [0,1]$ define

$$f_{\boldsymbol{\pi}}(s) = \sum_{i\geq 0} \boldsymbol{\pi}(\{i\}) s^i \quad \text{and} \quad \hat{f}_{\boldsymbol{\pi}}(s) = \frac{f'_{\boldsymbol{\pi}}(s)}{f'_{\boldsymbol{\pi}}(1)}.$$

Our second assumption is then:

- The probability measure $\boldsymbol{\pi}$ is supercritical in the following sense:

$$\hat{f}'_{\boldsymbol{\pi}}(1) > 1. \tag{A2}$$

Denote by $\rho_{\boldsymbol{\pi}}$ the smallest positive solution in $(0,1]$ of the fixed point equation

$$1 - \rho_{\boldsymbol{\pi}} = \hat{f}_{\boldsymbol{\pi}}(1 - \rho_{\boldsymbol{\pi}}) \tag{7.1}$$

and write

$$\xi_{\boldsymbol{\pi}} = 1 - f_{\boldsymbol{\pi}}(1 - \rho_{\boldsymbol{\pi}}). \tag{7.2}$$

Under Assumptions **(A1)** and **(A2)**, the giant connected component of $\mathscr{C}(\mathbf{d}^{(N)})$ has size $\xi_{\boldsymbol{\pi}} N + o(N)$ with probability tending to 1 as $N \to \infty$.

Finally, we make two additional technical assumptions on the degree sequences $(\mathbf{d}^{(N)})_N$:

- The following convergence holds:

$$\lim_{N\to+\infty} \frac{{d_1^{(N)}}^2 + \cdots + {d_N^{(N)}}^2}{N} = \sum_{k\geq 0} k^2 \boldsymbol{\pi}(\{k\}). \tag{A3}$$

- There exists $\gamma > 2$ such that:

$$\max\left\{d_1^{(N)}, \ldots, d_N^{(N)}\right\} \leq N^{1/\gamma}. \tag{A4}$$

We are now ready to state our main result:

Theorem 14. *Let $\boldsymbol{\pi}$ be a probability measure on $\mathbb{Z}_+$ with generating series $f_{\boldsymbol{\pi}}$ and $(\mathscr{C}(\mathbf{d}^{(N)}))_{N\geq 1}$ be a configuration model with supercritical asymptotic degree distribution $\boldsymbol{\pi}$ satisfying assumptions* **(A1)** **(A2)**, **(A3)** *and* **(A4)**. *Denote by $\mathcal{H}_N$ be the length of the longest induced cycle or induced path in $\mathscr{C}(\mathbf{d}^{(N)})$.*

Let α_c be the smallest positive solution of the equation

$$\frac{f''_{\boldsymbol{\pi}}\left(f_{\boldsymbol{\pi}}^{-1}(1-\alpha)\right)}{f'_{\boldsymbol{\pi}}(1)} = 1.$$

For every $\alpha \in [0, \alpha_c]$ and $s \in [0,1]$, we define the following functions:

$$g(\alpha, s) = \frac{1}{1-\alpha} f_\pi \left(f_\pi^{-1}(1-\alpha) - (1-s) \frac{f_\pi' \left(f_\pi^{-1}(1-\alpha) \right)}{f_\pi'(1)} \right) \quad \text{and} \quad \hat{g}(\alpha, s) = \frac{\partial_s g(\alpha, s)}{\partial_s g(\alpha, 1)}.$$

There exists an implicit function $\alpha(\rho)$ defined on $[0, \rho_\pi]$ such that $1 - \rho = \hat{g}(\alpha(\rho), 1-\rho)$.

Then,

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\frac{\mathcal{H}_N}{N} \geq \int_0^{\rho_\pi} \frac{u\, \alpha'(u)}{\partial_s \hat{g}(\alpha(u), 1)} \mathrm{d}u - \varepsilon \right) \underset{N \to +\infty}{\longrightarrow} 1.$$

In addition, this bound still holds if we condition the graphs $\mathscr{C}(\mathbf{d}^{(N)})$ to be simple by standard arguments (see for example Corollary 7.17 of [vdH17]).

Although the formulation of Theorem 14 is implicit, explicit computations are easy in specific models. Indeed, for d-regular random graphs with $d \geq 3$ we find:

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\frac{\mathcal{H}_N}{N} \geq \frac{d}{2(d-1)} \left(1 - \int_0^1 \left(\frac{1 - x^{\frac{1}{d-1}}}{1-x} \right)^{\frac{2}{d-2}} \mathrm{d}x - \varepsilon \right) \right) \underset{N \to +\infty}{\longrightarrow} 1.$$

For Erdős-Rényi random graphs with connection probability c/N ($c > 1$), denoting by ρ_c the smallest positive solution of $1 - \rho_c = \exp(-c\rho_c)$, we find:

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\frac{\mathcal{H}_N}{N} \geq \frac{\rho_c}{-\ln(1-\rho_c)} \left(\gamma + \rho_c + \ln(-\ln(1-\rho_c)) - \mathrm{Li}_2(1-\rho_c) \right) - \varepsilon \right) \underset{N \to +\infty}{\longrightarrow} 1,$$

where $\gamma \approx 0.577\ldots$ is Euler's constant and Li_2 is the dilogarithm function.

To the best of our knowledge, the only other general bound on $\mathcal{H}_N$ is due to Frieze and Jackson in [FJ87] in the setting where the degrees of the graph are bounded below by 3 and uniformly bounded above but do not necessarily satisfy our assumptions. For instance their lower bound for 3-regular graphs is approximately equal to 0.07 while our bound is approximately equal to 0.45. Their proof relies on a greedy algorithm which allows them to construct a macroscopic induced path in the graph. However, they do not establish the exact asymptotic length of this path, but rather a lower bound. Building on this result they also obtained a lower bound on $\mathcal{H}_N/N$ for Erdős-Rényi random graphs having large fixed averaged degree.

To prove Theorem 14, we introduce a different algorithm that is amenable to a detailed analysis with the framework developed in [EFMN19] by the authors. This allows us to exhibit an induced path with explicit macroscopic length. As we will see in Section 7.4, it turns out that both algorithms provide the same long induced path (and cycle) up to $o(N)$ vertices.

The paper is organized as follows. In Section 7.2 we present an algorithm that constructs a configuration model and spanning trees of its connected components for which ancestral lines form induced paths in the graph. We also state a result for the limiting profile of the spanning forest constructed by the algorithm in Theorem 11 from which Theorem 14 follows easily. In Section 7.3, we give a detailed analysis of our algorithm using the framework of our previous works [EFM20, EFMN19]. In Section 7.4, we show the induced path constructed by our algorithm is roughly the same as the induced path constructed by the algorithm of Frieze and Jackson [FJ87]. Finally, in Section 7.5, we analyse an extension of our algorithm. This extension mixes depth-first and breadth-first explorations and constructs $m-$induced paths for any fixed m.

7.2 Constructing the graph while discovering induced paths

We now introduce an algorithm which, from the knowledge of the sequence of degree $\mathbf{d}^{(N)}$, simultaneously constructs a configuration model $\mathscr{C}(\mathbf{d}^{(N)})$ together with an exploration of it. The latter exploration, which is a modification of the depth-first exploration, is designed to discover long induced paths in the graph. At each step n of the construction, we consider the following objects, defined by induction:

- A_n, the set of active vertices, is an ordered list of pairs $(v, \mathbf{m}_v)$ where v is a vertex of V_N and $\mathbf{m}_v$ is an ordered list of elements of the form $(u, (u^1, \dots, u^l))$ where u is a vertex that will be matched to v and $u^1, \dots, u^l$ vertices that will be matched to u during the exploration.
- S_n, the set of sleeping vertices, which consists of vertices that do not appear in A_n.
- R_n, the set of retired vertices, which consists of vertices that appear neither in A_n nor S_n.

At the initial step $n = 0$ of the algorithm, we choose a vertex v uniformly at random and pair each of its $d_v^{(N)}$ half-edges to uniform half-edges of the graph. Denote by $v_1, \dots, v_l$ the corresponding vertices. For each $1 \leq i \leq l$, we successively match the half-edges of v_i to uniform half-edges of the graph and denote by $v_i^1, \dots, v_i^{k_i}$ the corresponding vertices (without repeat). Let $\mathbf{m}_v = \Big((v_1, (v_1^1, \dots, v_1^{k_1})), \dots, (v_l, (v_l^1, \dots, v_l^{k_l}))\Big)$ and set

$$\begin{cases} A_0 = ((v, \mathbf{m}_v)), \\ S_0 = \mathrm{V}_N \setminus \{v, v_1, \dots, v_l\}, \\ R_0 = \emptyset. \end{cases}$$

Suppose that A_n, S_n and R_n are constructed. Three cases are possible:

1. If $A_n = \emptyset$, the algorithm has just finished exploring and building a connected component of $\mathscr{C}(\mathbf{d}^{(N)})$. In this case, we select v_{n+1} uniformly at random inside S_n and define $\mathbf{m}_{v_{n+1}}$ exactly as before, except that the matched vertices are in S_n. Denoting $v_{n+1}^1, \dots, v_{n+1}^l$ the vertices of S_n matched to v_{n+1}, we then set:

$$\begin{cases} A_{n+1} = (v_{n+1}, \mathbf{m}_{v_{n+1}}), \\ S_{n+1} = S_n \setminus \{v_{n+1}, v_{n+1}^1, \dots, v_{n+1}^l\}, \\ R_{n+1} = R_n. \end{cases}$$

2. If $A_n \neq \emptyset$ and if its last element $(v, \mathbf{m}_v)$ is such that $\mathbf{m}_v = \emptyset$, the exploration backtracks and we set:

$$\begin{cases} A_{n+1} = A_n - (v, \mathbf{m}_v), \\ S_{n+1} = S_n \\ R_{n+1} = R_n \cup \{v\}. \end{cases}$$

3. If $A_n \neq \emptyset$ and if its last element $(v, \mathbf{m}_v)$ is such that $\mathbf{m}_v \neq \emptyset$, we denote by $(v_{n+1}, (v_{n+1}^1, \dots, v_{n+1}^l))$ the first element of $\mathbf{m}_v$. In that case, the exploration goes to v_{n+1} and we construct $\mathbf{m}_{v_{n+1}}$ as before using the v_{n+1}^i's and the vertices of S_n that are matched to them. We finally set:

$$\begin{cases} A_{n+1} = A_n + (v_{n+1}, \mathbf{m}_{v_{n+1}}), \\ S_{n+1} = S_n \setminus \{v_{n+1}^1, \dots, v_{n+1}^l\}, \\ R_{n+1} = R_n. \end{cases}$$

In words, this algorithm is an interpolation between depth-first and breadth-first explorations of the graph. More precisely, our procedure first constructs the 1-neighborhood of the current vertex v, which consists in some vertices $v_1, \dots, v_l$ and it then matches sequentially the half-edges of $v_1, \dots, v_l$ that are not yet matched. In particular, this ensures that every ancestral line of the trees constructed by the algorithm is in fact an induced path in the graph, because the discovered vertices v_i^j, $1 \leq i \leq j$, $1 \leq j \leq k_i$, are distinct from $v_1, \dots, v_l$.

Since each matching of half-edges is uniform during the construction, this algorithm constructs a random graph $\mathscr{C}(\mathbf{d}^{(N)})$. Furthermore, at each step n, the subgraph of $\mathscr{C}(\mathbf{d}^{(N)})$ induced by the vertices of S_n is a configuration model. The sequence of vertices corresponding to the first component of the last element of A_n provides a spanning tree of each connected component of the graph together with a contour process of each of these trees.

We will denote by $(X_n)_{0\leq n\leq 2N}$ the concatenated contour processes of the successive covering trees constructed by the algorithm. Our main result, Theorem 14, will be an easy consequence of the following fluid limit for the process $(X_n)_{0\leq n\leq 2N}$:

Theorem 15. *Recall the definition of ρ_π and ξ_π given in equations (7.1) and (7.2), and the definition of the functions $\alpha(\rho)$, $g(\alpha, s)$ and $\hat{g}(\alpha, s)$ given in Theorem 14. Under assumptions* **(A1)**, **(A2)**, **(A3)** *and* **(A4)**, *the following limit holds in probability for the topology of uniform convergence:*

$$\forall u \in [0,2], \quad \lim_{N\to\infty} \frac{X_{\lceil uN \rceil}}{N} = h(u),$$

where the function h is continuous on $[0,2]$, null on the interval $[2\xi_\pi, 2]$ and defined hereafter on the interval $[0, 2\xi_\pi]$.

The graph $(u, h(u))_{u\in[0,2\xi_\pi]}$ can be divided into a first increasing part and a second decreasing part. These parts are respectively parametrized for $\rho \in [0, \rho_\pi]$ by :

$$\begin{cases} x^{\uparrow}(\rho) & := \displaystyle\int_\rho^{\rho_\pi} \frac{(2-r)\,\alpha'(r)}{\partial_s \hat{g}(\alpha(r), 1)} \mathrm{d}r, \\ y^{\uparrow}(\rho) & := \displaystyle\int_\rho^{\rho_\pi} \frac{r\,\alpha'(r)}{\partial_s \hat{g}(\alpha(r), 1)} \mathrm{d}r, \end{cases}$$

for the increasing part and

$$\begin{cases} x^{\downarrow}(\rho) := x^{\uparrow}(\rho) + 2\,(1-\alpha(\rho))\left(1 - g(\alpha(\rho), 1-\rho)\right), \\ y^{\downarrow}(\rho) := y^{\uparrow}(\rho), \end{cases}$$

for the decreasing part.

Theorem 14 is obtained by computing the maximal value of h, which is given by $y^{\uparrow}(0)$. The proof of Theorem 15 is an adaptation of the article [EFMN19] by the authors and is the object of the next section. It relies on Wormald's differential equations method [Wor95] via the study of ladder times of the exploration and the law of the graph induced by the sleeping vertices at these times.

7.3 Analysis of the algorithm

The overall strategy follows the guidelines of the previous work [EFMN19]. In particular, we start by identifying a good event that makes possible a decomposition of the exploration at

ladder times. For every $n \in \{0, \dots, 2N\}$, let $D_n^{(N)}$ be the degree of a uniform vertex in the graph induced by S_n. For every $\varepsilon > 0$ we define

$$n_\epsilon = n_\varepsilon^{(N)} = \sup \left\{ n \in [\![0, 2N]\!] : \forall m \in [\![0, n]\!], \frac{\mathbb{E}[D_n^{(N)}(D_n^{(N)} - 1)]}{\mathbb{E}[D_n^{(N)}]} > 1 + \varepsilon \right\}.$$

For $n < n_\varepsilon$, the subgraphs induced by S_n are all supercritical. For $0 < \delta < 1/2$, let $\mathbf{G}_\varepsilon = \mathbf{G}_\varepsilon^{(N)}(\delta)$ be the event that, for all $n < n_\varepsilon$,

- there is at least one connected component with size greater than $N^{1-\delta}$ in the graph induced by S_n;
- there is no connected component of size between N^δ and $N^{1-\delta}$ in the graph induced by S_n.

Under our assumptions **(A1)**, **(A2)**, **(A3)** and **(A4)**, we have for every $\lambda > 0$,

$$\mathbb{P}(\mathbf{G}_\varepsilon) = 1 - \mathcal{O}(N^{-\lambda}). \tag{7.3}$$

7.3.1 Ladder times

Fix $\delta \in (0, 1)$. Let $T_0 = 0$ and define, for $k \in \{0, \dots, K\}$,

$$T_{k+1} := \min \left\{ i > T_k,\ X_i = k + 1 \text{ and } \forall i \le j \le i + N^\delta,\ X_j \ge k + 1 \right\},$$

where K is the last index for which this definition makes sense (i.e. the set for which the min is taken is not empty). Of course, this sequence of times will only be useful to analyse our algorithm when K is of macroscopic order, which is indeed the case on the event $\mathbf{G}_\varepsilon$. Indeed, as long as $T_k < n_\varepsilon$, we can define T_{k+1} on the event $\mathbf{G}_\varepsilon$. Therefore we set

$$K_\varepsilon = \sup\{k \in [\![0, K]\!] : T_k < n_\varepsilon\}. \tag{7.4}$$

Thanks to (6.2), we have $K_\varepsilon < K$ with probability $1 - \mathcal{O}(N^{-\lambda})$.

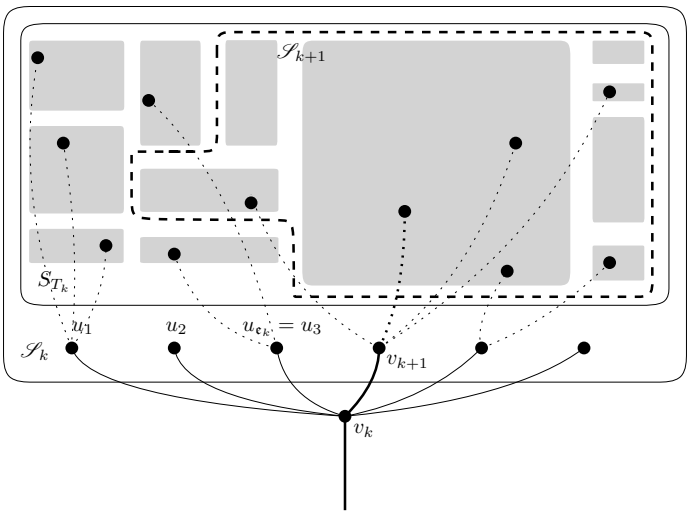


Figure 7.1 – Structure of the remaining graph at a ladder time. The first half edges of $\mathbf{v}_k$ are numbered according to their matching order during the construction.

For all $k \in \{0, \dots, K\}$, let $\mathbf{v}_k$ be the vertex corresponding to the first component of the last element of A_{T_k}. There exists some $l \geq 1$ such that the sequence $\mathbf{m}_{\mathbf{v}_k}$ can be written

$$\mathbf{m}_{\mathbf{v}_k} = \left((u_1, (u_1^1, \dots, u_1^{k_1})), \dots, (u_l, (u_l^1, \dots, u_l^{k_l})) \right), \tag{7.5}$$

and we will denote by $\mathfrak{e}_k$ the index in $\{1, \dots, l\}$ such that $u_{\mathfrak{e}_k+1} = \mathbf{v}_{k+1}$. We consider $\mathscr{S}_k$ the graph induced by the set of vertices S_{T_k-1}. See Figure 7.1 for an illustration of these definitions. As it turns out, the analysis of the structural changes of $\mathscr{S}_k$ will be crucial for our purpose. For instance, the difference between $T_{k+1} - T_k$ is equal to the time spent exploring the connected components associated to the vertices $u_1, \dots, u_{\mathfrak{e}_k}$ inside the graph $\mathscr{S}_k$.

7.3.2 Analysis of the graphs $\mathscr{S}_k$

For all $k < K$, let $N_i(k)$ be the number of vertices of degree i in $\mathscr{S}_k \cup \{\mathbf{v}_k\}$ which are different from $\mathbf{v}_k$. Then, by definition of the exploration, the graph $\mathscr{S}_k$ has the law of a configuration model with vertex degrees given by the sequence $(\widetilde{N}_i(k))_{i\geq 0}$. Using the notation of (7.5), the contribution to $\widetilde{N}_i(k)$ of the edges belonging to v_k is given by

$$\sum_{p=1}^{l} \left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(u_p)=i} + \mathbf{1}_{\deg_{\mathscr{S}_k}(u_p)=i-1} \right),$$

and we write

$$\widetilde{N}_i(k) = N_i(k) + \sum_{p=1}^{l} \left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(u_p)=i} + \mathbf{1}_{\deg_{\mathscr{S}_k}(u_p)=i-1} \right).$$

Assumption (**A4**) ensures that $\widetilde{N}_i(k) - N_i(k)$ is of order $o(N)$ uniformly along the whole exploration with high probability. Henceforth, we focus our analysis on the $N_i(k)$'s whose analysis is more clear. This evolution is ideed given by:

$$N_i(k+1) - N_i(k) = -V_i(\mathscr{S}_k \setminus \mathscr{S}_{k+1}) \tag{7.6}$$

$$+ \sum_{p=\mathfrak{e}_k+2}^{l} \sum_{q=1}^{k_p} \left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(u_p^q)=i} + \mathbf{1}_{\deg_{\mathscr{S}_k}(u_p^q)=i+1} \right), \tag{7.7}$$

where $V_i(S)$ stands for the number of vertices with degree i in S and l is defined in Equation (7.5). Indeed, the first contribution corresponds to the complete removal of vertices belonging to $\mathscr{S}_k$ but not $\mathscr{S}_{k+1}$. The second contribution corresponds to the removal of edges connecting the vertices $u_{\mathfrak{e}_k+2}, \dots, u_l$ to $\mathscr{S}_{k+1}$.

The crucial step of the proof is the asymptotic analysis of the variables T_k and $N_i(k)$ for large N. This is the object of the forthcoming Theorem 16. In order to state it, we first need to introduce some notation.

Let $(z_i)_{i\geq 0} \in \mathbb{R}^{\mathbb{Z}_+}$ be such that $\sum_{i\geq 0} z_i \leq 1$ and $\sum_{k\geq 0} i z_i < \infty$. for any $i \geq 0$ let $\hat{z}_i = (i+1)z_i / \sum_j j z_j$ and define:

$$\begin{cases} g_{(z_i)_{i\geq 0}}(s) &= \sum_{i\geq 0} \frac{z_i}{\sum_{l\geq 0} z_l} s^i \\ \hat{g}_{(z_i)_{i\geq 0}}(s) &= \sum_{i\geq 0} \hat{z}_i s^i = \frac{g'_{(z_i)_{i\geq 0}}(s)}{g'_{(z_i)_{i\geq 0}}(1)} \end{cases} \tag{7.8}$$

respectively the generating series associated to $(z_k)_{k\geq 0}$ and its sized biased version. Let also $\rho_{(z_i)_{i\geq 0}}$ be the largest solution in $[0, 1]$ of

$$1 - s = \hat{g}_{(z_i)_{i\geq 0}}(1 - s). \tag{7.9}$$

Remark 15. *Since $\hat{g}$ is the generating function of a probability distribution on the integers, it is convex on $[0,1]$. Therefore, Equation (7.9) has a positive solution in $(0,1]$ if and only if $\hat{g}'(1) > 1$, which is equivalent to $\frac{\sum_{l\geq 1}(l-1)lz_l}{\sum_{l\geq 1} lz_l} > 1$.*

We also define the following functions:

$$f(z_0, z_1, \ldots) = \frac{2 - \rho_{(z_i)_{i\geq 0}}}{\rho_{(z_i)_{i\geq 0}}} \tag{7.10}$$

$$f_i(z_0, z_1, \ldots) = -\frac{1}{\rho_{(z_j)_{j\geq 0}}} \frac{iz_i}{\sum_{j\geq 0} kz_j} + \frac{1}{\rho_{(z_j)_{j\geq 0}}} \left(\frac{\sum_{j\geq 0}(j-1)jz_j}{\sum_{n\geq 0} jz_j} - 1 \right) \times \left(-\frac{iz_i}{\sum_{j\geq 0} jz_j} + \frac{\sum_{j\geq 0}(j-1)jz_j}{\sum_{n\geq 0} jz_j} \left(\frac{(i+1)z_{i+1}}{\sum_{j\geq 0} jz_j} - \frac{iz_i}{\sum_{j\geq 0} jz_j} \right) \right). \tag{7.11}$$

The asymptotic behaviour of the variables T_k and $N_i(k)$ will be driven by the solution of an infinite system of differential equations whose uniqueness and existence is provided by the following lemma.

Lemma 11. *Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_i)_{i\geq 0} \in [0,1]^{\mathbb{N}}$ be a probability measure which satisfies the supercriticality assumption (**A2**). Then, the following system of differential equations has a unique solution which is well defined on $[0, t_{\max})$ for some $t_{\max} > 0$:*

$$\begin{cases} \frac{\mathrm{d}z_i}{\mathrm{d}t} &= f_i(z_0, z_1, \ldots); \\ z_i(0) &= \boldsymbol{\pi}_i. \end{cases} \tag{S}$$

Proof. Let

$$F(t,s) := \sum_{i\geq 0} z_i(t) s^i \tag{7.12}$$

be the generating series associated to the z_i's, which is well defined for all $s \in [0,1]$ and all $t \in [0, t'_{\max})$, where $[0, t'_{\max})$ is the maximal interval where the functions z_i are well defined. Then, it is easy to check that F satisfies the following partial differential equation:

$$\frac{\partial F}{\partial t}(t,s) = \frac{1}{\rho_{(z_j(t))_{j\geq 0}}} \frac{\frac{\partial^2 F}{\partial s^2}(t,1)}{\frac{\partial F}{\partial s}(t,1)} \frac{\partial F}{\partial s}(t,s) \left((1-s) \frac{\frac{\partial^2 F}{\partial s^2}(t,1)}{\frac{\partial F}{\partial s}(t,1)} - 1 \right).$$

Using the appropriate time change, we end up with a new generating series $f(t,s) = \sum_{j\geq 0} \zeta_j(t) s^i$ satisfying

$$\frac{\partial f}{\partial t}(t,s) = \frac{\frac{\partial f}{\partial s}(t,s)}{\frac{\partial f}{\partial s}(t,1)} \left((1-s) \frac{\frac{\partial^2 f}{\partial s^2}(t,1)}{\frac{\partial f}{\partial s}(t,1)} - 1 \right). \tag{7.13}$$

Notice that, up to a time change, this corresponds to the differential equation that already appeared in [EFMN19] – see the beginning of the proof of Proposition 1. It was proved in Section 6.2 of [EFMN19] that it has a unique solution under our assumptions. □

We are now ready to state the main result of this section.

Theorem 16. *Fix $\varepsilon > 0$. With high probability, for all $k \leq K_\varepsilon$:*

$$T_k = Nz\left(\frac{k}{N}\right) + o(N)$$
$$N_i(k) = Nz_i\left(\frac{k}{N}\right) + o(N),$$

where $(z_0, z_1, \dots)$ is the unique solution of (S) and z is the unique solution of $\frac{\mathrm{d}z}{\mathrm{d}t} = f(z_0, z_1, \dots)$ with initial condition given by $z(0) = 0$.

In addition, if $w(k)$ denotes the number of vertices that are not in the graph $\mathscr{S}_k$, then

$$w(k) = N\tilde{z}\left(\frac{k}{N}\right) + o(N),$$

where $\tilde{z}$ satisfies $\tilde{z}'(t) = \hat{g}'_{(z_j(t))_{j\geq 0}}(1)/\rho_{(z_j(t))_{j\geq 0}}$ and $\tilde{z}(0) = 1$.

Proof. Our main tool is an adaptation of Wormald's differential equations method, which is the content of Corollary 2 of [EFMN19]. To apply this result we need to check the following two points:

1. There exists $0 < \beta < 1/2$ such that with high probability for all $k \leq K_\varepsilon$,
$$|T_{k+1} - T_k| \leq N^\beta \text{ and for all } i \geq 0,\ |N_i(k+1) - N_i(k)| \leq N^\beta.$$
2. We denote by $(\mathcal{F}_k)_{k\geq 0}$ the canonical filtration associated to the sequence $((N_i(k))_{i\geq 0})_{k\geq 0}$. There exists $\lambda > 0$ such that for every k and n,
$$\mathbb{E}[T_{k+1} - T_k \,|\, \mathcal{F}_k] = f\left(\frac{N_0(k)}{N}, \frac{N_1(k)}{N}, \dots\right) + O\left(N^{-\lambda}\right), \tag{7.14}$$
$$\mathbb{E}[N_i(k+1) - N_i(k) \,|\, \mathcal{F}_k] = f_i\left(\frac{N_0(k)}{N}, \frac{N_1(k)}{N}, \dots\right) + O\left(N^{-\lambda}\right). \tag{7.15}$$

The first point is a direct consequence of the definition of the times T_k's and of Equation (6.2) by chosing δ small enough.

We now turn to the second point. Since the computations are very similar to those made during the proof of Theorem 3 of [EFMN19], we will only sketch them and point out the corresponding details in [EFMN19]. For all $k \geq 0$, let

$$\begin{array}{lll} p_i = p_i(k) = \frac{N_i(k)}{\sum_{j\geq 0} N_j(k)} & ; & g_k = g_{(p_j)_{j\geq 0}} \\ \hat{p}_i = \hat{p}_i(k) = \frac{(i+1)p_{i+1}(k)}{\sum_{j\geq 0} j p_j(k)} & ; & \hat{g}_k = \hat{g}_{(p_j)_{j\geq 0}} = g_{(\hat{p}_j)_{j\geq 0}} \end{array}, \tag{7.16}$$

and let $\rho_k = \rho_{(p_j(k))_{j\geq 0}}$ be the largest solution in $[0,1]$ of $1 - s = \hat{g}_k(1-s)$.

Recall the notation of (7.5). With high probability, the first $\mathfrak{e}_k$ neighbors of $\mathbf{v}_k$ belong to distinct connected components of $\mathscr{S}_k$. Denoting $W^{(1)}, \dots, W^{(k)}$ these connected components, we deduce that

$$\begin{aligned} \mathbb{E}[T_{k+1} - T_k \,|\, \mathcal{F}_k] &= 1 + 2\mathbb{E}\left[\sum_{j=1}^{k} |W^{(j)}| \,\middle|\, \mathcal{F}_k\right] \\ &= \frac{2 - \rho_k}{\rho_k} + \mathcal{O}(N^{-\lambda}), \end{aligned}$$

where the last equality is the content of Equation (12) in [EFMN19]. This proves (7.14).

We now fix $i \geq 0$, $k \geq 0$, and prove (7.15) by examining separately the contributions (7.6) and (7.7). The contribution (7.6) corresponds to the removal of vertices of degree i inside the small components $W^{(j)}$'s, to the removal of $\mathbf{v}_{k+1}$, and to the removal of $u_{\mathfrak{e}_k+2}, \dots, u_l$. It is given by

$$\begin{aligned}&\mathbb{E}\left[V_i\left(\mathscr{S}_k \setminus \mathscr{S}_{k+1}\right) \mid \mathcal{F}_k\right] \\ &= \mathbb{E}\left[V_i\left(\cup_{j=1}^{\mathfrak{e}_k} W^{(j)}\right) \mid \mathcal{F}_k\right] + \mathbb{P}(\deg_{\mathscr{S}_k}(\mathbf{v}_{k+1}) = i \mid \mathcal{F}_k) + \mathbb{E}\left[\sum_{j=\mathfrak{e}_k+2}^{l} \mathbf{1}_{\{\deg_{\mathscr{S}_k}(u_j)=i\}} \mid \mathcal{F}_k\right]\end{aligned} \tag{7.17}$$

The terms $\mathbb{E}\left[V_i\left(\cup_{j=1}^{\mathfrak{e}_k} W^{(j)}\right) \mid \mathcal{F}_k\right]$ and $\mathbb{P}(\deg_{\mathscr{S}_k}(\mathbf{v}_{k+1}) = i \mid \mathcal{F}_k)$ were respectively computed in Equations (14) and (10) of [EFMN19]. They are given by

$$\mathbb{E}\left[V_i\left(\cup_{j=1}^{\mathfrak{e}_k} W^{(j)}\right) \mid \mathcal{F}_k\right] = \frac{\widehat{p}_{i-1}}{\rho_k}(1-\rho_k)^{i-1} + \mathcal{O}(N^{-\lambda}) \tag{7.18}$$

$$\mathbb{P}(\deg_{\mathscr{S}_k}(\mathbf{v}_{k+1}) = i \mid \mathcal{F}_k) = \frac{\hat{p}_{i-1}}{\rho_k}\left(1-(1-\rho_k)^{i-1}\right) + \mathcal{O}(N^{-\lambda}). \tag{7.19}$$

For the last term, we use that with high probability, $u_{\mathfrak{e}_k+2}, \dots, u_l$ are distinct vertices. All of them are connected to $\mathbf{v}_k$ through a uniform matching of half-edges. Therefore:

$$\begin{aligned}\mathbb{E}\left[\sum_{j=\mathfrak{e}_k+2}^{l} \mathbf{1}_{\{\deg_{\mathscr{S}_k}(u_j)=i\}} \,\middle|\, \mathcal{F}_k\right] &= \mathbb{E}\left[\deg_{\mathscr{S}_k}(\mathbf{v}_k) - \mathfrak{e}_k - 1 \mid \mathcal{F}_k\right] \hat{p}_{i-1} + \mathcal{O}(N^{-\lambda}) \\ &= \frac{1}{\rho_k}\left(\hat{g}_k'(1) - 1\right)\hat{p}_{i-1} + \mathcal{O}(N^{-\lambda}).\end{aligned} \tag{7.20}$$

where the computation of $\mathbb{E}\left[\deg_{\mathscr{S}_k}(\mathbf{v}_k) - \mathfrak{e}_k - 1 \mid \mathcal{F}_k\right]$ can be found in Equation (15) of [EFMN19].

Finally, the contribution (7.7) is given by

$$\begin{aligned}&\mathbb{E}\left[\sum_{p=\mathfrak{e}_k+2}^{l}\sum_{q=1}^{k_p}\left(-\mathbf{1}_{\deg_{\mathscr{S}_k}(u_p^q)=i} + \mathbf{1}_{\deg_{\mathscr{S}_k}(u_p^q)=i+1}\right)\,\middle|\,\mathcal{F}_k\right] \\ &\qquad= \mathbb{E}\left[\deg_{\mathscr{S}_k}(\mathbf{v}_k) - \mathfrak{e}_k - 1 \mid \mathcal{F}_k\right]\hat{g}_k'(1)\left(-\hat{p}_{i-1} + \hat{p}_i\right) + \mathcal{O}(N^{-\lambda}) \\ &\qquad= \frac{1}{\rho_k}\left(\hat{g}_k'(1) - 1\right)\hat{g}_k'(1)\left(-\hat{p}_{i-1} + \hat{p}_i\right) + \mathcal{O}(N^{-\lambda})\end{aligned} \tag{7.21}$$

Combining (7.18), (7.19), (7.20) and (7.21) we obtain

$$\mathbb{E}\left[N_i(k+1) - N_i(k) \mid \mathcal{F}_k\right] = \frac{\hat{g}_k'(1)}{\rho_k}\left(-\hat{p}_{i-1}\hat{g}_k'(1) + \hat{p}_i(-1+\hat{g}_k'(1))\right) + \mathcal{O}(N^{-\lambda}) \tag{7.22}$$

giving Equation (7.15).

We finally turn to the last claim of Theorem 16. For all $k < K_\varepsilon$, let $w(k)$ be the number of vertices that are not in the graph $\mathscr{S}_k$. Using the notation of (7.5), the evolution of $w(k)$ is given by

$$\begin{aligned}w(k+1) - w(k) &= V(\mathscr{S}_k \setminus \mathscr{S}_{k+1}) \\ &= V\left(\cup_{j=1}^{\mathfrak{e}_k} W^{(j)}\right) + 1 + (l - \mathfrak{e}_k - 1).\end{aligned}$$

Therefore,

$$\mathbb{E}\left[w(k+1) - w(k) \mid \mathcal{F}_k\right] = \frac{1-\rho_k}{\rho_k} + 1 + \frac{1}{\rho_k}\left(\hat{g}_k'(1) - 1\right) = \frac{1}{\rho_k}\hat{g}_k'(1). \tag{7.23}$$

Using Wormald's differential equations method as before we deduce that with high probability

$$w(k) = N\tilde{z}(k/N) + o(N),$$

where $\tilde{z}$ is the only solution of $y'(t) = \hat{g}'_{(z_j(t))_{j\geq 0}}(1)/\rho_{(z_j(t))_{j\geq 0}}$ with initial condition $y(0) = 1$. □

7.3.3 The asymptotic degree distribution inside $\mathscr{S}_k$

As in [EFMN19], Theorem 16 allows us to identify the laws of the graphs induced by sleeping vertices remaining after having explored a given proportion of vertices. Remarkably, these laws are the same as the laws appearing in the Depth First Search algorithm in Theorem 1 of [EFMN19]. We will see in the next section that the speed at which the graph is explored is however different.

Theorem 17. *Recall the definition of α_c and of $g(\alpha, s)$ given in Theorem 14. For every $\alpha \in [0, \alpha_c]$, let $\boldsymbol{\pi}_\alpha$ be the probability distribution on $\mathbb{Z}_+$ with generating series $g(\alpha, s)$. Then, for every $\alpha \in [0, \alpha_c]$, denoting $\tau^{(N)}(\alpha) = \inf\{n \geq 1,\ |S_n| \leq (1-\alpha)N\}$, the empirical degree distribution of the graph induced by the vertices of $S_{\tau^{(N)}(\alpha)}$ converges to $\boldsymbol{\pi}_\alpha$.*

Proof. Fix $\alpha \in [0, \alpha_c)$, by definition, a configuration model with asymptotic degree distribution $\boldsymbol{\pi}_\alpha$ is supercritical. Therefore,

$$\frac{\mathbb{E}\left[D_{\tau^{(N)}(\alpha)}(D_{\tau^{(N)}(\alpha)} - 1)\right]}{\mathbb{E}\left[D_{\tau^{(N)}(\alpha)}\right]} > 1 + \varepsilon$$

for some $\varepsilon > 0$ and N large enough, which ensures that $\tau^{(N)}(\alpha) < n_\varepsilon$.

By definition of $\tilde{z}$, this also ensures that $N\tilde{z}^{-1}(\alpha) < K_\varepsilon$ and that the proportion of explored vertices at time $T_{N\tilde{z}^{-1}(\alpha)}$ is $\alpha N + o(N)$. Therefore, $\tau^{(N)}(\alpha) = \inf\{n \geq 0,\ |S_n| \leq (1-\alpha)N\} = T_{N\tilde{z}^{-1}(\alpha)} + o(N)$ and for all $i \geq 0$,

$$\begin{aligned}\left|V_i\left(S_{\tau^{(N)}(\alpha)}\right)\right| &= N_i\left(T_{N\tilde{z}^{-1}(\alpha)} + o(N)\right)\\ &= Nz_i\left(\tilde{z}^{-1}(\alpha)\right) + o(N).\end{aligned}$$

It is easy to check that the generating series associated to the sequence of functions $(z_i \circ \tilde{z}^{-1})_{i\geq 0}$ is solution to the partial differential equation (7.13). □

7.3.4 Proof of Theorem 15

Let $N \geq 1$. By definition, for all $1 \leq k \leq K_\varepsilon$, the contour process of the tree constructed by our algorithm at time T_k is located at point (T_k, k). Furthermore, by Theorem 16,

$$(T_k, k) = N\left(z\left(\frac{k}{N}\right) + o(1), \frac{k}{N}\right).$$

Note that $|T_{k+1} - T_k| = o(N)$ and that, between two consecutive T_k's, the contour process cannot fluctuate by more than $o(N)$. Hence, after normalization by N, in its increasing part

the limiting contour process converges to the curve $(z(t), t)$ where t ranges from 0 to $t_{\max} = \sup\{t > 0,\, z'(t) < +\infty\}$. In the remaining, we provide a parametrization $(x^{\uparrow}(\rho), y^{\uparrow}(\rho))_{\rho\in(0,\rho_\pi]}$ of this increasing curve. Recalling the definition of f given in Equation 7.10, our Theorem 16 gives us

$$\frac{(x^{\uparrow})'(\rho)}{(y^{\uparrow})'(\rho)} = \frac{2}{\rho} - 1.$$

exactly as in [EFMN19].

To obtain a second equation involving $(x^{\uparrow})'(\rho)$ or $(y^{\uparrow})'(\rho)$, the paper [EFMN19] uses the implicit function $\alpha(\rho)$ given by the only solution of $1-\rho = \hat{g}(\alpha, 1-\rho)$. The link between $\alpha, x^{\uparrow}, y^{\uparrow}$ is however different in our setting. In order to establish this link, let us first notice that $\left(\rho(t) = \rho_{(z_i(t))_{i\geq 0}}\right)_t$ is the fluid limit of the survival probability $(\rho_k)_k$ as N tends to infinity, where we recall that the ρ_k's are defined in terms of the functions g_k (see Equation (7.16)). Indeed, since for all $t \geq 0$, we have

$$1 - \rho_{\lfloor tN \rfloor} = \hat{g}_{\lfloor tN \rfloor}(1 - \rho_{\lfloor tN \rfloor}),$$

the fact that the sequence of generating series $(\hat{g}_{\lfloor tN \rfloor}(s))_{N\geq 0}$ converges to $\partial_s F(t,s)/\partial_s F(t,1)$ (recall that F is defined in Equation (7.12)) as N tends to infinity and an application of Dini's Theorem ensures that $\rho_{\lfloor tN \rfloor}$ converges to $\rho(t)$. For all $N \geq 0$, we define the function $\alpha^{(N)}$ by

$$\forall \rho \in (0, \rho_\pi], \quad \alpha^{(N)}(\rho) = 1 - \frac{|\mathscr{S}_{k(\rho)}|}{N} \quad \text{where} \quad k(\rho) = \underset{0\leq k\leq N}{\operatorname{argmin}}\{|1 - \rho - \hat{g}_k(1-\rho)|\}.$$

From this definition, it is clear that $\alpha^{(N)}$ converges to the implicit function α. Moreover, by definition of w in Section 7.3.3, we have that

$$\forall k \in [\![0, K_\varepsilon]\!], \quad w(k) = N\alpha^{(N)}(\rho_k),$$

giving that $\tilde{z}(t) = \alpha(\rho(t))$. On the other hand, using Theorem 16, we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{z}(t) = \frac{1}{\rho(t)}\partial_s \hat{g}(\alpha(\rho(t)), 1),$$

which finally yields

$$y'(\rho) = \frac{\rho\alpha'(\rho)}{\partial_s \hat{g}(\alpha(\rho), 1)}. \tag{7.24}$$

The decreasing part, is obtained by translating horizontally each point $(x^{\uparrow}(\rho), y^{\uparrow}(\rho))$ of the ascending phase to the right by twice the asymptotic proportion of the giant component of the remaining graph of parameter ρ, which is $2(1 - g(\alpha(\rho), 1-\rho))$. Indeed, the time it takes to the DFS to return at a given height k attained during the ascending phase corresponds to the time of exploration of the giant component of the unexplored graph at time T_k. The latter is given by twice the number of vertices of the giant component which is equal to $2(1 - g_k(1-\rho_k))$.

7.3.5 From induced paths to induced cycles

We are going to show that, with high probability, one of the first vertices of the spine of the tree constructed by our algorithm shares a common neighbor with one of the last vertices of the spine and that this common neighbor is not connected to another vertex of the spine. See Figure 7.2 left for an illustration. This ensures that our bound on the length of the longest induced path is also valid for the longest induced cycle.

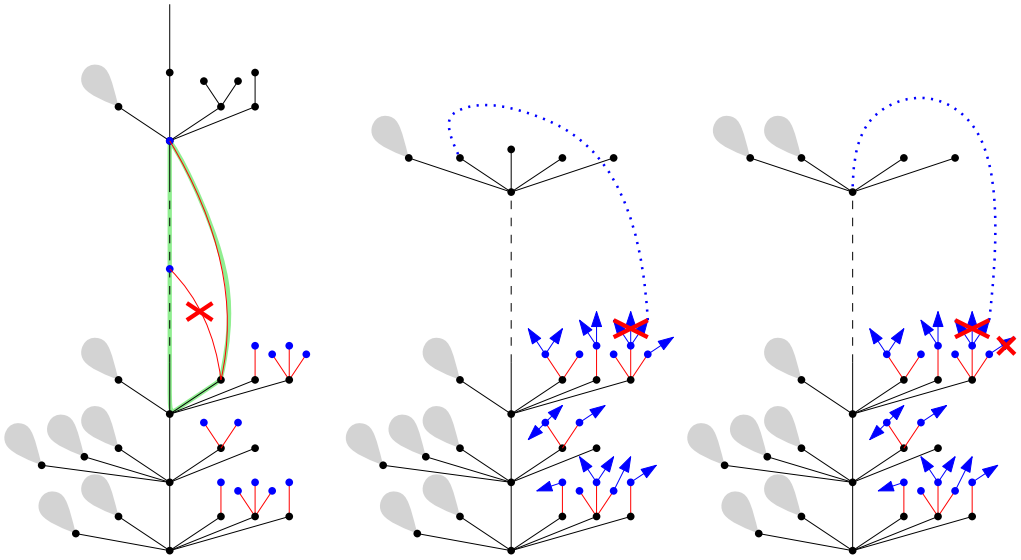


Figure 7.2 – Left: an induced cycle constructed by the algorithm in green. Center: a matching to a half edge of E^ε for a vertex that is not on the spine. Right: a matching to a half edge of E^ε for a vertex of the spine.

Recall the definition of the ladder times T_k, of the associated vertices $\mathbf{v}_k$ on the spine of the tree and of $\mathbf{m}_{\mathbf{v}_k} = \left((u_1,(u_1^1,\dots,u_1^{k_1})),\dots,(u_l,(u_l^1,\dots,u_l^{k_l}))\right)$ given in Equation (7.5). The candidates for a common neighbor between vertices of the spine are $u_{\mathfrak{e}_k+2},\dots,u_l$ (recall that $u_{\mathfrak{e}_k+1} = v_{k+1}$). These vertices are connected to $\mathbf{v}_k$ and to a half edge of each of the vertices u_i^j for $i \in [\![\mathfrak{e}_k+2, l]\!]$. The induced cycles we are interested in are when one of these vertices u_i^j is on the spine of the tree. We denote E_k the set of half edges of the vertices u_i^j for $i \in [\![\mathfrak{e}_k+2, l]\!]$ not yet matched. For $\varepsilon > 0$ small enough, the number of half edges of $E^\varepsilon = \bigcup_{k=1}^{\varepsilon N} E_k$ that are still unmatched at time $T_{\varepsilon N}$ is larger than ηN for some $\eta > 0$. For $\varepsilon N \le k \le K_\varepsilon$, we denote by $\eta_k N$ the number of half edges of E^ε that are still unmatched at time T_k and such that the corresponding vertex has not been connected to a vertex of the spine between heights εN and k. The evolution of η_k can be studied with Wormald's differential equation method in a similar fashion as in the proof of Theorem 16. We have:

$$\mathbb{E}\left[\eta_{k+1}N - \eta_k N | \mathcal{F}_k\right] \ge -\left[\left(\frac{1-\rho_k}{\rho_k} + \frac{1}{\rho_k}(\hat{g}_k'(1)-1)\right)\frac{\hat{g}_0'(1)}{1-\hat{g}_0(0)} + \left(\frac{\hat{g}_0'(1)}{1-\hat{g}_0(0)}\right)^2\right]\frac{\eta_k N}{|\mathscr{S}_k|\, g_k'(1)} + O\left(N^{-\lambda}\right).$$

Indeed, each new matching of the algorithm has a probability $\frac{\eta_k N}{|\mathscr{S}_k|\, g_k'(1)}$ to be with an available half edge of E^ε. The first term $\left(\frac{1-\rho_k}{\rho_k} + \frac{1}{\rho_k}(\hat{g}_k'(1)-1)\right)\frac{\hat{g}_0'(1)}{1-\hat{g}_0(0)}$ corresponds to the number of matchings along the exploration between T_k and T_{k+1}, except for the matching of $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$ (see Figure 7.2, center). For such matchings, we have to remove every other half edge of the discovered vertex from E^ε, whose expected number is roughly the expectation of $\hat{\pi}$ conditioned to be nonzero since it was fixed at the very begining of the construction (note that these half edges are still available for the construction). If $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$ are matched by a half edge of E^ε, this means that v_{k+1} is a neighbor of a vertex u' at distance 1 from the first εN first vertices of the spine. In this case, we have to remove from E^ε every half edge of $\mathbf{v}_{k+1}$ and every half edge of E^ε belonging to a vertex connected to u' (see Figure 7.2, right). The expectation of the number

of these removed edges is bounded from above by the square expectation of $\hat{\pi}$ conditioned to be non 0 (since some some half edges may have been matched previously).

As long as $k \leq K_\varepsilon$, the factor in front of η_k in the above equation is uniformly bounded in absolute value by a finite constant c_ε depending of ε. This ensures that $(\eta_{\lfloor tN \rfloor})_t$ is stochastically dominated by a process that has a fluid limit that is smaller than the solution of the differential equation $y' = -c_\varepsilon y$ as long as $t < K_\varepsilon/N$, thus $\eta_{K_\varepsilon} > \eta \exp(-c_\varepsilon t_{max})$ with high probability.

We established that at the ladder time T_{K_ε}, with high probability, at least $\eta \exp(-c_\varepsilon t_{max})N$ available half edges that belong to some vertex u are such that:

- The vertex u is connected to a vertex u' which is itself connected to the spine at height lower than εN.
- The vertex u' is not connected to any vertex of the spine between heights $\varepsilon N + 1$ and K_ε.

For $k > K_\varepsilon$, each vertex of the spine $\mathbf{v}_k$ has a probability roughly $\frac{\eta \exp(-c_\varepsilon t_{max})}{|\mathscr{S}_k|/N} = \mathcal{O}(1)$ to form with E^ε an induced cycle of length at least $K_\varepsilon - \varepsilon N$. Taking first the limit $N \to \infty$ and then $\varepsilon \to 0$ proves that our algorithm constructs an induced cycle of the same macroscopic length as the spine.

7.4 Comparison with Frieze and Jackson's algorithm

In [FJ87], Frieze and Jackson study the following variant of the Depth-First algorithm that constructs a subtree of each connected component of the graph: Perform a depth-first exploration of the graph, and each time a new vertex is discovered, ask if it is connected to a vertex belonging to its ancestral line in the exploration tree. If it is the case, delete this vertex from the exploration tree and backtrack. The ancestral lines of trees constructed by this algorithm are induced paths by construction, however they are not necessarily spanning trees of the corresponding connected components.

Our algorithm and Frieze and Jackson's algorithm do not, in general, provide the same induced paths on some deterministic examples of graphs. However, for configuration graphs satifying assumptions **(A1), (A2), (A3), (A4)**, with high probability, they will construct two trees with identical spines up to a microscopic number of vertices at the top. To see this, we first construct the graph and our tree with the algorithm of Section 7.2, and then perform Frieze and Jackson's algorithm on the resulting graph and starting at the same initial vertex of the giant component. The construction of the graph gives a total order on its vertices according to their first appearance in the contour of the spanning tree. We use this order to choose between neighbors in the DFS performed by Frieze and Jackson.

Fix $\varepsilon > 0$. The following statement on the Frieze and Jackson's algorithm can be proved by induction. With high probability, for all $k < K_\varepsilon$,

- The exploration goes from $\mathbf{v}_k$ to $\mathbf{v}_{k+1}$,
- The vertices explored between the first visit of $\mathbf{v}_k$ and the first visit of $\mathbf{v}_{k+1}$ form a subset of the vertices explored between the times T_k and T_{k+1} of our algorithm.

Indeed, if Frieze and Jackson's algorithm visits $\mathbf{v}_k$, at the first visit of this vertex, the vertices that have not been explored by this algorithm are the vertices of $\mathscr{S}_k$, a subset of the vertices that our algorithm has explored before time T_k, and the vertices attached to the spine before $\mathbf{v}_k$. The vertices already explored by our algorithm that are still available for Frieze and Jackson's algorithm belong to small connected components inside past graphs $\mathscr{S}_i$ for some $i < k$ and will

never be explored by Frieze and Jackson's algorithm. The vertices attached to the spine form cycles in the graph. They are therefore deleted if explored by Frieze and Jackson's algorithm. Thus, after the first visit of $\mathbf{v}_k$, the vertices that are truly available for Frieze and Jackson's algorithm are the vertices of $\mathscr{S}_k$ and the induction follows.

Finally, from $\mathbf{v}_{K_\varepsilon}$, Frieze and Jackson's algorithm explores a subgraph of $\mathscr{S}_{K_\varepsilon}$ and the length of the induced path cannot be increased by more than the size of its giant component. This giant component has a size $C_\varepsilon N$ with $C_\varepsilon \to 0$ as $\varepsilon \to 0$.

7.5 Extension to m-induced cycles

Let $m \geq 1$. An m-induced cycle inside a given graph G is a cycle such that two vertices separated by k edges of the cycle have a distance at least $\min\{m, k\}$ in the graph G. When $m = 1$, we retrieve the definition of induced cycles. In this last section, we briefly explain how to adapt our arguments in order to find the following lower bound on the length of the longest m-induced cycle in a configuration model:

Proposition 10. *Let $\mathcal{H}_N^m$ be the length of the longest m-induced cycle in in $\mathscr{C}(\mathbf{d}^{(N)})$ satisfying assumptions* **(A1), (A2), (A3), (A4)**. *Then,*

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\frac{\mathcal{H}_N^m}{N} \geq m \int_0^{\rho_\pi} \frac{u\,\alpha'(u)}{\sum_{j=1}^m (\partial_s \hat{g}(\alpha(u), 1))^j} \mathrm{d}u - \varepsilon \right) \underset{N \to +\infty}{\longrightarrow} 1.$$

Proof. This bound comes from the following generalization of the algorithm defined in Section 7.2. The idea is to interpolate between a depth-first and a breadth-first exploration. Each time the exploration goes to a new vertex v, it discovers its m-neighborhood by a breadth-first exploration creating a rooted plane tree $\mathcal{T}_v$ with root v and height at most m. We denote by $v_1, \dots, v_l$ the vertices of height m in $\mathcal{T}_v$ listed in lexicographic order (that is in order of their discovery). Similarly as for in Section 7.2, for each $1 \leq i \leq l$, we successively match the half-edges of v_i to uniform half-edges of the graph and denote by $v_i^1, \dots, v_i^{k_i}$ the corresponding vertices and write $\mathbf{m}_v = \Big((v_1, (v_1^1, \dots, v_1^{k_1})), \dots, (v_l, (v_l^1, \dots, v_l^{k_l}))\Big)$. The evolution of active and sleeping vertices follows the same rules as before except that when a new vertex v is explored, every vertex of the tree $\mathcal{T}_v$ is removed from the set of sleeping vertices.

We can define similar ladder times (T_k) for when the exploration discovers a vertex $\mathbf{v}_k$ of the giant component of the graph induced by the sleeping vertices. We denote by $\mathbf{w}_{k,0} = \mathbf{v}_k; \mathbf{w}_{k,2}; \dots; \mathbf{w}_{k,m} = \mathbf{v}_{k+1}$ the ancestral line between $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$. Between T_k and T_{k+1}, the algorithm explores completely the connected components of the vertices on the left hand side of this ancestral line, and up to a distance between m and 1 for the right hand side (see Figure 7.3 for an illustration). Fix m' between 0 and $m-1$ and denote $u_1, \dots, u_l$ the children of $\mathbf{w}_{k,m'}$ in $\mathcal{T}_{\mathbf{v}_k}$. Let $\mathfrak{e}_{m'}$ be such that $u_{\mathfrak{e}_{m'}+1} = \mathbf{w}_{k,m'+1}$. For $j \leq \mathfrak{e}_{m'}$, we also denote by $W^{(j)}$ the connected components of u_j explored by the algorithm. Finally, for $j > \mathfrak{e}_{m'} + 1$, we denote by $\mathcal{T}_{\mathbf{v}_k}(u_j)$ the subtree of $\mathcal{T}_{\mathbf{v}_k}$ emanating from u_j. The contribution of vertices attached to $\mathbf{w}_{k,m'}$ to the evolution of vertices of degree i is given by:

$$\begin{aligned}
\Delta_i(\mathbf{w}_{k,m'}) = &- V_i\left(\cup_{j=1}^{\mathfrak{e}_{m'}} W^{(j)}\right) - \mathbf{1}_{\{\deg(w_{k,m'+1})=i\}} \\
&- \sum_{x \in \mathcal{T}_{\mathbf{v}_k}(u_{\mathfrak{e}_{m'}+2}) \cup \dots \cup \mathcal{T}_{\mathbf{v}_k}(u_l)} \mathbf{1}_{\deg(x)=i} \\
&+ \sum_{x \in \partial\mathcal{T}_{\mathbf{v}_k}(u_{\mathfrak{e}_k+2}) \cup \dots \cup \partial\mathcal{T}_{\mathbf{v}_k}(u_l)} \left(-\mathbf{1}_{\deg(x)=i} + \mathbf{1}_{\deg(x)=i+1}\right).
\end{aligned}$$

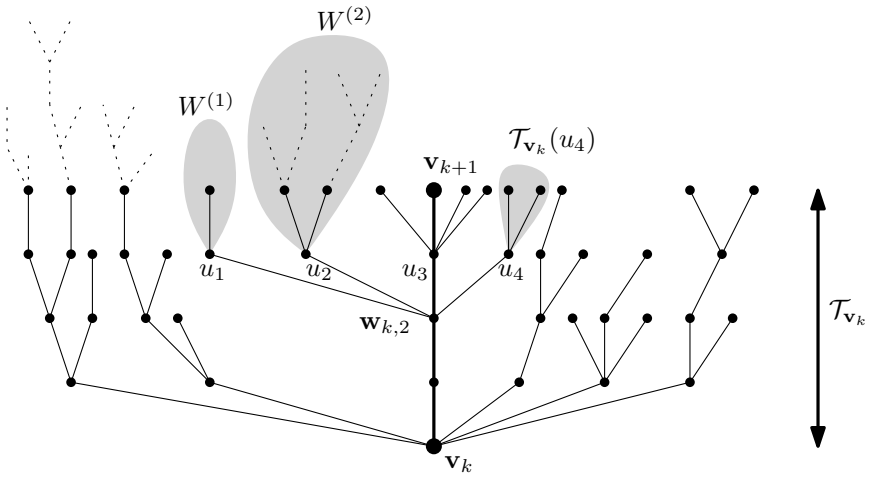


Figure 7.3 – Illustration of the proof of Proposition 10 for $m = 4$ and $m' = 2$. Dotted lines are explored between times T_k and T_{k+1}.

Computations analogous to those made during the proof of Theorem 16 yield

$$\begin{aligned}\mathbb{E}\left[\Delta_i(\mathbf{w}_{k,m'}) \mid \mathcal{F}_k\right] = &-\frac{\hat{p}_{i-1}}{\rho_k}\\ &-\frac{\hat{p}_{i-1}}{\rho_k}\left(\hat{g}_k'(1)-1\right)\left[\sum_{j=0}^{m-m'-1}\left(\hat{g}_k'(1)\right)^j\right]\\ &-\frac{1}{\rho_k}\left(\hat{g}_k'(1)-1\right)\left(\hat{g}_k'(1)\right)^{m-m'}\left(\hat{p}_{i-1}-\hat{p}_i\right)\\ &+\mathcal{O}(N^{-\lambda}).\end{aligned}$$

Simplifying the above expression, we obtain an equation similar to (7.22) :

$$\begin{aligned}\mathbb{E}\left[\Delta_i(\mathbf{w}_{k,m'}) \mid \mathcal{F}_k\right] &= \frac{\hat{p}_{i-1}}{\rho_k}\left(-\left(\hat{g}_k'(1)\right)^{m-m'+1}\right)+\frac{\hat{p}_i}{\rho_k}\left(\hat{g}_k'(1)\right)^{m-m'}\left(-1+\hat{g}_k'(1)\right)+\mathcal{O}(N^{-\lambda})\\ &= \frac{\left(\hat{g}_k'(1)\right)^{m-m'}}{\rho_k}\left(-\hat{p}_{i-1}\hat{g}_k'(1)+\hat{p}_i(-1+\hat{g}_k'(1))\right)+\mathcal{O}(N^{-\lambda}).\end{aligned}$$

Summing for m' between 0 and $m-1$ finally gives:

$$\mathbb{E}\left[N_i(k+1)-N_i(k) \mid \mathcal{F}_k\right] = \frac{\sum_{j=1}^{m}\left(\hat{g}_k'(1)\right)^j}{\rho_k}\left(-\hat{p}_{i-1}\hat{g}_k'(1)+\hat{p}_i(-1+\hat{g}_k'(1))\right)+\mathcal{O}(N^{-\lambda}).$$

Using the same arguments as in the proof of Theorem 16, we deduce that for all $i \geq 0$, the function $t \to N_i(\lfloor tN \rfloor)/N$ converges pointwise towards a function z_i with high probability. The corresponding sequence of functions $(z_i)_{i\geq 0}$ is the unique solution of an infinite system of differential equations, and their generating series $F(t,s) = \sum_{i\geq 0} z_i(t)s^i$ satisfies the following equation:

$$\frac{\partial F}{\partial t}(t,s) = \frac{1}{\rho_{(z_j(t))_{j\geq 0}}}\left(\sum_{m'=1}^{m}\left(\frac{\frac{\partial^2 F}{\partial s^2}(t,1)}{\frac{\partial F}{\partial s}(t,1)}\right)^{m'}\right)\frac{\frac{\partial F}{\partial s}(t,s)}{\frac{\partial F}{\partial s}(t,1)}\left((1-s)\frac{\frac{\partial^2 F}{\partial s^2}(t,1)}{\frac{\partial F}{\partial s}(t,1)}-1\right).$$

Up to a time change, this is exactly Equation 7.12. As it turns out, the resulting new time scale corresponds to the proportion of explored vertices during the exploration of the graph whose derivative is given by the analog of Equation 7.23 which amounts here to the prefactor:

$$\frac{1}{\rho_{(z_j(t))_{j\geq 0}}}\left(\sum_{m'=1}^{m}\left(\frac{\frac{\partial^2 F}{\partial s^2}(t,1)}{\frac{\partial F}{\partial s}(t,1)}\right)^{m'}\right).$$

The contour process of the spanning tree constructed by this new algorithm has therefore a fluid limit with two parametrized arcs $(x^{\uparrow}(\rho), y^{\uparrow}(\rho))$ and $(x^{\downarrow}(\rho), y^{\downarrow}(\rho))$ as in Theorem 15. The derivative of $y^{\uparrow}$ is given by

$$(y^{\uparrow})'(\rho) = m \cdot \frac{\rho\,\alpha'(\rho)}{\sum_{j=1}^{m}(\partial_s \hat{g}(\alpha(\rho),1))^j}$$

which is the analog of Equation (7.24) in the current setting. Notice the factor m which comes from the fact that the ancestral line between two vertices $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$ has length m. □

Bibliography

[AB14] Romain Allez and Jean-Philippe Bouchaud. Eigenvector dynamics under free addition. *Random Matrices Theory Appl.*, 3(3):1450010, 17, 2014.

[ABBG12] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields*, 152(3-4):367–406, 2012.

[AF20] Michael Anastos and Alan Frieze. A scaling limit for the length of the longest cycle in a sparse random digraph. *arXiv preprint arXiv:2001.06481*, 2020.

[Akh65] N. I. Akhiezer. *The classical moment problem and some related questions in analysis*. Translated by N. Kemmer. Hafner Publishing Co., New York, 1965.

[AKS81] Miklós Ajtai, János Komlós, and Endre Szemerédi. The longest path in a random graph. *Combinatorica*, 1(1):1–12, 1981.

[AL07] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[Ald97] David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

[Arn67] Ludwig Arnold. On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.*, 20:262–268, 1967.

[Arn71] L. Arnold. On Wigner's semicircle law for the eigenvalues of random matrices. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 19:191–198, 1971.

[Asm03] Søren Asmussen. *Applied probability and queues*, volume 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[BABP16] Joël Bun, Romain Allez, Jean-Philippe Bouchaud, and Marc Potters. Rotational invariant estimator for general noisy matrices. *IEEE Trans. Inform. Theory*, 62(12):7475–7490, 2016.

[BAG08] Gérard Ben Arous and Alice Guionnet. The spectrum of heavy tailed random matrices. *Comm. Math. Phys.*, 278(3):715–751, 2008.

[Bat54] Harry Bateman. *Tables of integral transforms [volumes I & II]*, volume 1. McGraw-Hill Book Company, 1954.

[BBAP05] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

[BBCF17] Serban T. Belinschi, Hari Bercovici, Mireille Capitaine, and Maxime Février. Outliers in the spectrum of large deformed unitarily invariant models. *Ann. Probab.*, 45(6A):3571–3625, 2017.

[BBR+12] Lars Backstrom, Paolo Boldi, Marco Rosa, Johan Ugander, and Sebastiano Vigna. Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 33–42, 2012.

[BC78] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[BDG09] Serban Belinschi, Amir Dembo, and Alice Guionnet. Spectral measure of heavy tailed band and covariance random matrices. *Comm. Math. Phys.*, 289(3):1023–1055, 2009.

[BEK+14] Alex Bloemendal, László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.*, 19:no. 33, 53, 2014.

[BFF84] B. Bollobás, T. I. Fenner, and A. M. Frieze. Long cycles in sparse random graphs. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 59–64. Academic Press, London, 1984.

[BGCD12] Florent Benaych-Georges and Thierry Cabanal-Duvillard. Marčenko-Pastur theorem and Bercovici-Pata bijections for heavy-tailed or localized vectors. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):685–715, 2012.

[BGEM18] Florent Benaych-Georges, Nathanaël Enriquez, and Alkéos Michaïl. Eigenvectors of a matrix under random perturbation. *arXiv preprint arXiv:1801.10512*, 2018.

[BGK17] Florent Benaych-Georges and Antti Knowles. Local semicircle law for Wigner matrices. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 1–90. Soc. Math. France, Paris, 2017.

[BGN11] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.*, 227(1):494–521, 2011.

[BH12] Andries E. Brouwer and Willem H. Haemers. *Spectra of graphs*. Universitext. Springer, New York, 2012.

[Bia97] Philippe Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.*, 46(3):705–718, 1997.

[Bia03] Philippe Biane. Free probability for probabilists. In *Quantum probability communications, Vol. XI (Grenoble, 1998)*, QP-PQ, XI, pages 55–71. World Sci. Publ., River Edge, NJ, 2003.

[Big93] Norman Biggs. *Algebraic graph theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 1993.

[BKYY16] Alex Bloemendal, Antti Knowles, Horng-Tzer Yau, and Jun Yin. On the principal components of sample covariance matrices. *Probab. Theory Related Fields*, 164(1-2):459–552, 2016.

[BL10] Charles Bordenave and Marc Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37(3):332–352, 2010.

[BLS11] Charles Bordenave, Marc Lelarge, and Justin Salez. The rank of diluted random graphs. *Ann. Probab.*, 39(3):1097–1121, 2011.

[BMP07] Z. D. Bai, B. Q. Miao, and G. M. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.*, 35(4):1532–1572, 2007.

[BN17] Mogens Bladt and Bo Friis Nielsen. *Matrix-exponential distributions in applied probability*, volume 81. Springer, 2017.

[Bol80] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[Bol82] Béla Bollobás. Long paths in sparse random graphs. *Combinatorica*, 2(3):223–228, 1982.

[Bol83] Béla Bollobás. Almost all regular graphs are Hamiltonian. *European J. Combin.*, 4(2):97–106, 1983.

[Bol84] Béla Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, 286(1):257–274, 1984.

[Bol01] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[BR15] Béla Bollobás and Oliver Riordan. An old approach to the giant component problem. *J. Combin. Theory Ser. B*, 113:236–260, 2015.

[Bre92] Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.

[BS98] Z. D. Bai and Jack W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1):316–345, 1998.

[BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001.

[BSV17] Charles Bordenave, Arnab Sen, and Bálint Virág. Mean quantum percolation. *J. Eur. Math. Soc. (JEMS)*, 19(12):3679–3707, 2017.

[BSY88] Z. D. Bai, Jack W. Silverstein, and Y. Q. Yin. A note on the largest eigenvalue of a large-dimensional sample covariance matrix. *J. Multivariate Anal.*, 26(2):166–168, 1988.

[BY88] Z. D. Bai and Y. Q. Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.*, 16(4):1729–1741, 1988.

[BY12] Zhidong Bai and Jianfeng Yao. On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.*, 106:167–177, 2012.

[Cap13] M. Capitaine. Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices. *J. Theoret. Probab.*, 26(3):595–648, 2013.

[CB94] Pierre Cizeau and Jean-Philippe Bouchaud. Theory of lévy matrices. *Physical Review E*, 50(3):1810, 1994.

[CDM17] Mireille Capitaine and Catherine Donati-Martin. Spectrum of deformed random matrices and free probability. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 151–190. Soc. Math. France, Paris, 2017.

[CDMF09] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Ann. Probab.*, 37(1):1–47, 2009.

[CDMFF11] M. Capitaine, C. Donati-Martin, D. Féral, and M. Février. Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices. *Electron. J. Probab.*, 16:no. 64, 1750–1792, 2011.

[CDS95] Dragoš M. Cvetković, Michael Doob, and Horst Sachs. *Spectra of graphs*. Johann Ambrosius Barth, Heidelberg, third edition, 1995. Theory and applications.

[Cha06] Sourav Chatterjee. A generalization of the Lindeberg principle. *Ann. Probab.*, 34(6):2061–2076, 2006.

[CS18] Simon Coste and Justin Salez. Emergence of extended states at zero in the spectrum of sparse random graphs. *arXiv preprint arXiv:1809.07587*, 2018.

[DE02] Ioana Dumitriu and Alan Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[Dei99] P. A. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, volume 3 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

[Don97] R. A. Doney. One-sided local large deviation and renewal theorems in the case of infinite mean. *Probab. Theory Related Fields*, 107(4):451–465, 1997.

[Dur10] Rick Durrett. *Random graph dynamics*, volume 20 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010.

[DvdHvLS17] Souvik Dhara, Remco van der Hofstad, Johan S. H. van Leeuwaarden, and Sanchayan Sen. Critical window for the configuration model: finite third moment degrees. *Electron. J. Probab.*, 22:Paper No. 16, 33, 2017.

[EFM20] Nathanaël Enriquez, Gabriel Faraud, and Laurent Ménard. Limiting shape of the depth first search tree in an Erdős-Rényi graph. *Random Structures & Algorithms*, 56(2):501–516, 2020.

[EFMN19] Nathanaël Enriquez, Gabriel Faraud, Laurent Ménard, and Nathan Noiry. Depth first exploration of a configuration model. *arXiv preprint arXiv:1911.10083*, 2019.

[EFMN20] Nathanaël Enriquez, Gabriel Faraud, Laurent Ménard, and Nathan Noiry. Long induced paths in a configuration model. *In preparation*, 2020.

[EM16] Nathanaël Enriquez and Laurent Ménard. Spectra of large diluted but bushy random graphs. *Random Structures Algorithms*, 49(1):160–184, 2016.

[EN20] Nathanaël Enriquez and Nathan Noiry. A solvable class of renewal processes. *To appear in Electronic Communications in Probability*, 2020.

[ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[Erd75] Paul Erdős. Problems and results on finite and infinite graphs. In *Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*, pages 183–192. (loose errata), 1975.

[ES07] Alan Edelman and Brian D. Sutton. From random matrices to stochastic operators. *J. Stat. Phys.*, 127(6):1121–1165, 2007.

[ESY09] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009.

[EY17] László Erdős and Horng-Tzer Yau. *A dynamical approach to random matrix theory*, volume 28 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017.

[FdlV79] W. Fernandez de la Vega. Long paths in random graphs. *Studia Sci. Math. Hungar.*, 14(4):335–340, 1979.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[Fel71] William Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[FF77] M. Fixman and J. J. Freire. Theory of DNA melting curves. *Biopolymers*, 16:2693–2704, 1977.

[FJ87] A. M. Frieze and B. Jackson. Large holes in sparse random graphs. *Combinatorica*, 7(3):265–274, 1987.

[FK81] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[FK16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

[Fri86] A. M. Frieze. On large matchings and cycles in sparse random graphs. *Discrete Math.*, 59(3):243–256, 1986.

[GCK20] Christina Goldschmidt and Guillaume Conchon-Kerjan. The stable graph: the metric space scaling limit of a critical random graph with iid power-law degrees. *arXiv preprint arXiv:2002.04954*, 2020.

[Gem80] Stuart Geman. A limit theorem for the norm of random matrices. *Ann. Probab.*, 8(2):252–261, 1980.

[Gia07] Giambattista Giacomin. *Random polymer models*. Imperial College Press, London, 2007.

[Gia08] Giambattista Giacomin. Renewal convergence rates and correlation decay for homogeneous pinning models. *Electronic Journal of Probability*, 13:513–529, 2008.

[Gil59] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

[GNR16] Fabrice Gamboa, Jan Nagel, and Alain Rouault. Sum rules via large deviations. *J. Funct. Anal.*, 270(2):509–559, 2016.

[GNR17] Fabrice Gamboa, Jan Nagel, and Alain Rouault. Sum rules and large deviations for spectral measures on the unit circle. *Random Matrices Theory Appl.*, 6(1):1750005, 49, 2017.

[GNR19] Fabrice Gamboa, Jan Nagel, and Alain Rouault. Sum rules and large deviations for spectral matrix measures. *Bernoulli*, 25(1):712–741, 2019.

[GR11] Fabrice Gamboa and Alain Rouault. Large deviations for random spectral measures and sum rules. *Appl. Math. Res. Express. AMRX*, (2):281–307, 2011.

[Gre63] Ulf Grenander. *Probabilities on algebraic structures*. John Wiley & Sons, Inc., New York-London; Almqvist & Wiksell, Stockholm-Göteborg-Uppsala, 1963.

[HM12] Hamed Hatami and Michael Molloy. The scaling window for a random graph with a given degree sequence. *Random Structures Algorithms*, 41(1):99–123, 2012.

[JL09] Svante Janson and Malwina J. Luczak. A new approach to the giant component problem. *Random Structures Algorithms*, 34(2):197–216, 2009.

[Joh01a] Kurt Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.*, 215(3):683–705, 2001.

[Joh01b] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2):295–327, 2001.

[Jon82] Dag Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982.

[Jos14] Adrien Joseph. The component sizes of a critical random graph with given degree sequence. *Ann. Appl. Probab.*, 24(6):2560–2594, 2014.

[KS81] Richard M Karp and Michael Sipser. Maximum matching in sparse random graphs. In *22nd Annual Symposium on Foundations of Computer Science (sfcs 1981)*, pages 364–375. IEEE, 1981.

[KS13] Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: a simple proof. *Random Structures Algorithms*, 43(2):131–138, 2013.

[KY14] Antti Knowles and Jun Yin. The outliers of a deformed Wigner matrix. *Ann. Probab.*, 42(5):1980–2031, 2014.

[KY17] Antti Knowles and Jun Yin. Anisotropic local laws for random matrices. *Probab. Theory Related Fields*, 169(1-2):257–352, 2017.

[Len15] Romuald Lenczewski. Random matrix model for free Meixner laws. *Int. Math. Res. Not. IMRN*, (11):3499–3524, 2015.

[LP11] Olivier Ledoit and Sandrine Péché. Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Related Fields*, 151(1-2):233–264, 2011.

[LSSY16] Ji Oon Lee, Kevin Schnelli, Ben Stetler, and Horng-Tzer Yau. Bulk universality for deformed Wigner matrices. *Ann. Probab.*, 44(3):2349–2425, 2016.

[Łuc90] Tomasz Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms*, 1(3):287–310, 1990.

[Mal17] Camille Male. The limiting distributions of large heavy Wigner and arbitrary random matrices. *J. Funct. Anal.*, 272(1):1–46, 2017.

[Meh60] M. L. Mehta. On the statistical properties of the level-spacings in nuclear spectra. *Nuclear Phys.*, 18:395–419, 1960.

[Meh04] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[MG60] M. L. Mehta and M. Gaudin. On the density of eigenvalues of a random matrix. *Nuclear Phys.*, 18:420–427, 1960.

[MP67] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

[MR95] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 161–179, 1995.

[MR98] Michael Molloy and Bruce Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7(3):295–305, 1998.

[MS06] Enzo Marinari and Guilhem Semerjian. On the number of circuits in random graphs. *J. Stat. Mech. Theory Exp.*, (6):P06019, 41, 2006.

[Nag15] S. V. Nagaev. Local renewal theorems in the absence of an expectation. *Theory Probab. Appl.*, 59(3):388–414, 2015.

[Noi18] Nathan Noiry. Spectral asymptotic expansion of Wishart matrices with exploding moments. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(2):897–911, 2018.

[Noi20] Nathan Noiry. Spectral measures of spiked random matrices. *Journal of Theoretical Probability*, 2020.

[NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[NS10] Raj Rao Nadakuditi and Jack W Silverstein. Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples. *IEEE Journal of Selected Topics in Signal Processing*, 4(3):468–480, 2010.

[Pé06] Sandrine Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields*, 134(1):127–173, 2006.

[Pas72] L. A. Pastur. The spectrum of random matrices. *Teoret. Mat. Fiz.*, 10(1):102–112, 1972.

[Pau07] Debashis Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica*, 17(4):1617–1642, 2007.

[PS11] Leonid Pastur and Mariya Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[PSZ10] Alexei Poltoratski, Barry Simon, and Maxim Zinchenko. The Hilbert transform of a measure. *Journal d'Analyse Mathématique*, 111(1):247–265, 2010.

[Rio12] Oliver Riordan. The phase transition in the configuration model. *Combin. Probab. Comput.*, 21(1-2):265–299, 2012.

[Rog73] B. A. Rogozin. An estimate of the remainder term in limit theorems of renewal theory. *Teor. Verojatnost. i Primenen.*, 18:703–717, 1973.

[RRV11] José A. Ramírez, Brian Rider, and Bálint Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.*, 24(4):919–944, 2011.

[Rya98] Øyvind Ryan. On the limit distributions of random matrices with independent or free entries. *Comm. Math. Phys.*, 193(3):595–626, 1998.

[Sal11] Justin Salez. *Some implications of local weak convergence for sparse random graphs*. PhD thesis, 2011.

[Sal15] Justin Salez. Every totally real algebraic integer is a tree eigenvalue. *J. Combin. Theory Ser. B*, 111:249–256, 2015.

[Sal20] Justin Salez. Spectral atoms of unimodular random trees. *J. Eur. Math. Soc. (JEMS)*, 22(2):345–363, 2020.

[SC95] Jack W. Silverstein and Sang-Il Choi. Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):295–309, 1995.

[Sil95] Jack W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.*, 55(2):331–339, 1995.

[Sim05] Barry Simon. *Orthogonal polynomials on the unit circle*. American Mathematical Soc., 2005.

[Sos99] Alexander Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.*, 207(3):697–733, 1999.

[Sos02] Alexander Soshnikov. A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *Journal of Statistical Physics*, 108(5-6):1033–1056, 2002.

[Tao12] Terence Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[TV10] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010.

[TV12] Terence Tao and Van Vu. Random covariance matrices: universality of local statistics of eigenvalues. *Ann. Probab.*, 40(3):1285–1315, 2012.

[TW94] Craig A. Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994.

[vdEvdHH08] Henri van den Esker, Remco van der Hofstad, and Gerard Hooghiemstra. Universality for the distance in finite variance random graphs. *Journal of Statistical Physics*, 133(1):169–202, 2008.

[vdH17] Remco van der Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017.

[vdHHVM05] Remco van der Hofstad, Gerard Hooghiemstra, and Piet Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.

[Ven14] V. Vengerovsky. Eigenvalue distribution of a large weighted bipartite random graph. *Zh. Mat. Fiz. Anal. Geom.*, 10(2):240–255, 259, 261, 2014.

[VV09] Benedek Valkó and Bálint Virág. Continuum limits of random matrices and the Brownian carousel. *Invent. Math.*, 177(3):463–508, 2009.

[Wig55] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, 62:548–564, 1955.

[Wig58] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

[Wis28] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.

[Wor80] Nicholas C Wormald. Some problems in the enumeration of labelled graphs. *Bulletin of the Australian Mathematical Society*, 21(1):159–160, 1980.

[Wor95] Nicholas C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.*, 5(4):1217–1235, 1995.

[YBK88] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah. On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields*, 78(4):509–521, 1988.

[Zak06] Inna Zakharevich. A generalization of Wigner's law. *Comm. Math. Phys.*, 268(2):403–414, 2006.