universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

Mixing and cover time profiles for random walks on graphs

verfasst von / submitted by

Lucas Teyssier, licence master

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 796 605 405
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Mathematik
Betreut von / Supervisor:	Univ.-Prof. Dr. Nathanaël Berestycki

Acknowledgements

First, I am immensely grateful to my supervisor, Nathanaël Berestycki for your multidimensional guidance and the insightful discussions we had in the last years. I learnt a lot at your contact, thank you!

Thanks also to my former teachers Christophe Kapoudjian, Denis Choimet, Thomas Budzinski, and Justin Salez, for making me like abstract algebra, Fourier analysis, probability, and mixing times respectively; to my collaborators Amaury Freslon, Jonathan Hermon, and Simeng Wang for the interesting ideas you shared with me and your support in my growth as a young researcher; and to Antoine Jego, Pierre-Loïc Méliot, Evita Nestoridi, Sam Olesker–Taylor, and Dominik Schmid for the enriching conversations we had.

I am furthermore very thankful to Persi Diaconis and Perla Sousi for accepting to read and examine this thesis, and for your interest in my work.

Next, to all present and former members of the 6-th floor (and some of the 9-th floor as well, and the goat!), for the excellent atmosphere we had over all those years, both at work and in the evenings!

Finally, a big thank to my family and Thurian for your presence and support throughout these years.

List of included articles

Chapters 2, 3, and 4 are respectively based on the following articles.

- [BHT22]: Nathanaël Berestycki, Jonathan Hermon, and Lucas Teyssier, *On the universality of fluctuations for the cover time.* ArXiv e-prints, March 2023
- [Tey20]: Lucas Teyssier, *Limit profile for random transpositions*, Ann. Probab., 48(5):2323–2343, 2020.
- [FTW22]: Amaury Freslon, Lucas Teyssier, and Simeng Wang, *Cutoff profiles for quantum Lévy processes and quantum random transpositions*, Probab. Theory Related Fields, 183(3-4):1285–1327, 2022.

Abstract

In this thesis, we contribute to the quantitative theory of random walks on graphs.

We first consider the cover time problem. On finite vertex-transitive graphs of bounded degree, we find a necessary and sufficient condition for the properly rescaled cover time to have asymptotically Gumbel fluctuations. More precisely, this holds if and only if $\mathrm{Diam}(\Gamma)^2 = o(n/\log n)$, where $n = |\Gamma|$. Surprisingly, the condition is therefore purely in terms of the global geometry of the graph. Furthermore, we prove that this condition is equivalent to the decorrelation of the uncovered set.

The arguments rely on recent breakthroughs by Tessera and Tointon on finitary versions of Gromov's theorem on groups of polynomial growth, which we leverage into strong heat kernel bounds. At the technical level the main innovation consists in obtaining refined quantitative estimates for an exponential approximation of hitting times, first put forward by Aldous and Brown.

In the next part we study the problem of mixing times, mostly with representation theoretic techniques. First, we find an improvement of the Diaconis–Shahshahani upper bound lemma, which is suitable for investigating limit profiles. Combining this with a novel identity involving the characters of the symmetric group, this allows us to find the cutoff profile for random transpositions. This admits an explicit expression in terms of Poisson laws.

Finally, we study an analogue of Brownian motion on free orthogonal quantum groups. We find its cutoff profile, which involves free Poisson distributions and the semi-circle law.

Zusammenfassung

In dieser Doktorarbeit, tragen wir zur quantitativen Theorie der Zufälligen Irrfahrten auf Graphen bei.

Wir betrachten zunächst das Problem der Überdeckungszeit. Auf endlichen knoten-transitiven Graphen mit beschränktem Grad, finden wir eine notwendige und hinreichende Bedingung dafür, dass die korrekt skalierte Überdeckungszeit asymptotisch Gumbel-Schwankungen hat. Genauer gesagt, gilt dies dann und nur dann, wenn $\mathrm{Diam}(\Gamma)^2 = o(n/\log n)$, wo $n = |\Gamma|$. Überraschenderweise bezieht sich die Bedingung also ausschließlich auf die globale Geometrie des Graphen. Außerdem beweisen wir, dass diese Bedingung äquivalent zur Dekorrelation der unüberdeckten Menge ist.

Die Argumente stützen sich auf die jüngsten Durchbrüche von Tessera und Tointon zu finitären Versionen vom Satz von Gromov über Gruppen mit polynomialem Wachstum, die wir in starke Schranken für den Wärmeleitungskern umwandeln. Auf technischer Ebene besteht die wichtigste Neuerung darin, verfeinerte quantitative Schätzungen für eine exponentielle Annäherung der Trefferzeiten zu erhalten, die erstmals von Aldous und Brown vorgeschlagen wurden.

Im nächsten Teil untersuchen wir das Problem der Mischzeiten , hauptsächlich mit darstellungstheoretischen Techniken. Zunächst finden wir eine Verbesserung des Lemmas der oberen Schranke von Diaconis und Shahshahani, die für die Untersuchung von Grenzprofilen geeignet ist. Kombiniert man dies mit einer neuen Identität, die die Zeichen der symmetrischen Gruppe einbezieht, kann man das Schnittprofil für zufällige Transpositionen finden. Dieses lässt einen expliziten Ausdruck in Form der Poisson-Verteilungen zu.

Schließlich untersuchen wir ein Analogon der Brownschen Bewegung auf freien orthogonalen Quantengruppen. Wir finden dessen Schnittprofil, das freie Poisson-Verteilungen und die Halbkreisverteilung involviert.

Résumé

Dans cette thèse, nous contribuons à la théorie quantitative des marches aléatoires sur les graphes.

Nous considérons tout d'abord le problème du temps de couverture. Sur les graphes sommet-transitifs finis de degré borné, nous trouvons une condition nécessaire et suffisante pour que les fluctuations du temps de couverture correctement renormalisé suivent asymptotiquement une loi de Gumbel. Plus précisément, les fluctuations sont Gumbel si et seulement si $\mathrm{Diam}(\Gamma)^2 = o(n/\log n)$, où $n = |\Gamma|$. De manière surprenante, la condition dépend donc uniquement de la géométrie globale du graphe. Nous prouvons par ailleurs que cette condition est équivalente à la décorrélation de l'ensemble des points non couverts.

Les arguments reposent sur des percées récentes de Tessera et Tointon, prouvant des versions finitaires du théorème de Gromov sur les groupes à croissance polynomiale, desquelles nous déduisons des bornes précises sur le noyau de la chaleur associé à la marche. Au niveau technique, la principale innovation est une amélioration des estimées d'Aldous et Brown sur l'approximation exponentielle des temps d'atteinte.

Dans la partie suivante, nous étudions le problème des temps de mélange, principalement à l'aide de la théorie des représentations. Nous trouvons une amélioration du lemme de majoration de Diaconis–Shahshahani qui permet d'obtenir des profils limites. Combinée avec une nouvelle identité sur les caractères du groupe symétrique, cela nous permet de trouver le profil de coupure pour les transpositions aléatoires, qui peut s'écrire avec des lois de Poisson.

Enfin, nous étudions un analogue du mouvement brownien sur les groupes quantiques orthogonaux libres. Nous trouvons son profil de coupure, qui implique des distributions de Poisson libres et la loi du demi-cercle.

Resumo

En tiu disertaĵo, ni kontribuas al kvanteca teorio de hazardaj promenadoj sur grafeoj.

Ni unue konsideras kovrotempojn. Sur finiaj vertictranzitivaj baritgradaj grafeoj, ni trovas necesan kaj sufiĉan kondiĉon por ke la konvene reskalita kovrotempo asimptote havu fluktuaĵojn de Gumbel. Pli precize, la fluktuaĵoj Gumbel-as se kaj nur se $\mathrm{Diam}(\Gamma)^2 = o(n/\log n)$, kie $n = |\Gamma|$. Surprize, tiu kondiĉo do nur dependas de la makroskopa geometrio de la grafeoj. Ni krome pruvas ke ĝi ekvivalentas al la nekorelacio de la nekovrita aro.

La argumentoj baziĝas sur freŝdataj sukcesegoj de Tessera kaj Tointon pri finitaraj versioj de la teoremo de Gromov pri grupoj kun polinoma kresko, kiujn ni konvertas al baroj sur la varmeca kerno de la promenado. Tekniknivele, la ĉefa novaĵo konsistas en plibonigo de la kvantecaj taksoj de Aldous kaj Brown pri eksponenciala aproksimado de atingotempoj.

Sekvaparte ni studas la problemon de miksotempoj, ĉefe per reprezentteoriaj metodoj. Ni unue trovas plibonigon de la superbara lemo de Diaconis–Shahshahani, kiu konvenas por kalkuli limesajn profilojn. Kombinante tiun kun nova identeco pri la karakteroj de la simetria grupo, ni trovas la tranĉprofilon por hazardaj du-cikloj. Ĝi eksplicite esprimeblas per distribuoj de Poisson.

Ni fine studas analogon de braŭna movado sur liberaj ortaj kvantumaj grupoj. Ni trovas ĝian tranĉprofilon, kiu enhavas liberajn distribuojn de Poisson kaj la duoncirklan leĝon.

Contents

Chapter 1

Introduction

How many times does one need to shuffle a deck of 52 cards to mix it *well*? What if the number of cards tends to infinity? More generally, how long does it take for a random walk on a graph to be *close to stationarity*? Precise answers to those questions started to be given in the 80's in the founding works of Aldous and Diaconis of what is now called the theory of mixing times.

Another question we consider is the following. How many steps does a random walk on a graph need to *cover* the graph, that is to visit every vertex at least once? This is the problem of cover times, which has received more attention since the works of Aldous in the 80's.

In this introduction, we will first introduce cover times, which are simpler to define, and then mixing times, with an emphasis on representation theoretic techniques.

1.1 Cover times

1.1.1 Definition and first examples

Let us consider the simple random walk $X = (X_t)_{t \geq 0}$ on a graph Γ. How many steps does X need to visit all vertices? This random time is called the *cover time* of Γ and is denoted by

$$\tau_{\text{cov}} = \max_{x \in \Gamma} T_x,$$

where $T_x = \inf \{t \geq 0 : X_t = x\}$ is the hitting time of the vertex x. Of course, for τ_{cov} to be finite, the graph Γ has to be finite and connected. We will always make this assumption.

Let us give a few examples.

Example 1.1.1 (The complete graph)**.** Consider the complete graph K_n. At each step, the walk moves to an (almost) uniform vertex, so the cover time problem on K_n is equivalent to the well-studied coupon-collector problem. It follows from [ER61] that we have for each fixed $s \in \mathbb{R}$

$$\mathbb{P}(\tau_{\text{cov}} \leq n(\log n + s)) \xrightarrow[n \to \infty]{} e^{-e^{-s}}. \tag{1.1.1}$$

Example 1.1.2 (The cycle)**.** Consider the cycle $\mathbb{Z}/n\mathbb{Z}$. In this case the cover time is of order n^2 and is not concentrated.

Example 1.1.3 (The complete graph with an extra edge)**.** Consider the complete graph $K_n = (V_n, E_n)$ and let $o \in K_n$ be any vertex. By attaching a vertex z to K_n at o, we get the graph $\Gamma_n := (V_n \cup \{z\}, E \cup \{\{o, z\}\})$. If the random walk on Γ_n starts at z then the asymptotic behaviour is as for K_n. However, if the walk starts at any other vertex, then the walk will visit all points in the complete graph after about $n \log n$ steps, but it will need order n^2 steps to visit z: while it is in the complete graph it spends $\frac{1+o(1)}{n}$-th of the time at o, and each time it is at o it has probability $\frac{1+o(1)}{n}$ to jump to z. More precisely,

$$\frac{\tau_{\text{cov}}}{n^2} \xrightarrow[n\to\infty]{(d)} E, \tag{1.1.2}$$

where E is an exponential random variable of parameter 1. This illustrates that cover times are very sensitive to small perturbations of the graphs, and to the starting vertex of the walk.

1.1.2 Some history and related problems

Cover times are one of the most natural quantities associated to random walks to consider. Aldous found in [Ald83b], for random walks on finite groups and under mild geometric assumptions, the value of $\mathbb{E}\,\tau_{\text{cov}}$ and proved that the cover time is asymptotically concentrated around its mean:

$$\frac{\tau_{\text{cov}}}{\mathbb{E}\,\tau_{\text{cov}}} \xrightarrow[|\Gamma|\to\infty]{} 1,$$

where we use the notation Γ also for the vertex set of Γ, with an abuse of notation which we will also make below. A few years later Matthews [Mat88a] gave a sharp upper bound on the expected cover time valid for all graphs:

$$\mathbb{E}\,\tau_{\text{cov}} \leq \mathrm{t}_{\text{hit}}\, H(n-1), \tag{1.1.3}$$

where $\mathrm{t}_{\text{hit}} := \max_{x,y\in\Gamma} \mathbb{E}_x\,[T_y]$ is the maximal hitting time, and $H(n-1) = 1 + \frac{1}{2} + \ldots + \frac{1}{n-1}$ is the $(n-1)$-st harmonic number.

Studying the cover time can be very hard, even on elementary graphs. It was only in 2004 that the first order of the cover time for the two-dimensional torus $(\mathbb{Z}/n\mathbb{Z})^2$ was found by Dembo, Peres, Rosen and Zeitouni in [DPRZ04].

$$\frac{\tau_{\text{cov}}}{(n\log n)^2} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{4}{\pi}. \tag{1.1.4}$$

This result was later refined, see [BK17, Abe21, BRZ20], but the precise profile is still to be discovered.

More is known in higher dimensions: In 2013, Belius found in [Bel13], building on Sznitman's random interlacement model [Szn10], the fluctuations of the cover time variable around its mean for tori of dimension $d \geq 3$: they follow a standard Gumbel law, as for the complete graph.

A related and more general question is to understand the structure of the uncovered set

$$U(t) = \Gamma \setminus \{X_u : u \leq t\} \tag{1.1.5}$$

for t slightly to long before the cover time. This problem has been studied at a fraction of the cover time on tori of dimension ≥ 3 in [MS17], and refined results were proved in [OTS20], with an elementary Chen–Stein approach. They prove that the uncovered set at time $a\,\mathrm{t}_{\text{cov}}$ is decorrelated, that is close to a product measure, if $a > 7/8$ (and $a > 3/4$ as $d \to \infty$) and correlated if $a < \frac{1+p_d}{2}$, where $p_d = \mathbb{P}_o(T_o^+ < \infty)$ for the simple random walk on $\mathbb{Z}^d$. It nevertheless still remains a fascinating open question whether there is a phase transition, i.e. whether there exists a critical value a^* such that the uncovered set is close to (resp. far from) a product measure at time $a\,\mathrm{t}_{\text{cov}}$ if $a > a^*$ (resp. $a < a^*$). There was recent progress on this question in [PRS23] where a similar type of phase transition was proved at $a^* = \frac{1+p_d}{2}$.

Cover times are also related to the maximum of the *Gaussian free field*. Ding, Lee, and Peres proved in[DLP12] that for any connected graph $\Gamma = (V, E)$, if $(\eta_v)_{v\in V}$ is the Gaussian free field,

$$\max_{v\in V} \mathbb{E}_v\, \tau_{\text{cov}} \asymp |E| \left(\mathbb{E} \max_{v\in V} \eta_v\right)^2 . \tag{1.1.6}$$

This was later generalised in [Din14, Zha18].

Finally, cover times of graphs are linked the mixing times of the associated lamplighter graphs [PR04, DDMP19], where intriguing phase transitions can occur.

1.1.3 Universality of Gumbel fluctuations: main results of Chapter 2

The purpose of Chapter 2 is to prove that Gumbel fluctuations are universal in a large class of graphs. Our main result characterises the cover time fluctuations with a geometric condition on graphs. Write $|\Gamma|$ for the number of vertices of Γ, and let $D = \max_{x,y\in\Gamma} d(x, y)$ be its diameter. Recall that $\mathrm{t_{hit}} := \max_{x,y\in\Gamma} \mathbb{E}_x [T_y]$ is the maximal expected hitting time of the chain.

Theorem 1.1.1 (Theorem 2.1.1)**.** *Let* (Γ) *be a collection of finite (connected) vertex-transitive graphs of fixed degree* d*, and let* χ *be a standard Gumbel variable, i.e.* $\mathbb{P}(\chi \le s) = e^{-e^{-s}}$ *for* $s \in \mathbb{R}$*. Then*

$$\frac{\tau_{cov}}{t_{hit}} - \log|\Gamma| \xrightarrow[|\Gamma|\to\infty]{(d)} \chi$$

if and only if

$$\frac{D^2 \log|\Gamma|}{|\Gamma|} \xrightarrow[|\Gamma|\to\infty]{} 0. \tag{DC}$$

We could also complement it by a characterisation of the structure of the uncovered set. For $s \in \mathbb{R}$, set $t_s := \mathrm{t_{hit}}(\log(|\Gamma|) + s)$ and let $\mathcal{L}(U(t_s))$ denote the law of the uncovered set at time t_s.

Theorem 1.1.2 (Theorem 2.1.2)**.** *Let* (Γ) *be a collection of finite (connected) vertex-transitive graphs of fixed degree* d*, and* $s \in \mathbb{R}$*. Let* $\mu_s^{\otimes\Gamma}$ *denote the product over all the vertices of the graph of the Bernoulli law* μ_s *with parameter* $e^{-s}/|\Gamma|$*. Then*

$$d_{TV}(\mathcal{L}(U(t_s)), \mu_s^{\otimes\Gamma}) \xrightarrow[|\Gamma|\to\infty]{} 0,$$

if and only if (DC) *holds.*

In words the uncovered set at t_s is *decorrelated* if and only if (DC) holds, i.e. if and only if the graphs Γ are at least slightly more than *2-dimensional*. The local structure of the graph affects $\mathrm{t_{hit}}$ and hence the renormalisation, but it is the macroscopic condition (DC) which characterises whether the cover time is close to a product measure and the cover time fluctuations are Gumbel.

Let us now explain the main ingredients of our proof. If $|\Gamma| = D^q$, we will informally call q the *dimension* of Γ. The hardest cases to prove decorrelation of the uncovered set are low dimensional graphs, especially as the dimension approaches 2. To make things more concrete, let (Γ) be a collection of vertex-transitive graphs of degree 8 and dimension 2.3, that is $n := |\Gamma| = D^{2.3}$. We consider the simple random walk (in continuous time) on Γ started at a (fixed) vertex $o \in \Gamma$.

As a first main ingredient, we use recent breakthroughs by Tessera and Tointon [TT21, TT20] on the geometry of finite vertex-transitive graphs. Those results tell us that the balls $B(r)$ of radius r in our Γ have to grow at least as fast as r^3, up to some radius $r_0 \asymp D^{0.3}$, and then (for $r \ge r_0$) at least as fast as $r_0 r^2$. This means that the *slowest growing* graphs, among graphs of dimension 2.3, are the thin tori of side lengths $\asymp D, \asymp D, \asymp D^{0.3}$.

We then make use of the evolving set technique of Morris and Peres [MP05], to convert those bounds on the growth of balls into diagonal heat-kernel estimates. In our example case, what we obtain is that the probability to be at o at time t satisfies

$$p_t(o,o) \lesssim \frac{1}{t^{3/2}} \tag{1.1.7}$$

if $t \leq r_0^2$ and

$$p_t(o,o) \lesssim \frac{1}{r_0 t} \tag{1.1.8}$$

if $r_0^2 \leq t \leq D^2$.

Next, we use a refinement from Folz [Fol11] of the Carne–Varopoulos inequality, which also takes into account that balls can grow faster than linearly. As a result we obtain (essentially) the following off-diagonal heat kernel bounds: for $x, y \in \Gamma$,

$$p_t(x,y) \lesssim \frac{1}{t^{3/2}} \exp\left(-c\frac{d(x,y)^2}{t}\right), \tag{1.1.9}$$

if $t \leq r_0^2$, and

$$p_t(x,y) \lesssim \frac{1}{r_0 t} \exp\left(-c\frac{d(x,y)^2}{t}\right), \tag{1.1.10}$$

if $r_0 \leq t \leq D^2$, where c is some fixed constant.

Integrating those bounds, we finally deduce precise estimates on the expected local time before time t at any vertex $x \in \Gamma$

$$\mathbb{E}_o L_x(t) = \int_0^t p_t(o,x)\mathrm{d}t. \tag{1.1.11}$$

We can also deduce from the heat kernel bound (1.1.7), that whenever (DC) holds, $\mathbb{E}_o L_o(D^2)$ is finite. In other words, with positive probability, the walk doesn't come back to o before time D^2 (the mixing time). This good property of the walk is called *local transience*.

Observe that under local transience, expected local times $\mathbb{E}_o L_x(t)$ are proportional to the hitting probabilities $\mathbb{P}_o(T_x \leq t)$. We can therefore (under (DC)) precisely estimate $\mathbb{P}_o(T_x > t)$.

Another key tool is the Aldous and Brown [AB92] estimates on exponential approximation of hitting times. Vaguely stated (see Theorem 2.1.4 for precise statement), they give for any finite subset set $A \subset \Gamma$ an approximation

$$\mathbb{P}_\pi(T_A > t) \approx e^{-t/\mathbb{E}_\pi T_A}, \tag{1.1.12}$$

where π is the uniform probability measure on Γ, and $\mathbb{E}_\pi T_A$ is the hitting time of the set A from started at a uniformly random vertex. Using our estimates on $\mathbb{P}_o(T_x > t)$, we can deduce for sets A which have two points at distance $o(D)$ that even if individually their $\mathbb{P}_\pi(T_A > t)$ is larger than average, their overall contribution to the uncovered set is negligible.

Finally, we use the Bonferroni inequalities to convert our estimates on $\mathbb{P}_\pi(T_A > t)$ into estimates on $\mathbb{P}_\pi(T_A = t)$. We obtain that under (DC), we have as $|\Gamma| \to \infty$, for each $k \geq 0$ and $s \in \mathbb{R}$, uniformly over all *macroscopic* subsets $A \subset \Gamma$ of size k,

$$\mathbb{P}(U(t_s) = A) = \frac{e^{-e^{-s}} e^{-ks}}{n^k}(1 + o(1)), \tag{1.1.13}$$

where $t_s = \mathrm{t_{hit}}(\log(|\Gamma|) + s)$, which gives the desired product structure.

When (DC) is not satisfied, the Aldous–Brown approximations are not precise enough to estimate the cover time fluctuations.

One of our main technical contributions is an improvement of those estimates in Theorem 2.1.5. Those refined estimates are precise also for low dimensional graphs. Combined with our estimates on Green functions, we were able to prove that when (DC) is not satisfied, the fluctuations of the cover time cannot be Gumbel.

1.2 Mixing times

1.2.1 Preliminaries

Let us come back to our deck of cards. Our intuition tells us that the more we shuffle the deck, the better it is mixed, and that after many shuffles the deck of cards is very *close to being uniform*, that the *distance* between the distribution of our shuffled deck of cards and a uniformly random deck is small. We now see how to make such statements rigorous. There are different possible notions of distance between probability measures, of which the most natural and standard for our purpose is the following. The *total variation distance* between two probability measures μ and ν on a finite set S is defined by the formula

$$\mathrm{d_{TV}}(\mu,\nu) = \max_{A\subset S} |\mu(A)-\nu(A)|. \tag{1.2.1}$$

Elementarily but importantly, it coincides with the L^1 distance (up to a factor 2): we have

$$\mathrm{d_{TV}}(\mu,\nu) = \frac{1}{2}\sum_{x\in S} |\mu(x)-\nu(x)|\,. \tag{1.2.2}$$

Consider an irreducible aperiodic finite state Markov chain X. Let $\mu(t)$ be the distribution of X after t steps, and π be the stationary measure of X. We will only consider simple random walks on graphs, started at a point, that is when $\mu(0)$ is a Dirac measure. We define the distance to equilibrium at time t by

$$\mathrm{d}(t) = \mathrm{d_{TV}}(\mu(t),\pi).$$

The convergence theorem [LP17, Corollary 1.18] ensures that $d(\cdot)$ is (weakly) decreasing and that

$$\mathrm{d}(t) \xrightarrow[t\to\infty]{} 0. \tag{1.2.3}$$

We are somewhat reassured: if we shuffle our deck many times, it is well mixed. But what if 1000 riffle (also named faro, weave, or dovetail) shuffles are needed to mix properly a deck of 52 cards and prevent any cheating? This would be practically very inconvenient! More than knowing that $d(t)$ tends to 0 as t tends to infinity, we are therefore interested in understanding how large t should be to have reasonably well mixed, for example to bring $d(t)$ under, say, 1/10. This motivates the following definition.

Let $\varepsilon \in (0,1)$. We call (total variation) *mixing time* at level ε the first time t at which $d(t) \le \varepsilon$:

$$t_{\mathrm{mix}}(\varepsilon) = \min\left\{t\in\mathbb{N} : d(t)\le\varepsilon\right\}. \tag{1.2.4}$$

We will also (and mostly) be interested in asymptotic results, as the number of cards tends to infinity. Consider now a sequence $(X_n)_{n\in\mathbb{N}}$ of irreducible aperiodic finite state Markov chains. For each n, we define as above $\mu_n(t)$, π_n and $\mathrm{d}_n(t)$.

Let us give some examples with different behaviours.

Example 1.2.1 (The complete graph)**.** Let $o \in K_n$ and consider the simple random walk the complete graph started at o. $\mu_n(0)$ is concentrated at o, and $\mu_n(1)$ is already almost perfectly uniform, so $d_n(0) = 1 - o(1)$ and $d_n(1) = o(1)$, which means that the walk mixes in one step.

Example 1.2.2 (The cycle)**.** Consider the lazy random walk on the cycle started at 0, which goes clockwise or anticlockwise each with probability 1/4, and doesn't move with probability 1/2. It is possible to show using Fourier analysis that there exists a function ψ such that for every $\alpha > 0$,

$$d_n(\alpha n^2) \xrightarrow[n\to\infty]{} \psi(\alpha). \tag{1.2.5}$$

See [Dia88, Chapter 3C] for the technique and [Sal20] for the formula for ψ.

Remark 1.2.1. Except in some examples, we will only consider random walks on *vertex-transitive* graphs, that is graphs which have the same structure viewed from any of their vertices. For those graphs the starting point of the walk does not affect the distance to stationarity, and the stationary measure is uniform.

However, for graphs $\Gamma = (V, E)$ which are not vertex-transitive, the starting point matters, and it is standard to define the worst case distance to stationarity instead.

$$d(t) = \max_{x \in V} \mathrm{d_{TV}}(\mu_{x,t}, \pi), \tag{1.2.6}$$

where $\mu_{x,t}$ is the distribution of the simple random walk on Γ started at x after t steps.

Example 1.2.3 (Two glued complete graphs)**.** Consider two complete graphs with n vertices, $K_n^{(i)} = (V^{(i)}, E^{(i)})$, and $o^{(i)} \in V^{(i)}$, for $i \in \{1, 2\}$. Let $\Gamma_n := (V^{(1)} \cup V^{(2)}, E^{(1)} \cup E^{(2)} \cup \{\{o^{(1)}, o^{(2)}\}\})$ be the graph constructed by adding the edge $\{o^{(1)}, o^{(2)}\}$ between $K_n^{(1)}$ and $K_n^{(2)}$. The mixing time of Γ_n is analog to the cover time of the complete graph with one extra edge from Paragraph 1.1.3: for each $c > 0$,

$$d(cn^2) \xrightarrow[n\to\infty]{} e^{-c}. \tag{1.2.7}$$

1.2.2 The cutoff phenomenon

Let us give a last example, which is particularly important to us.

Example 1.2.4. Take a deck of n cards. At each step pick two cards uniformly at random and swap them. This process can be realised (up to a small correction due to laziness) as a simple random walk on the Cayley graph $\Gamma_n = (\mathfrak{S}_n, \{\{\sigma_1, \sigma_2\} : \sigma_1\sigma_2^{-1} \text{ is a transposition}\})$, where $\mathfrak{S}_n$ is the symmetric group of index n. Diaconis and Shahshahani proved in their landmark paper [DS81] that for all $\varepsilon \in (0, 1)$,

$$\mathrm{d}_n((1-\varepsilon)\frac{1}{2}n\log n) \xrightarrow[n\to\infty]{} 1 \quad \text{and} \quad \mathrm{d}_n((1+\varepsilon)\frac{1}{2}n\log n) \xrightarrow[n\to\infty]{} 0.$$

It means that the walk goes from not mixed at all to very well mixed in a short amount of time: there is a phase transition around time $\frac{1}{2}n\log n$, called *cutoff*. The method to prove this result is also very interesting. For the upper bound, Diaconis and Shahshahani used non-commutative Fourier analysis, that is group representations. Both their result and their method led to a vast amount of developments, including ours in Chapter 3.

Let us give a more general definition of the *cutoff phenomenon*. Consider again a sequence $(X_n)_{n\in\mathbb{N}}$ of irreducible aperiodic finite state Markov chains, and let $(t_n)_{n\in\mathbb{N}}$ be a sequence of times. We say that $(X_n)_{n\in\mathbb{N}}$ exhibits a *cutoff* in total variation distance at time $(t_n)_{n\in\mathbb{N}}$ (or simply t_n to lighten notations) if for all $\varepsilon \in (0,1)$,

$$\mathrm{d}_n((1-\varepsilon)t_n) \xrightarrow[n\to\infty]{} 1 \quad \text{and} \quad \mathrm{d}_n((1+\varepsilon)t_n) \xrightarrow[n\to\infty]{} 0,$$

where we skipped integer parts for readability.

Once we know that a chain exhibits cutoff at some time t_n, the next natural question is to know what happens if we zoom in around t_n.

If there exists a sequence $(w_n)_{n\in\mathbb{N}}$ and a continuous function f decreasing from 1 to 0 such that $w_n = o(t_n)$ and for all $c \in \mathbb{R}$,

$$\mathrm{d}_n(t_n + cw_n) \xrightarrow[n\to\infty]{} f(c),$$

then we say that f is the *cutoff profile* of $(X_n)_{n\in\mathbb{N}}$. Exhibiting a cutoff means that the convergence to equilibrium occurs through a sharp phase transition, falling rapidly from 1 to 0 around time t_n, while the cutoff profile describes how the distance to stationarity falls withing the *cutoff window*. Our main result from Chapter 3 is the cutoff profile for random transpositions. We will detail those results in Section 1.2.5.

1.2.3 Some history and related problems

The theory of mixing times has considerably developed since the foundational works of Diaconis and Shahshahani [DS81], Aldous [Ald83a], and Aldous and Diaconis [AD86]. We make below a non-exhaustive list of results on mixing times. See also [LP17] for an introduction and more references.

1.2.3.1 On random transpositions and related problems

The random transposition shuffle and its generalisations have attracted a lot of attention. The cutoff was proved for k-cycles (k finite) in [BSZ11], for k-cycles ($k = o(n)$) in [Hou16], and for conjugacy classes of *support size* $o(n)$ in [BŞ19]. We can also mention an interesting proof for transpositions via a strong stationary time [Whi19], and a generalisation to biased transpositions [BCMR21]. Some weaker results we also found for random involutions [Ber18].

More recently, we found in [Tey20] a way to compute the cutoff profile for random transpositions. Our technique was later generalised and led to several other cutoff profiles, including the cutoff profiles for k-cycles [NOT21], star-random transpositions [Nes21], and the k-particle interchange process on the complete graph [NOT22].

1.2.3.2 On other card mixing problems

The most emblematic shuffle is certainly the riffle shuffle. Aldous [Ald83a] proved that it has a cutoff at $\frac{3}{2}\log_2 n$. A decade later, Bayer and Diaconis [BD92] found the cutoff profile, as well as very precise estimations for finite n, in particular for $n = 52$: the distance to uniformity is a bit above 0.9 after 5 shuffles, and a bit below 0.1 after 9 shuffles.

Another common way to shuffle a deck of cards is by putting the cards on a table and *smooshing* them. It was recently modeled and studied by Diaconis and Pal [DP22]. There are other ways to shuffle a deck of cards, less convenient in practice, but mathematically very interesting. For instance, one can at each step pick a card at random and insert it at random. This is the random to random shuffle, which was proved in [BN19] to exhibit cutoff at time $\frac{3}{4}n\log n$, and for which the profile is yet to be discovered.

There are several open questions in this direction. What if we take two permutations $g, h \in \mathfrak{S}_n$ uniformly at random? It was proved in [Bab89] that they generate the symmetric group with probability $1 - o(1)$. Whenever g and h generate the symmetric group, what can we say about the geometry of the associated Cayley graph, whose vertex set is $\mathfrak{S}_n$ and has an edge between two permutations σ_1 and σ_2 if and only if $\sigma_1\sigma_2^{-1} \in \{g, h, g^{-1}, h^{-1}\}$? What is the typical behaviour of the mixing time on such a random graph? The best known estimates, in [HSZ15], show that the mixing time is typically $O(n^2(\log n)^c)$ for some constant c. Even in concrete cases this problem is very hard. The worst expected case is when $g = (1\,2)$ and $h = (1\,2\ldots n)$, and it is called the Rudvalis shuffle. Wilson proved in [Wil03] that its mixing time is $\Theta(n^3 \log n)$, but it is still an open question to prove that this shuffle exhibits cutoff.

1.2.3.3 On continuous groups

It is also possible to study mixing times on continuous objects, such as the unitary group $U(n)$ or the (special) othogonal group $SO(n)$. For instance, Méliot [Mé14] proved the cutoff phenomenon for the diffusion on all classical simple compact Lie groups and several other objects associated to them. Freslon [Fre19] also proved the cutoff phenomenon for the diffusion on the quantum orthogonal group O_n^+. In Chapter 4, we push the analysis further and find the profile for the diffusion on O_N^+, which involves laws from free probability with atoms. On classical Lie groups such as $U(n)$, the profiles are still unknown.

1.2.3.4 On models of statistical physics

Another way to see the random transposition shuffle is as an interchange process on the complete graph. On other graphs this process is very complicated to study and precise results are known only for a handful of graphs. On a segment it corresponds to adjacent transpositions, and the cutoff was proved by Lacoin [Lac16b]. The order of magnitude of the mixing time was also recently found on the hypercube [HS21].

More is known for the simple exclusion process, where particles have only two labels. We can mention for example the cutoff profile for the (symmetric) simple exclusion process on the circle [Lac16a], and the cutoff profile for the totally asymmetric simple exclusion process on the segment [BN22]. There was recently progress on variants of those models on the circle [SS], and on the segment with boundaries [Sch, SE].

Cutoffs were also proved for the Gibbs samplers on models of statistical physics, such as the Ising model [LS13, LS17] or the discrete Gaussian free field [GG]. Let us finally mention that a general technique was recently developed to obtain bounds on the mixing time of such walks [CE22].

1.2.3.5 On the general theory

It was proved that random walks exhibit cutoff on large classes of graphs. This was done for example for Ramanujan graphs [LP16a], or on random graphs [BLPS18]. An interesting open problem in this direction is to show that the cutoff phenomenon always occurs for vertex-transitive (bounded-degree) expanders.

Mixing times were also related to other canonical parameters associated to random walks, such as hitting times, first in [PS15] and then in [BHP17, HP18], or the entropy of the walks [Sal22]. The relations between mixing times in different distances are fairly well understood [CSC08, HLP16].

1.2.4 Representation theory of the symmetric group applied to random transpositions

In this section we introduce some ideas about the representations of the symmetric group. We will skip the algebraic framework (we refer to [Mé17] for a complete exposition of the representations of finite groups) and focus more on how one can use representation theory to study random walks on the symmetric group. We will illustrate this on the example of random transpositions.

The Fourier transform is an amazing tool in analysis. It can be generalised to finite groups, where frequences correspond to irreducible representations. In the specific case of the symmetric group, the Frobenius–Schur isomorphism ensures that there is a (bijective) correspondence between irreducible representations of $\mathfrak{S}_n$ and integer partitions $\lambda \vdash n$. We will therefore make a standard abuse of notations and use λ both for the partition and its corresponding irreducible representation.

To each representation λ are associated a dimension d_λ and a character $\mathrm{ch}^\lambda(\cdot) : \mathfrak{S}_n \to \mathbb{C}$.

A key point is that to study Markov chains such as random transpositions, one only needs information on d_λ and $\mathrm{ch}^\lambda(\cdot)$, and that both quantities can be expressed by combinatorial formulas, the hook-length formula and the Murnaghan–Nakayama formula respectively. Estimating dimensions is usually easy, but bounding characters can be much more challenging.

We will now briefly describe the arguments from [DS81] for random transpositions, for the result which we described in Example 1.2.4. The random transposition shuffle on a deck of n cards can be represented as the Markov chain X (we remove the dependence in n to lighten notations) on the symmetric group $\mathfrak{S}_n$ (started at the identity) such that $X_0 = Id$ and $X_t = \sigma_1 ... \sigma_t$, where the σ_i are i.i.d. random variables following the law P_n defined by

$$P_n(Id) = \frac{1}{n} \text{ and } P_n(\tau) = \frac{2}{n^2} \text{ if } \tau \text{ is a transposition.}$$

P_n is called the increment measure of the walk. The law μ_t of X_t is therefore the convolution product P_n^{*t}. Moreover, since P_n is constant on conjugacy classes (that is $P_n(ghg^{-1}) = P_n(h)$ for any $g, h \in \mathfrak{S}_n$), by the Schur lemma its Fourier transform satisfies, for each irreducible representation λ

$$\widehat{P_n}(\lambda) = \alpha_\lambda I_{d_\lambda} \tag{1.2.8}$$

with $\alpha_\lambda = \left(\frac{1}{n} + \frac{n-1}{n}\chi^\lambda(\tau)\right)$, where $\chi^\lambda(\cdot) = \frac{\mathrm{ch}^\lambda(\cdot)}{d_\lambda}$ is the reduced character, and $\tau \in \mathfrak{S}_n$ is any transposition. Note that if we remove the laziness $1/n$ (by picking two different cards at each step), we simply end up with $\alpha_\lambda = \chi^\lambda(\tau)$.

Applying the Cauchy–Schwarz inequality in (1.2.2), i.e. bounding the L^1-distance by the L^2-distance, and then applying the Parseval identity, we obtain, denoting by $\widehat{\mathfrak{S}_n}^*$ the set of nontrivial irreducible representations of $\mathfrak{S}_n$, for any $t \geq 0$

$$4d(t)^2 \leq d^{(2)}(t)^2 := n! \sum_{\sigma \in \mathfrak{S}_n} \left(\mu_t(\sigma) - \frac{1}{n!}\right)^2 = \sum_{\lambda \in \widehat{\mathfrak{S}_n}^*} (d_\lambda \alpha_\lambda^t)^2. \tag{1.2.9}$$

This last step is what is called the Diaconis–Shahshahani upper bound lemma. The rest of the proof of the upper bound constists of technical estimates based on the hook-length formula for d_λ, and on the formula

$$\chi^\lambda(\tau) = \frac{1}{\binom{n}{2}} \left(\sum_i \binom{\lambda_i}{2} - \binom{\lambda_i'}{2} \right). \tag{1.2.10}$$

The lower bound is probabilistic and much simpler: at time $t(\varepsilon) = (1-\varepsilon)\frac{1}{2}n\log n$, by a coupon-collector argument there are still around n^{ε} cards that have never been touched. Letting

$$A_{\varepsilon} := \left\{\sigma \in \mathfrak{S}_n : \sigma \text{ has at least } n^{\varepsilon/2} \text{ fixed points } \right\}, \tag{1.2.11}$$

we have $\mu_{t(\varepsilon)}(A_{\varepsilon}) = 1 - o(1)$ and $\pi(A_{\varepsilon}) = o(1)$ so by the definition (1.2.1) of the total variation distance, we have $d((1-\varepsilon)\frac{1}{2}n\log n) = 1 - o(1)$.

1.2.5 Profile for random transpositions: main result of Chapter 3

The main result of Chapter 3 is the following theorem.

Theorem 1.2.1 (Theorem 3.1.1)**.** *Let $c \in \mathbb{R}$, and set $t_{n,c} = \frac{1}{2}n\log(n) + cn$. Then we have,*

$$d_n\left(t_{n,c}\right) \xrightarrow[n\to\infty]{} d_{TV}\left(\text{Poiss}\left(1+e^{-2c}\right), \text{Poiss}(1)\right), \tag{1.2.12}$$

where Poiss(a) stands for the Poisson law of parameter a.

There are two main ingredients to prove this result. The first one is to obtain a sharper version of the Diaconis–Shahshahani upper bound lemma. Indeed, with computations similar to those in [DS81], it is possible to find the profile for the L^2 distance to stationarity (which is nothing else than the right hand side of (1.2.9)):

$$d^{(2)}(t_{n,c}) \xrightarrow[n\to\infty]{} \sqrt{e^{e^{-4c}} - 1}. \tag{1.2.13}$$

The right hand side of (1.2.13) is larger than that of (1.2.12), and can therefore not be used to bound sharply the total variation distance to stationarity within the cutoff window. In other words, to find this profile, we cannot bound the L^1 distance by the L^2 distance. What we can do instead is to use the inverse Fourier transform, which allows us to write directly

$$d^{(1)} = 2d(t) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda\in\widehat{\mathfrak{S}_n}^*} d_\lambda \alpha_\lambda^t \mathrm{ch}^\lambda(\sigma)\right|. \tag{1.2.14}$$

This is a complicated sum, since the inner sum depends also on σ. We therefore need to consider combinatorial formulas for $\mathrm{ch}^\lambda(\sigma)$, *a priori* for all λ and σ. However, by an analysis similar to that of [DS81], we can show that the only partitions λ which matter are those for which $n - \lambda_1$ is finite, that is, seeing λ as a Young diagram, for which the first line of λ is very long.

The second main ingredient is to understand the compensations between the characters of the symmetric group, or in other words the interactions between the eigenvectors of the transition matrix of the walk. An important observation is that most permutations have a long cycle, and that for those permutations the Murnaghan–Nakayama formula (which gives an explicit formula for $\mathrm{ch}^\lambda(\sigma)$) writes nicely. This allows us to express for most $\sigma \in \mathfrak{S}_n$ a weighted sum of characters as a polynomial depending only on the number $\mathrm{Fix}(\sigma)$ of fixed points. Let us denote by λ^* the truncated partition of λ, that is, if $\lambda = (\lambda_1, \lambda_2, ...)$ then $\lambda^* = (\lambda_2, \lambda_3, ...)$. Set

$$S_j(\sigma) := \sum_{\lambda\in\widehat{\mathfrak{S}_n}\,:\,\lambda_1 = n-j} d_{\lambda^*}\mathrm{ch}^\lambda(\sigma), \tag{1.2.15}$$

and define the polynomial T_j (on integers) by

$$T_j(z) = \sum_{i=0}^{j} \binom{z}{j-i} \frac{(-1)^i}{i!}. \tag{1.2.16}$$

Our polynomial identity for characters (Lemma 3.4.3) states that for most permutations, we have

$$\frac{1}{j!} S_j(\sigma) = T_j(\mathrm{Fix}(\sigma)). \tag{1.2.17}$$

We present in Section 3.5 a detailed proof of this result.

The rest of the proof is computational, and the Poisson laws arise naturally as the number of fixed points of random permutations.

1.2.6 Cutoff profile for the quantum orthogonal group: main result of Chapter 4

Let $O_N = \{M \in \mathcal{M}_N(\mathbb{R}) : {}^tMM = Id\}$ be the (classical) orthogonal group and

$$c_{ij} : M \mapsto [M]_{i,j}$$

be its N^2 coefficient functions. Let $\mathcal{O}(O_N)$ be the $\mathbb{C}$-algebra of *regular functions* on O_N, i.e. the algebra generated by the (restrictions to O_N of) the functions c_{ij}. We endow the algebra $\mathcal{O}(O_N)$ with an involution $*$, the complex conjugation, making it a commutative $*$-algebra.

Let $A := \mathbb{C}\langle u_{ij} : 1 \leqslant i,j \leqslant N \rangle$ be the $*$-algebra of non-commutative polynomials in the indeterminates $(u_{ij})_{1\leqslant i,j\leqslant N} =: U$, with $u_{ij}^* = u_{ij}$. Let I be the $*$-ideal of A generated by the coordinates of the matrices $U{}^tU - Id$ et ${}^tUU - Id$.

The **algebra of the orthogonal quantum group** is the quotient $*$-algebra

$$\mathcal{O}(O_N^+) := A/I.$$

Informally, $\mathcal{O}(O_N^+)$ is the algebra of non-commutative polynomials in an orthogonal matrix of indeterminates, $U = (u_{ij})_{1\leqslant i,j\leqslant N}$. Note that $\mathcal{O}(O_N)$ is the abelianisation of $\mathcal{O}(O_N^+)$, so in this regard quantum groups can be seen as a generalisation of classical groups.

We can also define a Brownian motion on those continuous objects, and defined a *quantum* distance to stationarity $\widetilde{\mathrm{d}}_N$. We refer to Chapter 4 for more comprehensive definitions.

The main result of Chapter 4 is the cutoff profile for the diffusion on O_N^+.

Theorem 1.2.2 (Theorem 4.3.9)**.** *For the diffusion on O_N^+, we have at time $t_c = N\ln N + cN$*

$$\widetilde{\mathrm{d}}_N(t_c) \xrightarrow[N\to\infty]{} d_{TV}\left(\mathrm{Meix}^+\left(-e^{-c}, 0\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,0)\right),$$

where Meix^+ *is the free Meixner law.*

The second main result is the profile for a quantum version of random transpositions.

Theorem 1.2.3 (Theorem 4.4.4)**.** *For quantum transpositions on the quantum symmetric group S_N^+, we have at time $t_c = \frac{1}{2}N\ln N + cN$*

$$\widetilde{\mathrm{d}}_N(t_c) \xrightarrow[N\to\infty]{} d_{TV}\left(D_{\sqrt{1+e^{-c}}}\left(\mathrm{Meix}^+\left(\frac{1-e^{-c}}{\sqrt{1+e^{-c}}}, \frac{-e^{-c}}{1+e^{-c}}\right)\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,1)\right).$$

Let us give some ideas on the proof of Theorem 1.2.2. The proof of Theorem 1.2.3 is similar. We use the simple structure of the representations of $\mathcal{O}(O_N^+)$. Indeed, they can be indexed by integers, and they verify the same recurrence relation as Chebyshev polynomials, so they are more tractable than those of SO_N for example. In this non-commutative setting we however have to replace probability measures by states, and find an appropriate analogue of Brownian motion. This leads to extra difficulties: before time $N \ln N + o(1)$, i.e. in the middle of the cutoff window, the density of the Brownian motion is not absolutely continuous with respect to the Haar state. This can be observed in the formula of the cutoff profile: the free Meixner law $\mathrm{Meix}^+(-e^{-c}, 0)$ has an atom when $c < 0$. Nevertheless, the profile is still continuous as a function of c.

1.3 Work in progress

Let $\mathcal{C}$ be a conjugacy class of $\mathfrak{S}_n$. Denote by $f = f(\mathcal{C})$ the number of fixed points of any $\sigma \in \mathcal{C}$ and by $k = n - f$ its support size. Call $\mathcal{C}$-shuffle the random walk on $\mathfrak{S}_n$ which starts at $\frac{1}{2}(\delta_{Id} + \delta_{(1\,2)})$, and at each step multiplies by a permutation picked uniformly at random in $\mathcal{C}$.

It was proved in [BŞ19] that if $k = o(n)$, the $\mathcal{C}$-shuffle exhibits cutoff at time $\frac{n}{k} \log n$, and if $\mathcal{C}$ consists of k-cycles (still $k = o(n)$), the profile is also known [NOT21].

On the other side of the spectrum, when $k \asymp n$, only partial results are known. Roichman found in [Roi96] estimates on characters via the Murnaghan–Nakayama rule, and deduced the correct order of magnitude of the mixing time when both k and f are of order n. Larsen and Shalev later developed in [LS08] the technique of virtual characters and gave sharp estimates on characters $\mathrm{ch}^\lambda(\sigma)$ when σ has very few fixed points. It allowed them to find the mixing time when $f = n^{o(1)}$, extending works from Lulov and Pak [LP02] and proving one of their conjectures. The picture is however still very incomplete.

In addition to the above works we include a preview and short description of some ongoing work with Sam Olesker-Taylor, which however was not sufficiently polished to be included in this thesis. We developed a technique to estimate characters which is much more precise. It already enabled us to find the mixing time profile for k-cycles for any $k \asymp n$, and therefore to complete the picture for all $2 \leq k \leq n$.

We are currently generalising our results to all conjugacy classes. We believe that our estimates will allow us to understand the mixing time sharply (including the profile) for all conjugacy classes $\mathcal{C}$ of support size $k \asymp n$, and in particular prove that whenever $t(n, f) := \frac{\log n}{\log(n/f)}$ diverges, the $\mathcal{C}$-shuffle exhibits cutoff at time $t(n, f)$. We also believe that the profile depends also on the number of transpositions: that for example the $\mathcal{C}$-shuffles corresponding to the cycle types $[n/2, 1^{n/2}]$ and $[2^{n/4}, 1^{n/2}]$ will both exhibit cutoff at time $\log_2(n)$, but with different profiles.

Chapter 2

On the universality of fluctuations for the cover time

We consider random walks on finite vertex-transitive graphs Γ of bounded degree. We find a simple geometric condition which characterises the cover time fluctuations: the renormalised cover time $\frac{\tau_{\text{cov}}}{\mathrm{t}_{\text{hit}}} - \log|\Gamma|$ converges to a standard Gumbel variable if and only if $\text{Diam}(\Gamma)^2 = o(n/\log n)$, where $n = |\Gamma|$. We prove that this condition is furthermore equivalent to the decorrelation of the uncovered set. The arguments rely on recent breakthroughs by Tessera and Tointon on finitary versions of Gromov's theorem on groups of polynomial growth, which we leverage into strong heat kernel bounds, and refined quantitative estimates on Aldous and Brown's exponential approximation of hitting times, which are of independent interest.

2.1 Introduction

2.1.1 Context

Let Γ be a finite (connected) vertex-transitive graph. Let $(X_t)_{t\geq 0}$ denote the simple random walk in continuous time, which at constant rate 1 jumps to a randomly chosen neighbour, starting from a designated vertex called the root and denoted by o. Let us denote by $T_x = \inf\{t \geq 0 : X_t = x\}$ the hitting time of the vertex x. The **cover time** variable of Γ by X is the first time that the walk has visited every vertex:

$$\tau_{\text{cov}} = \max_{x\in\Gamma} T_x,$$

where with a slight abuse of notation we have used Γ to also denote the vertex set of the graph Γ. In this article we are concerned with obtaining general fluctuation results for the cover time τ_{cov} of random walks on vertex-transitive graphs, as the vertex set size $|\Gamma|$ tends to infinity. As we will see this question is deeply intertwined with the study of the structure of the uncovered set $U(t) = \{x \in \Gamma : T_x > t\}$, for t close to the (expected) cover time.

Obtaining quantitative estimates on the cover time is a natural problem which parallels the intensively studied question of mixing time and cutoff for random walks on graphs (i.e. understanding how far the law of X_t is from the stationary distribution, see Section 2.2.2 for some definitions, and [LP17] and the references therein for an introduction). In common with much of the literature on the subject, we will focus in this paper on vertex-transitive graphs. The restriction to this class of graphs is natural in order to avoid pathological examples, which can be arbitrarily badly behaved. At the same time it allows for a very rich range of behaviours, as we are about to discuss.

Results in this direction go back at least to the seminal work of Aldous [Ald83b] who proved that for random walks on finite groups and under mild geometric conditions, the cover time τ_{cov} is concentrated around its mean $\mathrm{t}_{\text{cov}} = \mathbb{E}(\tau_{\text{cov}})$, i.e. that $\frac{\tau_{\text{cov}}}{\mathrm{t}_{\text{cov}}} \to 1$ in probability, and obtained the leading order behaviour of t_{cov}. This was complemented a few years later by Matthews [Mat88a] who obtained a general upper bound (valid for *any* graph) on the expected cover time. The problem of the correlation structure of the uncovered set was raised in the physics literature through the work of Brummelhuist and Hilhorst [BH91, BH92].

It is worth noting that even for basic graphs such as the d-dimensional torus of side-length n, i.e. when $\Gamma = (\mathbb{Z}/n\mathbb{Z})^d$, the problem is highly nontrivial. It was only in 2004 that the first order of the cover time for the two-dimensional torus $(\mathbb{Z}/n\mathbb{Z})^2$ was found by Dembo, Peres, Rosen and Zeitouni in [DPRZ04]. This result was later refined, see [BK17, Abe21, BRZ20], but obtaining a convergence in distribution for the fluctuations of the cover time in two dimensions remains an open problem to this day (it is widely believed however that the fluctuations will be in any case different from the higher dimensional regime described below).

In dimensions $d \geq 3$, it was proved by Belius [Bel13] in 2013 (solving an open question of Aldous and Fill [AF]), building on Sznitman's random interlacement model [Szn10] (who was in fact motivated by the work of Brummelhuist and Hilhorst [BH91, BH92]), that the fluctuations of the cover time are asymptotically distributed according to a standard Gumbel law:

$$\mathbb{P}\left(\frac{\tau_{\text{cov}}}{\mathrm{t}_{\text{hit}}} - \log|\Gamma| \leq s\right) \to e^{-e^{-s}},$$

where $\mathrm{t}_{\text{hit}} := \max_{x,y\in\Gamma} \mathbb{E}_x\left[T_y\right]$ is the maximal expected hitting time. (See also the results in Prata's thesis [PdP12] for partial results valid for more general graphs but under some restrictive assumptions.)

The Gumbel law is significant because it describes the asymptotic maximum of i.i.d. random variables, subject to some conditions on their common distribution. The Gumbel fluctuations in the result above therefore suggest that the uncovered set at time $t_s := \mathrm{t}_{\text{hit}}(\log(|\Gamma|) + s)$ might asymptotically be close to a product measure, where each vertex is uncovered with probability $e^{-s}/|\Gamma|$ independently of other vertices. More formally, if μ_s is a Bernoulli variable of parameter $e^{-s}/|\Gamma|$, we might expect that as $|\Gamma| \to \infty$,

$$\mathrm{d}_{\text{uncov}}(t_s) := \mathrm{d}_{\text{TV}}\left(\mathcal{L}(U(t_s)), \mu_s^{\otimes\Gamma}\right) \to 0, \tag{2.1.1}$$

where $\mathcal{L}(U(t_s))$ denotes the law of the uncovered set at time t_s, and the total variation distance between two probability measures μ and ν on a finite space S is given by

$$\mathrm{d}_{\text{TV}}(\mu,\nu) = \max_{A\subset S}|\mu(A) - \nu(A)|.$$

Part of the goal of this paper is to study the uncovered set and in particular prove (2.1.1) in a general framework.

We note that questions concerning the geometry of the uncovered set (even when the cover time is well understood) have recently become prominent. Even on the d-dimensional torus, a number of basic problems remain open. For instance, it was proved in [MS17] and [OTS20] that the uncovered set at time $a\,\mathrm{t}_{\text{cov}}$ is decorrelated in the above sense if $a > 7/8$ and correlated (in the sense that (2.1.1) does not hold) if $a < \frac{1+p_d}{2}$, where $p_d = \mathbb{P}_o(T_o^+ < \infty)$ for the simple random walk on $\mathbb{Z}^d$. It nevertheless still remains an open question whether there actually is a phase transition, i.e. whether there exists a critical value a^* such that for $a \neq a^*$, $\mathrm{d}_{\text{uncov}}(a\,\mathrm{t}_{\text{cov}}) \to 1_{a<a^*}$.

More generally, the study of the uncovered set fits into the theme of exceptional points for random walks. In two dimensions the structure of those exceptional points has recently been proved to be linked with Liouville quantum gravity, see [AB22], and, away from the cover time, to an even more singular and in some way intriguing object called Brownian multiplicative chaos, see [Jeg20].

Finally, we note that the expected cover time $\mathrm{t}_{\mathrm{cov}}$ of a graph is also closely related to the typical value of the maximum of the associated Gaussian free field, see for instance [DLP12], [Din14] and [Zha18].

2.1.2 Main results

Let Γ be a finite (connected) bounded degree vertex-transitive graph. Write Γ also for the vertex set of Γ as above. We denote by $d(x, y)$ the graph distance between the vertices x and y, by $D := \max_{x,y\in\Gamma} d(x, y)$ the diameter of Γ, and by π its stationary distribution.

Our first main results shows that Gumbel fluctuations are universal. Perhaps even more surprisingly we obtain a sharp (necessary and sufficient) geometric condition for this universality.

Theorem 2.1.1. Let (Γ) be a collection of finite (connected) vertex-transitive of fixed degree d, and let χ be a standard Gumbel variable, i.e. $\mathbb{P}(\chi \le s) = e^{-e^{-s}}$ for $s \in \mathbb{R}$. Then

$$\frac{\tau_{\mathrm{cov}}}{\mathrm{t}_{\mathrm{hit}}} - \log|\Gamma| \xrightarrow[|\Gamma|\to\infty]{(d)} \chi \tag{2.1.2}$$

if and only if

$$\frac{D^2 \log|\Gamma|}{|\Gamma|} \xrightarrow[|\Gamma|\to\infty]{} 0. \tag{DC}$$

For future reference we note that, trivially, if we let $n = |\Gamma|$ denote the number of vertices of Γ (which tends to infinity by assumption), then (DC) is equivalent to $D^2 = o(n/\log n)$, and (2.1.2) is equivalent to the condition that for all $s \in \mathbb{R}$, as $n \to \infty$,

$$\mathbb{P}(\tau_{\mathrm{cov}} \le \mathrm{t}_{\mathrm{hit}}(\log n + s)) \to \exp(-e^{-s}).$$

We complement this result with a refined statement on the structure of the uncovered set.

Theorem 2.1.2. Let (Γ) be a collection of finite (connected) vertex-transitive graphs of fixed degree d, and $s \in \mathbb{R}$. For $s \in \mathbb{R}$, let $t_s := \mathrm{t}_{\mathrm{hit}}(\log(|\Gamma|) + s)$ and let $\mu_s^{\otimes\Gamma}$ denote the product over all the vertices of the graph of the Bernoulli law μ_s with parameter $e^{-s}/|\Gamma|$. Then

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{L}(U(t_s)), \mu_s^{\otimes\Gamma}) \xrightarrow[|\Gamma|\to\infty]{} 0 \tag{2.1.3}$$

for every $s \in \mathbb{R}$ if and only if the diameter condition (DC) holds.

This result will be strengthened in various ways under the assumption (DC) in Section 2.4. For instance, in Theorem 2.4.2 we discuss the law of the uncovered set at the first time that exactly k points remain to be covered.

In particular, we recover that (2.1.1) holds for the tori $\mathbb{Z}^d$, for $d \ge 3$ (this is in fact already mentioned in the thesis of Prata [PdP12], although the result was never published).

To get a feel for the condition (DC), the reader might want to keep in mind a "thin" torus where $\Gamma = (\mathbb{Z}/m\mathbb{Z})^2 \times (\mathbb{Z}/h\mathbb{Z})$. This graph is nothing but a box (with periodic boundary conditions) of sidelength m and height $h = h_m$, where we assume that $1 \le h_m \le m$. (Thus the extreme cases

$h_m = 1$ and $h_m = m$ correspond to the familiar two- and three-dimensional situations respectively, while very little is known in general about intermediate situations). Then the diameter $D = m$, and the volume is $n = m^2 h_m$, so the condition (DC) holds if and only if $h_m \ggg \log m$ (we will write $a_n \ggg b_n$ when $a_n/b_n \to \infty$ as $n \to \infty$) .

The reader might find it surprising that the condition (DC) determines the behaviour (2.1.2) (and (2.1.3)). It would indeed be natural to expect that the local structure of the graphs plays an important role in the decorrelation of the uncovered set. Theorem 2.1.2 tells us that it is not the case: decorrelation in the uncovered set depends only on the relation of the diameter of the graph to its size, a completely global geometric condition.

We also point out that despite considerable effort over the last 50 years, in the parallel problem of the mixing time mentioned above, a simple characterisation of the cutoff phenomenon (and even more so of the limiting profile) is currently completely out of reach.

In the proof that (DC) is necessary for (2.1.3), we will see that even the first moment of $U(t_s)$ is different from that of $\mu_s^{\otimes\Gamma}$. In that sense the theorem above does not capture the correlations of the uncovered set close to the cover time. However, there are other natural ways to adjust the time scaling, for instance by considering (for any $s \in \mathbb{R}$) the uncovered set at the time t_s^* such that $|U(t_s^*)| = e^{-s}$. Note that such changes of normalisation would not affect the above results under (DC). That is, under (DC), both (2.1.2) and (2.1.3) hold with t_s replaced t_s^*. Indeed, we will prove in Corollary 2.2.2 and Proposition 2.2.4 that under the assumption (DC), we have $t_s = t_s^* + o(n)$, where $n = |\Gamma|$ is the number of vertices of Γ. It is therefore natural to ask whether $U(t_s^*)$ is close to a product measure. For this as well, we will show that the diameter condition (DC) is the sharp criterion for decorrelation. However, in order to state this stronger result we will need to make an additional geometric assumption.

Theorem 2.1.3. Let (Γ) be a collection of finite (connected) vertex-transitive graphs of fixed degree d, and let $n = |\Gamma|$. Suppose that $D^2/n \to 0$. For $s \in \mathbb{R}$, let $\mu_s^{\otimes\Gamma}$ denote the product over all the vertices of the graph of the Bernoulli law μ_s with parameter $e^{-s}/|\Gamma|$. Fix $s \in \mathbb{R}$. The following are equivalent:

(i) $\mathrm{d_{TV}}(\mathcal{L}(U(t_s^*)), \mu_s^{\otimes\Gamma}) \xrightarrow[|\Gamma|\to\infty]{} 0$,

(ii) The diameter condition (DC) holds.

Remark 2.1.1. In fact, we will prove an even stronger result: namely, the conclusion is valid as soon as $t_{\mathrm{rel}} = o(\mathrm{t_{hit}})$, where t_{rel} is the *relaxation time*, or inverse spectral gap. We will also show that these conditions are equivalent to a third one, where instead of t_s^* we use the time $t_{\langle s\rangle} := t_{\langle\mathrm{hit}\rangle}(\log n + s)$, where $t_{\langle\mathrm{hit}\rangle} := \mathbb{E}_\pi T_o$; this is in fact an integral part of our proof. The same statement also holds with t_s instead of t_s^*. In other words, if $t_{\mathrm{rel}} = o(\mathrm{t_{hit}})$ holds, we can strengthen Theorem 2.1.2, replacing "for every $s \in \mathbb{R}$" by "for some $s \in \mathbb{R}$". The theorem will be proved in these forms in Section 2.6.3, where we will also briefly explain why the assumption $D^2/n \to 0$ implies $t_{\mathrm{rel}} = o(\mathrm{t_{hit}})$.

We conjecture in fact that the theorem is valid without any assumption on t_{rel} or $\mathrm{t_{hit}}$.

Central to the arguments of this paper will be exponential approximations for hitting times of arbitrary sets of vertices (irrespective of the relative geometry of the points). Such exponential approximations go back to the work of Aldous and Brown [AB92], who obtained quantitative error bounds which we recall below. However, our proofs that the diameter condition (DC) is sharp require us crucially to strengthen the quantitative bounds in these approximations, as we now detail.

Let us recall that t_{rel} is the relaxation time of the chain and define, for $A \subset \Gamma$, the A-**quasistationary distribution** by, for $x \in \Gamma$,

$$\alpha_A(x) = \lim_{t\to\infty} \mathbb{P}(X_t = x \mid T_A > t). \tag{2.1.4}$$

(See, e.g., (3.86) in [AF] for a proof of the existence of this limit). It is not hard to see that, starting from the quasi-stationary distribution α_A, the hitting time of A is exactly an exponential random variable. A seminal result of Aldous and Brown in this direction is the following.

Theorem 2.1.4 ([AB92], Theorem 3 and Equation (1))**.** Let $(X_t)_{t\geq 0}$ be an irreducible reversible Markov chain on a finite state space V, and π be its stationary distribution. Let $A \subset V$, and denote by T_A the first hitting time of A. Then for all $t > 0$,

$$0 \leq 1 - \frac{\mathbb{P}_\pi(T_A > t)}{\mathbb{P}_{\alpha_A}(T_A > t)} = O\left(\frac{t_{\mathrm{rel}}}{\mathbb{E}_{\alpha_A} T_A}\right), \tag{2.1.5}$$

and

$$0 \leq 1 - \frac{\mathbb{E}_\pi T_A}{\mathbb{E}_{\alpha_A} T_A} = O\left(\frac{t_{\mathrm{rel}}}{\mathbb{E}_{\alpha_A} T_A}\right). \tag{2.1.6}$$

Among other things, these approximations are very useful to estimate the capacity $q_A := \frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A}$ of finite sets A, which play a crucial role in our analysis. Indeed, when the diameter condition (DC) holds, the error term above is $o(1/\log n)$, which will turn out to be sufficient. However, as soon as (DC) fails, this error term becomes too large (even to estimate the expected number of uncovered points). To prove our results when (DC) does not hold, we will instead rely on the following theorem, which is a refinement of Theorem 2.1.4, and which we view as the main technical innovation of our paper. We expect the error bounds be sharp up to a constant factor for *any* family of vertex-transitive graphs of fixed degree (without further assumptions).

Theorem 2.1.5. Let (Γ) be a collection of finite vertex-transitive graphs of fixed degree d, and let $k \geq 1$. Then, uniformly over all sets $A \subset \Gamma$ of size k, and $t \geq 0$, we have as $n = |\Gamma| \to \infty$

$$0 \leq 1 - \frac{\mathbb{P}_\pi(T_A > t)}{\mathbb{P}_{\alpha_A}(T_A > t)} = O_{d,k}\left(\frac{D^4}{(\mathbb{E}_\pi T_0)^2}\left(1 + \frac{n}{D^4}\int_0^D \frac{s^3 \mathrm{d}s}{V(s)}\right)\right), \tag{2.1.7}$$

and

$$0 \leq 1 - \frac{\mathbb{E}_\pi T_A}{\mathbb{E}_{\alpha_A} T_A} = O_{d,k}\left(\frac{D^4}{(\mathbb{E}_\pi T_0)^2}\left(1 + \frac{n}{D^4}\int_0^D \frac{s^3 \mathrm{d}s}{V(s)}\right)\right), \tag{2.1.8}$$

where $V(s)$ denotes the volume of the ball of radius $\lfloor s \rfloor \geq 0$. Moreover, the right-hand sides of (2.1.7) and (2.1.8) are $O_{d,k}\left(t_{\mathrm{rel}} / \mathrm{t}_{\mathrm{hit}}\right)^2$ if $D \gtrsim n^{1/4}\log n$ and $O_{d,k}(1/n)$ if $D \lesssim \frac{n^{1/4}}{\log n}$.

The full version of this result will be stated in Theorem 2.5.1 and Theorem 2.5.3. See also the other results in that section for complements.

We believe that those technical improvements will prove useful also for other problems, for instance in order to establish the transition phase of the uncovered set at time $a\,\mathrm{t}_{\mathrm{cov}}$ for $0 < a < 1$ on tori of dimension $d \geq 3$, which would answer the questions raised by Miller and Sousi in [MS17] (see also [OTS20]).

This type of improvement is fundamental to handle borderline cases. It is in fact even significant for 2-dimensional tori: if $\Gamma = (\mathbb{Z}/m\mathbb{Z})^2$, then this error term is $O(1/(\log n)^2) = o(1/\log n)$, compared to $O(1/\log n)$ in (2.1.5) and (2.1.6).

The setup of our improvements is actually fairly general: not all statements require the graphs Γ to be vertex-transitive, and the size of the sets A can be allowed to diverge. The interested reader can find more precise statements in Section 2.5.

2.1.3 Diameter condition and local transience

We have already mentioned that the sharpness of the diameter condition (DC) is a little surprising. Initially (and this was in fact our own belief when we began this work), one might have suspected that Gumbel fluctuations are perhaps more naturally linked with the following notions of local transience which we now define. For this it will be useful to recall the definition of the mixing time $t_{\text{mix}}(\varepsilon)$ at level $0 < \varepsilon < 1$ for a positive recurrent Markov chain $(X_t)_{t\in\mathbb{R}_+}$ on some state space Γ with invariant distribution π:

$$t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : \max_{x\in\Gamma} \mathrm{d}_{\text{TV}}(P_t(x,\cdot);\pi(\cdot)) \leq \varepsilon\}, \tag{2.1.9}$$

where $P_t(x,\cdot)$ denotes the law of the Markov chain at time t starting from $x \in \Gamma$.

Definition 2.1.6. We say that a sequence of finite Markov chains with state spaces Γ_n, transition matrices $P = P^{(n)}$ and stationary distributions $\pi = \pi_n$ satisfying $\lim_{n\to\infty} \max_{x\in\Gamma_n} \pi(x) = 0$ is *weakly uniformly transient* (**WUT**), or *uniformly locally transient*, if

$$\max_{o\in\Gamma_n} \mathbb{E}_o(L_o(t)) = O(1), \tag{2.1.10}$$

where $t = t_{\text{mix}}(1/4)$, and $L_x(t) := \int_0^t 1_{\{X_s=x\}}\mathrm{d}s$ is the *local time* of the walk at x up to time t. Again, writing $t = t_{\text{mix}}(1/4)$, we say that the sequence is (**SUT**) *strongly uniformly transient*, or *uniformly globally transient*, if

$$\lim_{s\to\infty} \limsup_{n\to\infty} \max_{o\in V_n} \mathbb{E}_o(L_o(t) - L_o(s)) = 0. \tag{2.1.11}$$

We say that a sequence of graphs $\Gamma_n := (V_n, E_n)$ is WUT (resp. SUT) if the sequence of simple random walks on Γ_n is WUT (resp. SUT).

Clearly, (2.1.11) implies (2.1.10). The reason why one might suspect that this notion might be related to Gumbel fluctuations is that uniform transience, especially strong uniform transience, should prevent clusterisation and hence lead to decorrelation in the uncovered set.

For example, a torus in dimension $d \geq 3$ is strongly uniformly transient, but a torus of dimension $d = 1, 2$ is not even weakly uniformly transient. More generally, let us return to the example of the thin torus $\Gamma = (\mathbb{Z}/m\mathbb{Z})^2 \times (\mathbb{Z}/h\mathbb{Z})$ discussed above, where $h = h_m$. For which values of h is this weakly/strongly uniformly transient? Since a two-dimensional random walk returns to the origin roughly $\log m$ times by time $t = t_{\text{mix}}(1/4) \asymp m^2$, it is not hard to see that the thin torus is strongly uniformly transient if and only if $h_m \ggg \log m$ (and weakly uniformly transient if and only if $h_m \gtrsim \log m$). This condition coincides with (DC), as already observed.

This immediately raises the following question: could it be the case that the diameter condition is equivalent to strong (or weak) uniform transience? We will see as part of our analysis that any sequence of vertex-transitive graphs $\Gamma_n = (V_n, E_n)$ satisfying the diameter condition (DC) satisfies that

$$\lim_{r\to\infty} \limsup_{n\to\infty} \max_{y\in V_n : d(o,y)\geq r} \mathbb{E}_y(L_o(2\, t_{\text{mix}}(1/4))) = 0,$$

where $d(o, y)$ is the graph distance between y and o (see, e.g., Lemma 2.2.3). Strong uniform transience can be deduced from this relatively easily (in Proposition 2.2.7 we give a more direct argument). Thus

$$\text{(DC) implies (SUT).} \tag{2.1.12}$$

However the converse is, perhaps surprisingly again, not true. In Section 2.7 we will construct a sequence of finite vertex-transitive graphs Γ_n of uniformly bounded degree which satisfy SUT but not (DC). In particular, as a consequence of Theorem 2.1.2, despite being locally transient in this strong sense, the uncovered set will display nontrivial correlations. This example will be constructed by considering the product of a Ramanujan graph and a cycle of suitable sizes. (Essentially, in this product, one direction is very recurrent, but the other is very transient). See Section 2.7 for the precise definition.

The implication (2.1.12) is closely related to a conjecture of Benjamini and Kozma [BK05]. This conjecture states that for vertex-transitive graphs of bounded degree, if we assume $D^2 \lesssim n/\log n$ then the effective resistance $\mathcal{R}_{x\leftrightarrow y}$ between vertices of the graph is uniformly bounded above by some constant. This conjecture was recently solved by Tessera and Tointon [TT21]. Furthermore, using the tools developed in their paper, it is not hard to show that the effective resistance is uniformly bounded if and only if (WUT) holds. The implication (2.1.12) is therefore the direct "strong" analogue of the Benjamini–Kozma conjecture (the left hand side replaces the assumption $D^2 \lesssim n/\log n$ by $D^2 \lll n/\log n$, and the weak uniform transience in the right hand side by the strong uniform transience).

We end this discussion with an instructive comparison with the results in a recent paper of Dembo, Ding, Miller and Peres [DDMP19]. This describes the first order (cutoff) behaviour of the mixing time of the lamplighter walk on the thin torus $\Gamma_m = (\mathbb{Z}/m\mathbb{Z})^2 \times (\mathbb{Z}/h\mathbb{Z})$ with $h = a \log m$. The (total variation) mixing time of the lamplighter group on the base graph Γ_m is known to be closely related to the cover time of the simple random walk on that base graph Γ_m. This led the authors of [DDMP19] to use the ratio between the mixing time of the lamplighter on Γ_n and the cover time on Γ_n as an indicator of low- vs. high-dimensional behaviour. Indeed, this ratio is asymptotically equal to 1 for a completely flat, two-dimensional torus by results of [DPRZ04] and [PR04], whereas it is asymptotically $1/2$ for a three-dimensional torus, by results of Miller and Peres [MP12] (as well as on the complete graph, where the lamplighter walk reduces to the well known walk on the hypercube). Surprisingly, the authors in [DDMP19] show that on a thin torus, this ratio is strictly contained in the interval $(1/2, 1)$ for $0 < a < a_*$ and becomes asymptotically equal to $1/2$ (as in the three-dimensional case) for $a \geq a_*$, for some explicit a_*. They interpret this as a phase transition between low-dimensional ("recurrent") and high-dimensional ("transient") behaviour; see the discussion after Theorem 1.3 in that paper. By contrast, our results show that, at the level of fluctuations, high-dimensional behaviour only kicks in when $a = a_m \to \infty$, arbitrarily slowly, rather than for $a \geq a_*$.

2.1.4 Discussion of proof ideas and organisation of paper

Our starting point for this paper is the remarkable series of papers by Tessera and Tointon [TT21], [TT20] which gives a **quantitative form of Gromov's theorem** on groups of polynomial growth. Recall that, since the graph is vertex-transitive, by results of Trofimov [Tro84] it is roughly isometric to the Cayley graph of a finite group. Recall also that, for infinite groups, Gromov's theorem [Gro81] shows that if the volume of balls of radius r grow polynomially in the radius r then the growth exponent α must be integer and the group is then (in the Gromov–Hausdorff sense) close to a d-dimensional Euclidean lattice.

The diameter condition (DC), which says that the graph is slightly-more than two-dimensional, combined with the results of Tessera and Tointon, therefore implies that at least for relatively moderate distances the volume growth is at least **three-dimensional**. In combination with isoperimetric profile bounds (coming for instance from the theory of **evolving sets** of Morris and Peres [MP05]), this translates into very good decay for the heat kernel at small times and so gives excellent control on the number of returns by random walk to its starting point in this time. Later visits to this point are controlled using two-dimensional estimates. As we will see now, it turns out this is at the root of the decorrelation in the uncovered set.

These heat kernel bounds are used as follows. Suppose that (DC) holds and we wish to prove (2.1.2). Let us compute the kth cumulant of the size of the uncovered set Z_s at time $t_{\langle s\rangle} = \mathrm{t}_{\mathrm{hit}}(\log n + s)$. It suffices to show that this cumulatnt converges to the kth cumulant of a Poisson random variable with parameter e^{-s}. Now, this cumulant can be expressed as a sum over sets A of size k of the probability that A was not visited by time $t_{\langle s\rangle}$, and we can assume without loss of generality that the starting distribution is the uniform distribution π, namely $\mathbb{P}_\pi(T_A > t_{\langle s\rangle})$. The Aldous–Brown approximation [AB92] (recall also Theorem 2.1.4) shows that this hitting time is approximately an exponential random variable.

It remains to get good approximation for the quasistationary exponential rate. We show (using again the Aldous–Brown approximation) that this quasistationary exponential rate is close to the ratio $q_A := \mathbb{E}_\pi(T_o)/\mathbb{E}_\pi(T_A)$ of expectations of hitting time of A compared to that of a single point; see Proposition 2.2.9. For sets A such that the points in A are well separated, we typically expect $q_A \approx k$, because the hitting time of A is close to the minimum of k independent exponential random variables. The challenge is to quantify this approximation and show that the contribution coming from sets where the points are not so well separated is negligible. It is here that the heat kernel bounds are very useful: indeed, the on-diagonal decay of $p_t(x, y)$ translates into a **strong off-diagonal decay** (using a subgaussian estimate, namely a recent variant due to Folz [Fol11]) and implies good quantitative bounds of the desired form for q_A. In the most delicate case where n is barely larger than $D^2 \log n$, the analysis uses a somewhat elaborate induction over scales which requires strong quantitative bounds.

In the opposite direction, when (DC) fails, the key task (both for Theorems 2.1.1 and 2.1.2) is to show that the expected number of uncovered points at time $t_{\langle s\rangle}$ is strictly smaller (by a factor c bounded away from 1) than e^{-s}. Indeed this immediately implies Theorem 2.1.2, and implies Theorem 2.1.1 by considering the tail at $+\infty$ of τ_{cov} and a simple union bound.

The proof of Theorem 2.1.3 is more complicated. It might initially be tempting to show that the uncovered set is "too" clustered, i.e., the probability that two relatively nearby points are uncovered is larger than it should be in an independent scenario. However, this turns out to be very difficult to control (moments of order at least two of the size of the uncovered set can explode when (DC) does not hold, precisely because of the contribution coming from nearby points). Instead we show that the uncovered set is negatively correlated at large distances. This requires controlling the macroscopic variations of the Green function, a task which occupies a good part of Section 2.6.

Organisation of the paper. We start in Section 2.2 with the preliminaries in which we setup the notations and obtain the heat kernel bounds (Corollary 2.2.1 for the on-diagonal term, and Section 2.2.3 for the off-diagonal terms). We also discuss the original Aldous–Brown approximation and some elementary consequences in Section 2.2.4.

Section 2.3 is the proof of the main theorem (Theorem 2.1.1) under the assumption (DC), and discusses the cumulants of the uncovered set. A detailed strategy is provided in Section 2.3.1.

Section 2.4 contains the proofs of Theorem 2.1.2 under the assumption (DC) and its refinements on the structure of the uncovered set.

Section 2.5 brings new insights to understand quasistationary distributions and proves Theorem 2.1.5 improving the Aldous–Brown approximations. Section 2.5.3 provides a somewhat shorter (at least given the results of that section) proof of Theorem 2.1.1 under the condition (DC).

Section 2.6 studies the cover time and uncovered set of graphs which do not satisfy (DC) and completes the proof of the theorems 2.1.1, 2.1.2, and 2.1.3.

Section 2.7 contains the construction of an explicit example of graphs that are (SUT) but do not satisfy (DC).

Acknowledgements. We thank Roberto Imbuzeiro Oliveira and Matt Tointon for some useful discussions. In particular we thank Matt Tointon for asking us about the relation between Definition 2.1.6 and the definition in [TT20]. This work started when the first two authors were at the University of Cambridge, during which time they were supported by EPSRC grant EP/L018896/1. The paper was finished while the first author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2022 semester on *Analysis and Geometry of Random Spaces*, which was supported by the National Science Foundation under Grant No. DMS-1928930. J.H.'s research is supported by NSERC grants.

2.2 Preliminaries

2.2.1 A priori bounds on the heat kernel

From now on and for the rest of the article to ease notations, we will write with a small abuse of notation Γ for the graph and will denote the size of the vertex set of Γ by n. Our constants will implicitly be allowed to depend on the degree d (and also on the size $k \geq 1$ of a set when we apply the moment method to compute the size of the uncovered set). We will not always recall this. This also applies to notations such as $a_n = o(b_n)$, $a_n = O(b_n)$. We will also sometimes use $a_n \lll b_n$ (which is by definition equivalent to $a_n = o(b_n)$, i.e., $a_n/b_n \to 0$) and $a_n \lesssim b_n$ (which by definition means $a_n = O(b_n)$, i.e., a_n/b_n is bounded).

It will be at times convenient to allow the graph Γ to not be vertex-transitive. Let $P(\cdot,\cdot)$ be the transition kernel of our walk. Define the conductance $\Phi(S)$ of a (non-empty) set $S \subset \Gamma$ by

$$\Phi(S) = \frac{Q(S,S^c)}{\pi(S)}, \tag{2.2.1}$$

where for $A, B \subset \Gamma$,

$$Q(A,B) = \sum_{a\in A, b\in B} \pi(a)P(a,b). \tag{2.2.2}$$

We call conductance profile the function ϕ defined, for $\min_{x\in\Gamma} \pi(x) \leq u \leq 1/2$, by

$$\phi(u) = \inf\left\{\Phi(S) : \pi(S) \leq u\right\}. \tag{2.2.3}$$

We first recall here a crucial consequence of the theory of evolving sets which allows us to get bounds on the heat kernel given a conductance profile. Let $\varepsilon > 0$. Then we know by [MP05, Theorem 13], considering only diagonal transitions, that for ε such that $4/\varepsilon \leq 1/2$ and every $x \in \Gamma$,

$$t \geq \int_{4\pi(x)}^{4/\varepsilon} \frac{8du}{u\phi(u)^2} \implies p_t(x,x) \leq (1+\varepsilon)\pi(x). \tag{2.2.4}$$

Consider now the vertex-transitive case where $\pi(x) = 1/n$ for every n. When applying this inequality it is essential to note that the condition on ε is such that $8 \leq \varepsilon \leq n$. It may be slightly perverse to name ε a quantity which by assumption is greater or equal to 8, but these notations are by now relatively well established so we do not tamper with them. ($w = n/\varepsilon$ plays the role of the relevant volume in the graph, namely we get $p_t(o,o) \asymp 1/w$. Therefore ε will typically be small compared to the volume n.)

Our first task will be to transform the condition on the conductance profile into a condition on the volume growth, or equivalently, on isoperimetry. (Combined with the work of Tessera and

Tointon on a quantitative form of Gromov's theorem giving strong control on the volume growth, this will give us excellent control on the return probabilities.)

Proposition 2.2.1. Recall that $n = |\Gamma|$ and assume that $n \geq D^q$. Then, we have uniformly over all $t \leq D^2$, writing $m = \lfloor q \rfloor$ and $R = D^{\{q\}}$,

$$p_t(o,o) \lesssim \max\left(\frac{1}{t^{(m+1)/2}}, \frac{1}{Rt^{m/2}}\right), \tag{2.2.5}$$

where the implicit constant depends on m and d.

Remark 2.2.1. For a given $m = \lfloor q \rfloor$ this upper bound is sharp, as is easily shown by considering "flat" tori of the form $\Gamma = (\mathbb{Z}/L\mathbb{Z})^m \times (\mathbb{Z}/R\mathbb{Z})$ with $1 \lll R \leq L$. In this case, the heat kernel initially decays like in dimension $(m+1)$; but eventually (for $t \ggg R^2$) the decay becomes only m-dimensional.

Remark 2.2.2. We will only apply this result with $m \leq 5$, so we don't need to track the dependence in m.

Remark 2.2.3. By using [MP05, Theorem 1] instead of [MP05, Theorem 13], the same estimate holds for the discrete time transition probability $p_t^{\#}(o,o)$, a fact which will also be useful in some of the steps of the proof of Lemma 2.3.5.

Proof. The key is to observe that we have the following bound on the isoperimetric profile: for every nonempty $S \subset \Gamma$ such that $\pi(S) \leq 1/2$, then

$$\Phi(S) \gtrsim \min\left(|S|^{-1/(m+1)}, R^{1/m}|S|^{-1/m}\right). \tag{2.2.6}$$

This is essentially [TT20, Theorem 6.1], which is in itself a simple consequence of their bound on the volume growth [TT20, Proposition 5.1] and an isoperimetric inequality [TT20, Proposition 4.1] for vertex-transitive graphs. (In fact such an inequality can also be found in [LP16b, Lemma 10.46], and for the infinite case in [LMS08, Lemma 7.2].) For $v \leq n/2$, considering sets S such that $v/n \leq \pi(S) \leq 1/2$, we get immediately that

$$\phi(v/n) \gtrsim \min\left(v^{-1/(m+1)}, R^{1/m}v^{-1/m}\right). \tag{2.2.7}$$

We deduce, making the change of variables $v = nu$ and setting $w = 8n/\varepsilon$ with $8 \leq \varepsilon < n$,

$$\begin{aligned}\int_{4/n}^{4/\varepsilon} \frac{8du}{u\phi(u)^2} = \int_4^{4n/\varepsilon} \frac{8dv}{v\phi(v/n)^2} &\lesssim \int_4^{4n/\varepsilon} \max\left(v^{2/(m+1)}, \frac{v^{2/m}}{R^{2/m}}\right)\frac{dv}{v} \\ &\lesssim \max\left(w^{2/(m+1)}, \frac{w^{2/m}}{R^{2/m}}\right).\end{aligned}$$

Consequently, there exists $C \geq 1$ such that if $t = C\max\left(w^{2/(m+1)}, \frac{w^{2/m}}{R^{2/m}}\right)$, we may apply (2.2.4) and obtain (recall that $w = 8n/\varepsilon$),

$$p_t(o,o) \leq \frac{1+\varepsilon}{n} = \frac{1}{n} + \frac{8}{w} \leq \frac{9}{w}.$$

As $t = C\max\left(w^{2/(m+1)}, \frac{w^{2/m}}{R^{2/m}}\right)$, we have in particular $w \gtrsim \min\left(t^{(m+1)/2}, Rt^{m/2}\right)$, and the bound on $p_t(o,o)$ can be rewriten as

$$p_t(o,o) \lesssim \max\left(\frac{1}{t^{(m+1)/2}}, \frac{1}{Rt^{m/2}}\right). \tag{2.2.8}$$

Finally, as we defined first $1 < w \leq n$, and then $t = C \max\left(w^{2/(m+1)}, \frac{w^{2/m}}{R^{2/m}}\right)$, (2.2.8) holds for all t such that $C < t \leq CD^2$, and hence, as we took $C \geq 1$ and up to changing the implicit constant because of the values of t smaller than C, for all $t \leq D^2$, as desired. ☐

To avoid separating the proof into too many cases and cutting the integrals into several pieces, the following Corollary will be helpful. It provides upper bounds on the diagonal return probabilities which are true for all t.

Corollary 2.2.1. *Recall that $n = D^2 f(n)$.*

(a) Uniformly over all finite vertex-transitive graphs of degree less or equal to d such that $n \geq D^2$, and $t \geq 0$, we have

$$p_t(o,o) \lesssim \frac{1}{t^{3/2}} + \frac{1}{f(n)t}, \qquad 1 \leq t \leq D^2, \tag{2.2.9}$$

and

$$p_t(o,o) \lesssim \frac{1}{D^3} + \frac{1}{n}, \qquad t \geq D^2. \tag{2.2.10}$$

(b) Uniformly over all finite vertex-transitive graphs of degree less or equal to d such that $n \geq D^5$, and $t \geq 0$, we have

$$p_t(o,o) \lesssim \frac{1}{t^{5/2}} + \frac{1}{D^5}. \tag{2.2.11}$$

Moreover, the implicit constants depend only on d.

Remark 2.2.4. The bounds we give are not necessarily optimal if we fix the value of $m = \lfloor q \rfloor$, but have the advantage that they do not depend on m and so can be used regardless. This leads to fewer cases to treat separately further down the proof and so makes the argument more unified.

Proof. (a) First assume that $f(n) < D$. Setting q such that $n = D^q$, we have from Proposition 2.2.1, noting that in this case $R = f(n)$, that for $1 \leq t \leq D^2$,

$$p_t(o,o) \lesssim \frac{1}{t^{3/2}} + \frac{1}{f(n)t}. \tag{2.2.12}$$

On the other hand, if $f(n) \geq D$, i.e. $n \geq 3$, we have, applying Proposition 2.2.1 with $q = 3$ that for all $1 \leq t \leq D^2$, $p_t(o,o) \lesssim \frac{1}{t^{3/2}}$. Observing moreover that as $t \leq D^2$, $\frac{1}{f(n)t} \leq \frac{1}{t^{3/2}}$, we get that uniformly over all vertex-transitive graphs (of degree less or equal to d) such that $n \geq D^2$, and $t \leq D^2$, we have

$$p_t(o,o) \lesssim \frac{1}{t^{3/2}} + \frac{1}{f(n)t}. \tag{2.2.13}$$

For $t \geq D^2$, as $t \mapsto p_t(o,o)$ is decreasing, we have, using the previous bound at time D^2,

$$p_t(o,o) \leq p_{D^2}(o,o) \lesssim \frac{1}{D^3} + \frac{1}{n}. \tag{2.2.14}$$

This concludes the proof of (a). (b) This follows by Proposition 2.2.1 with $q = 5$, proceeding exactly as in the proof of (a). ☐

2.2.2 Relaxation, mixing, and hitting times

Recall that the relaxation time of the chain is the inverse of the spectral gap:

$$t_{\mathrm{rel}} = \frac{1}{1-\lambda_2}, \tag{2.2.15}$$

where λ_2 is the second largest eigenvalue of the chain. By a classical argument (see e.g. [Chu97, Theorem 7.6]) the following bound on the relaxation time is valid on every finite (connected) vertex-transitive graph:

$$t_{\mathrm{rel}} \leq dD^2. \tag{2.2.16}$$

We now show that for graphs of polynomial growth, $t_{\mathrm{rel}} \asymp D^2$.

Lemma 2.2.1. *There exists a constant C such that for every finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \leq D^5$,*

$$CD^2 \leq t_{\mathrm{rel}} \leq dD^2. \tag{2.2.17}$$

Proof. We follow arguments from the proof of the lower bound of [DSC93, Theorem 3.1].

By the minimax characterisation of the spectral gap (see for instance [LP17, Lemma 13.7 and Remark 13.8]) we have, setting $g = d(o, \cdot)$,

$$1 - \lambda_2 \leq \frac{\mathcal{E}(g)}{\mathrm{Var}_\pi(g)}. \tag{2.2.18}$$

Since g is 1-Lipschitz, $\mathcal{E}(g) \leq 1/2$. We hence have

$$t_{\mathrm{rel}} \geq 2\mathrm{Var}_\pi(g) = \frac{1}{n^2}\sum_{x,y\in\Gamma}(g(x)-g(y))^2. \tag{2.2.19}$$

Let $z \in \Gamma$ such that $d(o,z) = D$. We deduce that

$$t_{\mathrm{rel}} \geq \frac{1}{n^2}\sum_{x\in B(o,D/4), y\in B(z,D/4)} (D/2)^2 = \frac{D^2}{4}\left(\frac{V(D/4)}{n}\right)^2. \tag{2.2.20}$$

Moreover, by [TT20, Proposition 5.1], there exists a (universal) constant c such that for all finite (connected) vertex-transitive graphs satisfying $n \leq D^5$,

$$V(D/4) \geq cn. \tag{2.2.21}$$

For those graphs, we hence have, setting $C = c^2/4$

$$CD^2 \leq t_{\mathrm{rel}} \leq dD^2. \tag{2.2.22}$$

□

Let us now define mixing times and the distance to stationarity. We restrict here to our framework of simple random walks on vertex-transitive graphs, so the walk is in particular transitive and the stationary distribution is the uniform distribution π. See [LP17, Chapter 4] for more general definitions. For $1 \leq p < \infty$, we define the L^p distance to stationarity of the chain $(X_t)_{t\geq 0}$ at time t by

$$d^{(p)}(t) = \left(\frac{1}{n}\sum_{x\in\Gamma}|np_t(o,x)-1|^p\right)^{1/p}. \tag{2.2.23}$$

We also define the L^∞ distance to stationarity as (see [LP17, Proposition 4.15]).

$$d^{(\infty)}(t) = np_t(o,o) - 1. \tag{2.2.24}$$

We define the L^p mixing time at level ε, for $\varepsilon < 0$ and $1 \le p \le \infty$, by

$$t^{(\mathrm{p})}_{\mathrm{mix}}(\varepsilon) = \inf\left\{t \ge 0 : d^{(p)}(t) \le \varepsilon\right\}. \tag{2.2.25}$$

Note that the total variation distance is just half of the L^1 distance. For all $x, y \in \Gamma$ and $t \ge 0$, we set

$$h_t(x,y) = p_t(x,y) - \frac{1}{n}. \tag{2.2.26}$$

Proposition 2.2.2. For $0 < \varepsilon < 1$, we have, uniformly over all finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \le D^5$,

$$t^{(\infty)}_{\mathrm{mix}}(\varepsilon) \asymp_{d,\varepsilon} t^{(1)}_{\mathrm{mix}}(\varepsilon) \asymp_{d,\varepsilon} t_{\mathrm{rel}} \asymp_d D^2. \tag{2.2.27}$$

Proof. We already proved in Lemma 2.2.1 that $t_{\mathrm{rel}} \asymp_d D^2$. By convexity, if $1 \le p \le q \le \infty$, and $t \ge 0$, $d^{(p)}(t) \le d^{(q)}(t)$, and hence for $0 < \varepsilon < 1$, we have $t^{(\mathrm{p})}_{\mathrm{mix}}(\varepsilon) \le t^{(\mathrm{q})}_{\mathrm{mix}}(\varepsilon)$. By [LP17, Lemma 20.11], we hence have $d^{(\infty)}(t) \ge d^{(1)}(t) \ge e^{-t/\,t_{\mathrm{rel}}}$ (since the total variation distance is exactly the half of the L^1 distance) and thus for $0 < \varepsilon < 1$,

$$t^{(\infty)}_{\mathrm{mix}}(\varepsilon) \ge t^{(1)}_{\mathrm{mix}}(\varepsilon) \ge \log(1/\varepsilon)\, t_{\mathrm{rel}}\,. \tag{2.2.28}$$

Let us now prove that $t^{(\infty)}_{\mathrm{mix}}(\varepsilon) \lesssim t_{\mathrm{rel}}$. By spectral estimates, we have for all $t, s \ge 0$.

$$h_{t+s}(o,o) \le e^{-s/\,t_{\mathrm{rel}}} h_t(o,o). \tag{2.2.29}$$

Moreover, by Proposition 2.2.1, there is a constant $C(d)$ such that uniformly over all finite connected vertex-transitive graphs of degree d such that $n \le D^5$, $h_{D^2}(o,o) \le C(d)$. □

Hence, at time $D^2 + i\, t_{\mathrm{rel}}$, where $i = \lceil \log(C(d)/\varepsilon) \rceil$,

$$h_{D^2 + i\, t_{\mathrm{rel}}}(o,o) \le e^{-i} h_{D^2}(o,o) \le \varepsilon, \tag{2.2.30}$$

so, recalling that $t_{\mathrm{rel}} \le dD^2$,

$$t^{(\infty)}_{\mathrm{mix}}(\varepsilon) \le D^2 + \lceil \log(C(d)/\varepsilon) \rceil\, t_{\mathrm{rel}} \le D^2 \left(1 + d \lceil \log(C(d)/\varepsilon) \rceil\right), \tag{2.2.31}$$

concluding the proof.

Remark 2.2.5. Proposition 2.2.2 shows that when $n \le D^5$, the relaxation time and the mixing time are both of order D^2. The explicit bounds (2.2.28) and (2.2.31) show in particular that there is no *cutoff*, i.e. that for some $0 < \varepsilon < 1$, $\frac{t^{(\infty)}_{\mathrm{mix}}(1-\varepsilon)}{t^{(\infty)}_{\mathrm{mix}}(\varepsilon)}$ does not converge to 1 as $n = |\Gamma| \to \infty$. See [LP17, Chapter 18] more details on the cutoff phenomenon.

The following proposition is a classical result which can be traced back to [Ald82].

Proposition 2.2.3. Uniformly over all finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \le D^5$, we have

$$\mathrm{t}_{\mathrm{hit}} - t_{\langle \mathrm{hit} \rangle} \asymp_d D^2. \tag{2.2.32}$$

Proof. Let τ_i for $1 \le i \le 4$ be as in [Ald82], and set $\tau_5 := \mathrm{t_{hit}} - t_{\langle \mathrm{hit} \rangle}$. From [Ald82, Equation (17)], we have $\tau_5 \le \tau_2$. On the other hand, we have

$$\tau_3 = \max_{x,y\in\Gamma} \sum_{z\in\Gamma} \pi(z) \left|\mathbb{E}_x T_z - \mathbb{E}_y T_z\right| = 2 \max_{x,y\in\Gamma} \sum_{z\in\Gamma \,:\, \mathbb{E}_x T_z \ge \mathbb{E}_y T_z} \pi(z) \left(\mathbb{E}_x T_z - \mathbb{E}_y T_z\right). \tag{2.2.33}$$

Moreover, for any $x, y, z \in \Gamma$, we have $\mathbb{E}_x T_z - \mathbb{E}_y T_z \le \mathrm{t_{hit}} - \mathbb{E}_y T_z$, and (by definition of $\mathrm{t_{hit}}$) $\mathrm{t_{hit}} - \mathbb{E}_y T_z \ge 0$. We deduce from this that

$$\tau_3 \le 2 \max_{x,y\in\Gamma} \sum_{z\in\Gamma} \pi(z)(\mathrm{t_{hit}} - \mathbb{E}_y T_z) = 2\tau_5, \tag{2.2.34}$$

so we have proved that $\tau_3/2 \le \tau_5 \le \tau_2$. It finally follows from [Ald82, Theorem 5] and Proposition 2.2.2 that

$$\tau_5 \asymp \tau_1 = \tfrac{1}{2}\, t^{(1)}_{\mathrm{mix}}\left((2e)^{-1}\right) \asymp_d D^2. \tag{2.2.35}$$

□

Remark 2.2.6. Proposition 2.2.3 also holds if we replace $t_{\langle \mathrm{hit} \rangle} = \mathbb{E}_\pi T_o$ by $\mathbb{E}_{\alpha_o} T_o$, where α_o is the quasi-stationary distribution (defined in (2.1.4)) associated to $A = \{o\}$. This will be shown in Proposition 2.6.8.

Corollary 2.2.2. *Under the assumptions of Theorem 2.1.2, we have when* (DC) *holds,*

$$t_{hit}((\log n) + s) = t_{\langle \mathrm{hit} \rangle}((\log n) + s + o(1)). \tag{2.2.36}$$

Proof. From Proposition 2.2.3, we have

$$\mathrm{t_{hit}} = t_{\langle \mathrm{hit} \rangle}\left(1 + O\left(\frac{D^2}{\mathrm{t_{hit}}}\right)\right). \tag{2.2.37}$$

Moreover, by [LP17, Proposition 1.19],

$$\mathrm{t_{hit}} = \max_{x,y} \mathbb{E}_x T_y \ge \max_x \mathbb{E}_x T_x^+ - 1 = \mathbb{E}_o T_o^+ - 1 = n - 1. \tag{2.2.38}$$

We finally deduce (using that (DC) holds for the last equality), that

$$\mathrm{t_{hit}}((\log n) + s) = t_{\langle \mathrm{hit} \rangle}\left((\log n) + s + O\left(\frac{D^2 \log n}{n}\right)\right) = t_{\langle \mathrm{hit} \rangle}\left((\log n) + s + o\left(1\right)\right). \tag{2.2.39}$$

□

Proposition 2.2.4. Recall that for $s \in \mathbb{R}$, $t_{\langle s \rangle} = t_{\langle \mathrm{hit} \rangle}((\log n) + s)$, and t^*_s is such that $\mathbb{E}|U(t^*_s)| = e^{-s}$. Then under (DC), we have

$$t^*_s = t_{\langle s \rangle} + o(n). \tag{2.2.40}$$

Proof. First observe that we have for every $t \ge 0$, setting $\alpha = \alpha_o$ the quasi-stationary distribution associated to $A = \{o\}$,

$$\mathbb{E}|U(t)| = n\mathbb{P}_\pi(T_o > t) = n\frac{\mathbb{P}_\pi(T_o > t)}{\mathbb{P}_\alpha(T_o > t)}\mathbb{P}_\alpha(T_o > t) = n\frac{\mathbb{P}_\pi(T_o > t)}{\mathbb{P}_\alpha(T_o > t)}e^{-t/\mathbb{E}_\alpha T_o}. \tag{2.2.41}$$

It follows, since $t \mapsto \mathbb{E}|U(t)|$ is decreasing, that t^*_s is unique and satisfies

$$t^*_s = \mathbb{E}_\alpha T_o\left((\log n) + s - \log\left(\frac{\mathbb{P}_\alpha(T_o > t^*_s)}{\mathbb{P}_\pi(T_o > t^*_s)}\right)\right). \tag{2.2.42}$$

Since (DC) holds, $t_{\text{rel}} / \mathrm{t}_{\text{hit}} \lesssim D^2/n = o(1/\log n)$ (by (2.2.16) and (2.2.38)). We deduce from the Aldous–Brown approximation (Theorem 2.1.4) that $\mathbb{E}_\alpha T_o = \mathbb{E}_\pi T_o(1 + o(1/\log n))$, and (using also that $\log(1+x) \leq x$ for $x > -1$) that $\log\left(\frac{\mathbb{P}_\alpha(T_o > t_s^*)}{\mathbb{P}_\pi(T_o > t_s^*)}\right) = o(1/\log n) = o(1)$. This implies

$$t_s^* = t_{\langle s\rangle} + o(t_{\langle \text{hit}\rangle}).$$

To conclude it suffices to show that $\mathrm{t}_{\text{hit}} = O(n)$. To see this, note that by the using reversibility and the commute-time identity (see [LP17, Proposition 10.7]), we have for all $x, y \in \Gamma$

$$\mathbb{E}_x(T_y) + \mathbb{E}_y(T_x) = d\mathcal{R}(x \leftrightarrow y)n,$$

so it suffices to show that the effective resistance between vertices is uniformly bounded. However, writing $\mathcal{R}_{\Gamma,2} := \max_{x,y\in\Gamma} \mathcal{R}(x \leftrightarrow y)$, we have from [TT20, Theorem 1.12], that

$$\mathcal{R}_{\Gamma,2} \lesssim \frac{1}{d} + \frac{D^2 \log n}{n} = \frac{1}{d} + o(1) = O(1). \tag{2.2.43}$$

This concludes the proof. □

A crucial part of the argument will be to obtain good bounds on the hitting time of a finite arbitrary set A of fixed cardinality $k \geq 1$. Problems usually arise when some points in A are relatively close to one another. The following proposition shows that under (DC), the expected hitting time of A is always of order n.

Proposition 2.2.5. Let A be a finite set of vertices of cardinality k of Γ. Then, as $n \to \infty$,

$$\mathbb{E}_\pi T_A \gtrsim n, \tag{2.2.44}$$

and besides assuming (DC) we have

$$\mathbb{E}_\pi T_A \asymp n, \tag{2.2.45}$$

where the implicit constants depend on k and d.

Proof. Averaging (2.2.49) with respect to x, and taking $y = o$, we have for every $t \geq 0$

$$\mathbb{P}_\pi(T_o \leq t) \leq e\mathbb{E}_\pi L_o(t+1) = e(t+1)/n. \tag{2.2.46}$$

It follows that for every $t \geq 1$, as $t + 1 \leq 2t$, and taking $t = n/(4ek)$,

$$\mathbb{P}_\pi(T_A \leq t) \leq k\mathbb{P}_\pi(T_o \leq t) \leq 2ekt/n \leq 1/2.$$

By Markov's inequality, we deduce

$$\mathbb{E}_\pi T_A \geq t\mathbb{P}_\pi(T_A > t) = t\left(1 - \mathbb{P}_\pi(T_A \leq t)\right) \geq t/2 \gtrsim n,$$

which concludes the proof of the lower bound.

The upper bound is straightforward, since $\mathbb{E}_\pi T_A \leq \mathbb{E}_\pi T_o \leq \mathrm{t}_{\text{hit}}$, and we have already observed in the proof of Proposition 2.2.4 that under (DC), $\mathrm{t}_{\text{hit}} = O(n)$.

□

2.2.3 Off-diagonal heat kernel bounds

Our first task will be to translate the on-diagonal bounds described in the previous section into off-diagonal bounds. A general upper bound can always be obtained with the Carne–Varopoulos inequality, but we will need a sharper recent result due Folz, see [Fol11, Corollary 1.2]. This general result takes as an input a time-dependent bound on the on-diagonal heat kernel and deduces from this a general off-diagonal bound. The result is actually quite general but we will only use it in the most simple case where the Lipschitz function is taken to be the graph distance, and the "volume" is measured with respect to cardinality.

Set $1/g(t)$ to be the function on the right hand side of (2.2.13), so that $1/g(t)$ is a bound on $p_t(o,o)$; that is,

$$1/g(t) = \begin{cases} \frac{C}{t^{3/2}} & \text{for } 1 \leq t \leq f(n)^2 \\ \frac{C}{tf(n)} & \text{for } f(n)^2 \leq t \leq D^2. \end{cases} \tag{2.2.47}$$

(Recall that if $m = 2$ then $R = f(n)$.) A consequence of his result is the following inequality. Assume that $f(n) < D$. Then there exists a constant $c > 0$ such that for every $x \neq y \in \Gamma$ and $d(x,y) \leq t \leq D^2$,

$$p_t(x,y) \lesssim \frac{1}{g(t)} \exp\left(-c\frac{d(x,y)^2}{t}\right), \tag{2.2.48}$$

where the implicit constant in $\lesssim$ depends only on the degree bound. For smaller values of t, we will simply bound the heat kernel via the Carne–Varopoulos bound,

$$p_t(x,y) \lesssim \exp(-cd(x,y)^2/t) \leq \exp(-cd(x,y))$$

when $t \leq d(x,y)$.

We will use (2.2.48) to get upper bounds on the off-diagonal heat kernel, and therefore by integrating, on the expected local time at a given vertex. In turn, this can be used to upper bound the probability to visit a vertex y far away from x in a relatively short time, via the following elementary lemma.

Lemma 2.2.2.

$$\mathbb{E}_x L_y(t+1) \geq \frac{1}{e}\mathbb{P}_x\left(T_y \leq t\right). \tag{2.2.49}$$

Proof. Let us define the event E by

$$E := \{\text{the walk stays for time at least 1 at } y \text{ just after } T_y\}. \tag{2.2.50}$$

Then we have

$$\begin{aligned} \mathbb{E}_x L_y(t+1) &\geq \mathbb{E}_x\left(L_y(t+1) | T_y \leq t, E\right) \mathbb{P}_x\left(T_y \leq t, E\right) \\ &\geq \mathbb{P}_x\left(T_y \leq t\right) \mathbb{P}\left(E\right). \end{aligned}$$

This proves the lemma. □

We now combine the above ideas to get the following bounds:

Proposition 2.2.6. Let $1 \leq \delta \leq D/2$. (a) Uniformly over all $x, y \in \Gamma$ such that $d(x,y) \geq \delta$, and $t \geq 1$,

$$\mathbb{E}_x L_y(t) \lesssim \frac{1}{\delta} + \frac{\log(D/\delta)}{f(n)} + \frac{t}{n} + \frac{t}{D^3}. \tag{2.2.51}$$

(b) If we moreover assume that $D^5 \leq n$, the bound becomes

$$\mathbb{E}_x L_y(t) \lesssim \frac{1}{\delta^3} + \frac{t}{D^5}. \tag{2.2.52}$$

Note that by Lemma 2.2.2, the same bounds hold also for $\mathbb{P}_x(T_y \leq t)$.

Proof. Let $x, y \in \Gamma$. Since for all s, $p_s(x,y) \leq p_s(o,o)$ (see [AF, Lemma 20, in particular Equation (3.60)])

$$\mathbb{E}_x L_y(t) = \int_0^t p_s(x,y)\mathrm{d}s \leq \int_0^{d(x,y)} p_s(x,y)\mathrm{d}s + \int_{d(x,y)}^{d(x,y)^2} p_s(x,y)\mathrm{d}s + 1_{t \geq d(x,y)^2} \int_{d(x,y)^2}^t p_s(o,o)\mathrm{d}s. \tag{2.2.53}$$

The first integral is by the Carne–Varopoulos bound smaller or equal to $\delta \exp(-c\delta) \lesssim 1/\delta$. Let us consider the second integral. By (2.2.48), we have

$$\int_{d(x,y)}^{d(x,y)^2} p_s(x,y)\mathrm{d}s \lesssim \int_0^{d(x,y)^2} \left(s^{-3/2} + \frac{1}{f(n)s} \right) \exp(-c\delta^2/s)\mathrm{d}s.$$

Now, studying the function $t \mapsto t^{-b}\exp(-A/t)$ we see that this is maximised at $t = A/b$ and so is always $< (A/b)^{-b}$. Applying this with $b = 1$ and $b = 3/2$ as well as $A \asymp \delta^2$, we get

$$\int_0^{d(x,y)^2} p_s(x,y)\mathrm{d}s \lesssim \delta^2 \frac{1}{\delta^3} + \delta^2 \frac{1}{\delta^2 f(n)} \leq \frac{1}{\delta} + \frac{1}{f(n)}.$$

For the third integral, using Corollary 2.2.1, (a), we immediately get

$$\begin{aligned}
\int_{d(x,y)^2}^t p_s(o,o)\mathrm{d}s &\lesssim \int_{\delta^2}^{D^2} \left(\frac{1}{s^{3/2}} + \frac{1}{f(n)s} \right) \mathrm{d}s + 1_{t \geq D^2} \int_{D^2}^t \left(\frac{1}{D^3} + \frac{1}{n} \right) \mathrm{d}s \\
&\lesssim \frac{1}{\delta} + \frac{\log(D/\delta)}{f(n)} + \frac{t}{n} + \frac{t}{D^3}.
\end{aligned}$$

We finally have, as $1 \leq \delta \leq D/2$,

$$\mathbb{E}_x L_y(t) \lesssim \frac{1}{\delta} + \frac{\log(D/\delta)}{f(n)} + \frac{t}{n} + \frac{t}{D^3} + \frac{1}{f(n)} \lesssim \frac{1}{\delta} + \frac{\log(D/\delta)}{f(n)} + \frac{t}{n} + \frac{t}{D^3}, \tag{2.2.54}$$

which proves (a). The second bound is proved the same way, using part (b) of Corollary 2.2.1 and taking $1/g(t) = C/t^{5/2}$ for $1 \leq t \leq D^2$. ☐

Corollary 2.2.3. *Let $t = t(n) = o\left(\min(n, D^3)\right)$ and fix a sequence $\omega_n \to \infty$. Then, uniformly over x, y with $d(x,y) \geq \omega_n$,*

$$\mathbb{E}_x L_y(t) \to 0. \tag{2.2.55}$$

In particular,

$$\mathbb{P}_x(T_y \leq t) \to 0. \tag{2.2.56}$$

Proof. It follows immediately from part (a) of Proposition 2.2.6, Lemma 2.2.2, and the assumption $\log n \lll f(n)$. ☐

When we consider the diagonal case we get a similar bound, but this time of order 1; in this case we can therefore afford to consider times that are as big as the volume. Such a bound is also a signature of our local transience condition. As mentioned in the introduction, Tessera and Tointon [TT20] proved (2.2.57). We present its proof for the sake of completeness.

Proposition 2.2.7. Let $\Gamma_n = (|V_n|, E_n)$ be a sequence of finite vertex-transitive graphs of uniformly bounded degree, and diameters $D = D_n$ satisfying $D^2 \asymp |V_n|/f(n) = n/f(n)$. If $f(n) \gtrsim \log n$ then

$$\mathbb{E}_o L_o(n) = O(1). \tag{2.2.57}$$

If $f(n) \ggg \log n$ then

$$\lim_{s\to\infty} \limsup_{n\to\infty} \max_{o\in V_n} \mathbb{E}_o(L_o(t_{\text{mix}}(1/4)) - L_o(s)) = 0. \tag{2.2.58}$$

Proof. We first prove (2.2.57). We start as in the proof of Proposition 2.2.6, cutting the integral at time 1 instead of time $d(x,y)^2$. At time $t = D^2$, we have

$$\mathbb{E}_o L_o(D^2) = \int_0^1 p_s(o,o)\mathrm{d}s + \int_1^{D^2} p_s(o,o)\mathrm{d}s \lesssim 1 + \left(1 + \frac{\log D}{f(n)} + \frac{t}{n} + \frac{t}{D^3}\right) = O(1). \tag{2.2.59}$$

Recall from (2.2.16) that $t_{\text{rel}} \leq dD^2$. Furthermore, we have, for every $t \geq D^2$, (see for instance [LP17, Lemma 4.18 and Equation 20.18])

$$p_t(o,o) - \frac{1}{n} \leq \exp\left(-\frac{t-D^2}{dD^2}\right) p_{D^2}(o,o). \tag{2.2.60}$$

Hence,

$$\begin{aligned}\int_{D^2}^n p_s(o,o)\mathrm{d}s &= \int_{D^2}^n \left(p_s(o,o) - \frac{1}{n}\right)\mathrm{d}s + \int_{D^2}^n \frac{1}{n}\mathrm{d}s \\ &\lesssim \frac{1}{n}\int_{D^2}^\infty \exp\left(\frac{-t}{dD^2}\right)\mathrm{d}s + 1 \\ &= O(1),\end{aligned}$$

which concludes the proof. We now prove (2.2.58). By Corollary 2.2.1 we have that $t_{\text{mix}}(1/4) \lll n$. The remainder of the proof is analogous to that of (2.2.57) and hence omitted. $\square$

With this, it is easy to deduce that any point (distinct from the starting point) can be avoided with positive probability up to a time slightly smaller than the volume.

Proposition 2.2.8. There exists a constant $\beta > 0$ such that for $t = t(n) \lll n$ and uniformly over x, y distinct points of Γ, we have

$$\mathbb{P}_x(T_y > t) \geq \beta + o(1). \tag{2.2.61}$$

Proof. This can be proved using [TT20, Theorem 1.8], but we give a simple, self-contained argument. We argue by contradiction and assume (without loss of generality, perhaps taking a subsequence if needed) that there exists x, y such that $\mathbb{P}_x(T_y \leq t) = 1 - o(1)$. As the law of T_x starting from y is the same as the law of T_y starting from x by symmetry (see the proof of [Ald89, Proposition 2]), this implies, using the Markov property, that we also have

$$\mathbb{P}_x(T_x^+ \leq 2t) \geq \mathbb{P}_x(T_y \leq t)\mathbb{P}_y(T_x \leq t) = 1 - o(1). \tag{2.2.62}$$

By the same argument, for every integer $m \geq 1$,

$$\mathbb{P}_x(\text{the walk returns to } x \text{ at least } m \text{ times before time } 2mt) \geq \mathbb{P}_x(T_x^+ \leq 2t)^m = 1 - o(1). \tag{2.2.63}$$

Since $t = t(n) = o(n)$, we also have $2mt = o(n)$ and hence we deduce that $\mathbb{E}_o(L_o(n)) \geq m(1-o(1))$ and so is unbounded. This contradicts Proposition 2.2.7. $\square$

2.2.4 Exponential approximation of the hitting time

The aim of this subsection is to approximate the hitting time of a fixed set A of cardinality $k \geq 1$ (starting from stationarity) by an exponential random variable with mean $\mathbb{E}_\pi T_A$. The key tool to do this will be to consider the quasi-stationary distribution α_A, for which the hitting distribution is *exactly* exponential. Such an idea goes back to the work of Aldous and Brown [AB92].

If $\varnothing \neq A \subsetneq \Gamma$, we will denote by $\alpha = \alpha_A$ its quasi-stationary distribution. This is a distribution whose support is contained in $B := A^c$, and is unique and of support B when $\mathbb{P}_a[T_b < T_A]$ for all $a, b \in B$. In fact, it is in principle possible for the set A to disconnect $\Gamma \setminus A$ into several connected components, in which case there are multiple quasi-stationary distribution. However, under our diameter assumption (in fact, as soon as $n \geq D^2$ say), only one of these components may be of macroscopic size as $n \to \infty$ while $k \geq 1$ is fixed (for a fixed k, under our diameter condition, for all sufficiently large n the rest of the components are of bounded sizes, where the bound on their sizes depends only on k and the degree). In such a case we therefore consider α_A to be the quasi-stationary distribution associated with that component, and we do so without further commenting on this case. We start with an elementary lemma to bound the error term in the Aldous–Brown approximation.

Lemma 2.2.3. *Uniformly over all sets $A \subset \Gamma$ of fixed size $k \geq 1$, we have*

$$\frac{t_{\text{rel}}}{\mathbb{E}_\alpha T_A} \lesssim \frac{D^2}{n}, \tag{2.2.64}$$

where $\alpha = \alpha_A$ denotes the quasistationary distribution associated to A.

Proof. From (2.2.16), we have $t_{\text{rel}} \lesssim D^2$. Moreover, from Theorem 2.1.4, $\mathbb{E}_\alpha T_A \geq \mathbb{E}_\pi T_A$. Finally, from Proposition 2.2.5, $\mathbb{E}_\pi T_A \gtrsim n$. This concludes the proof. □

The following lemma shows that the quantities $\mathbb{E}_\pi T_A$ and $\mathbb{E}_\alpha T_A$ are very similar.

Lemma 2.2.4. *Let $k \in \mathbb{N}^*$, and A a finite subset of Γ of cardinality k. Then*

$$\mathbb{E}_\pi T_A = \mathbb{E}_\alpha T_A \left(1 + O\left(\frac{1}{f(n)}\right)\right). \tag{2.2.65}$$

Proof. This follows immediately from the Aldous–Brown approximation (Theorem 2.1.4) and Lemma 2.2.3.

□

We now observe that $\mathbb{P}_\pi(T_A > t)$ and $\mathbb{P}_\alpha(T_A > t)$ are also very similar.

Lemma 2.2.5. *Let $k \in \mathbb{N}^*$ and A be a finite subset of Γ of cardinality k. Then for every $t \in \mathbb{R}^+$,*

$$\mathbb{P}_\alpha \left(T_A > t\right) = \mathbb{P}_\pi \left(T_A > t\right) \left(1 + O\left(\frac{1}{f(n)}\right)\right), \tag{2.2.66}$$

where the implicit constants do not depend on t, nor on A (among sets of same cardinality k).

Proof. This follows again immediately from the Aldous–Brown approximation (Theorem 2.1.4) and Lemma 2.2.3.

□

Remark 2.2.7. The most precise approximation stated in [AB92] (as Equation (1) and Theorem 3) is the following. For every $t \in \mathbb{R}^+$,

$$\mathbb{P}_\alpha \left(T_A > t\right) \left(1 - \frac{t_{\text{rel}}}{\mathbb{E}_\alpha T_A}\right) \leq \mathbb{P}_\pi \left(T_A > t\right) \leq \mathbb{P}_\alpha \left(T_A > t\right) (1 - \pi(A)). \tag{2.2.67}$$

Combining those two lemmas leads to an exponential approximation of $\mathbb{P}_\pi(T_A > t_{\langle s\rangle})$, which will later be useful to rewrite the moments of the uncovered set at time $t_{\langle s\rangle}$, which we recall from the introduction is given by, for $s \in \mathbb{R}$ fixed,

$$t_{\langle s\rangle} = \mathbb{E}_\pi T_o(\log(n) + s).$$

Proposition 2.2.9. Let A be a finite subset of Γ of cardinality $k \geq 1$. Then

$$\mathbb{P}_\pi\left(T_A > t_{\langle s\rangle}\right) = \exp\left(-\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A}(\log(n) + s)\right)(1 + o(1)), \tag{2.2.68}$$

where the $o(1)$ is uniform over all the A of cardinality k.

Proof. From Proposition 2.2.5,

$$\mathbb{P}_\pi\left(T_A > t_{\langle s\rangle}\right) = \mathbb{P}_\alpha\left(T_A > t_{\langle s\rangle}\right)(1 + o(1)). \tag{2.2.69}$$

Now, by the definition of quasi-stationarity and Proposition 2.2.4,

$$\mathbb{P}_\alpha\left(T_A > t_{\langle s\rangle}\right) = \exp\left(-\frac{t_{\langle s\rangle}}{\mathbb{E}_\alpha T_A}\right) = \exp\left(-\frac{t_{\langle s\rangle}}{\mathbb{E}_\pi T_A}\left(1 + O\left(\frac{1}{f(n)}\right)\right)\right). \tag{2.2.70}$$

Noting that $\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} = \Theta(1)$ and recalling that $t_{\langle s\rangle} = \mathbb{E}_\pi T_o(\log(n) + s)$, this leads to

$$\mathbb{P}_\alpha\left(T_A > t_{\langle s\rangle}\right) = \exp\left(-\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A}(\log(n) + s) + O\left(\frac{\log n}{f(n)}\right)\right). \tag{2.2.71}$$

As $f(n) \ggg \log n$, this concludes the proof. □

Since exponential functions are approximately linear for small times, we obtain a relation between $\mathbb{P}_\pi\left(T_A > t\right)$ and $\mathbb{E}_\pi T_A$ for not too large values of t.

Proposition 2.2.10. Let $k \in \mathbb{N}^*$, let A be a finite subset of vertices of cardinality k, and $t \lll n$. Then

$$\mathbb{E}_\pi T_A = \frac{t}{\mathbb{P}_\pi\left(T_A \leq t\right)}\left(1 + O\left(\frac{n}{tf(n)}\right) + O\left(\frac{t}{n}\right)\right). \tag{2.2.72}$$

Moreover, the $O(\cdot)$ are uniform over such A and t.

Proof. Let $t \geq 0$. From Lemma 2.2.5, which is true uniformly for all $t \geq 0$,

$$\mathbb{P}_\pi\left(T_A > t\right) = \exp\left(-\frac{t}{\mathbb{E}_\alpha T_A}\right)\left(1 + O\left(\frac{1}{f(n)}\right)\right).$$

Suppose that $t \lll n$, so that $t/\mathbb{E}_\alpha(T_A) = o(1)$. Then by Taylor expansion of the exponential,

$$\begin{aligned}
\mathbb{P}_\pi(T_A > t) &= \left(1 - \frac{t}{\mathbb{E}_\alpha(T_A)} + O\left(\frac{t}{\mathbb{E}_\alpha(T_A)}\right)^2\right)\left(1 + O(\frac{1}{f(n)})\right) \\
&= 1 - \frac{t}{\mathbb{E}_\alpha(T_A)} + O\left(\frac{t}{\mathbb{E}_\alpha(T_A)}\right)^2 + O\left(\frac{1}{f(n)}\right).
\end{aligned}$$

Hence, recalling that $\mathbb{E}_\alpha(T_A) \asymp n$, we get

$$\mathbb{P}_\pi(T_A \leq t) = \frac{t}{\mathbb{E}_\alpha(T_A)}\left(1 + O\left(\frac{t}{n}\right) + O\left(\frac{n}{tf(n)}\right)\right). \tag{2.2.73}$$

□

The following corollary will be particularly useful in what follows.

Corollary 2.2.4. *Uniformly over $A \subset \Gamma$ of size k and $t \lll n$, we have*

$$\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} = \frac{\mathbb{P}_\pi (T_A \le t)}{\mathbb{P}_\pi (T_o \le t)} \left(1 + O\left(\frac{n}{tf(n)}\right) + O\left(\frac{t}{n}\right)\right). \tag{2.2.74}$$

Remark 2.2.8. The ratio $q_A := \frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A}$ will play a crucial role in our analysis. In what follows we will want to apply Corollary 2.2.4 with $t = D^2\sqrt{f(n)} = n/\sqrt{f(n)}$. Hence for this choice of t we obtain:

$$\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} = \frac{\mathbb{P}_\pi (T_A \le t)}{\mathbb{P}_\pi (T_o \le t)} \left(1 + O\left(\frac{1}{\sqrt{f(n)}}\right)\right). \tag{2.2.75}$$

As we will see, this approximation is very useful, as the hitting probabilities $\mathbb{P}_\pi (T_A \le t)$ are easier to estimate than the expectations $\mathbb{E}_\pi T_A$. This approximation is particularly good when $n \ggg D^2(\log n)^2$, i.e. $f(n) \ggg (\log n)^2$, as in this case error term will be small:

$$\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} = \frac{\mathbb{P}_\pi (T_A \le t)}{\mathbb{P}_\pi (T_o \le t)} \left(1 + o\left(\frac{1}{\log n}\right)\right). \tag{2.2.76}$$

As we will soon see (see (2.3.7)), an error term of $o(1/\log n)$ is indeed sufficient for our purpose. However, when we merely have $f(n) \ggg \log n$, this approximation is not precise enough anymore, and more arguments will be required. We will show two related but distinct ways to deal with this difficulty in the general case. The first approach will be to work directly with the expectations $\mathbb{E}_\pi T_A$, and through a bootstrap argument. It is technical but rather elementary. This will be done in Section 2.3.7. The second will be to improve on the results of Aldous and Brown [AB92], so that the approximation (2.2.76) still holds, even when we merely have $f(n) \ggg \log n$. This simplifies some aspects of the proof (essentially, the base case of bootstrap argument is simpler, although the bootstrap itself remains needed). This second approach will be carried out in Section 2.5.

2.3 Convergence of the uncovered set

We now start the proof of the main result of this paper, and so recall our standing assumptions where we fix $s \in \mathbb{R}$, and assume that $f(n) \ggg \log n$. We also recall that $Z = Z_s$ denotes the size of the uncovered set at time $t_{\langle s \rangle}$, and we want to prove that for any fixed $k \ge 1$,

$$\mathbb{E}_\pi \left[Z^{\downarrow k}\right] \to e^{-ks}, \tag{2.3.1}$$

as $n \to \infty$, where we recall that $z^{\downarrow k} = z(z-1)\cdots(z-k+1)$. Indeed, once the convergence (2.3.1) is known, a standard application of the method of moments shows that Z_s converges in distribution to a Poisson random variable with parameter e^{-s}: that is,

$$\mathbb{P}(Z_s = k) \to \exp(-e^{-s})\frac{e^{-ks}}{k!}. \tag{2.3.2}$$

Consequently,

$$\mathbb{P}(\tau_{\text{cov}} \le t_{\langle s \rangle}) = \mathbb{P}(Z_s = 0) \to \exp(-e^{-s}),$$

as desired.

Our first task will be to rewrite these factorial moments in terms of the ratio

$$q_A = \frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} \tag{2.3.3}$$

mentioned in the previous section, and which will play a crucial role in all that follows.

2.3.1 Strategy

We recall that $t_{\langle s\rangle} = \mathbb{E}_\pi T_o(\log(n)+s)$. As $Z = \sum_{x\in\Gamma} \mathbb{1}_{T_x > t_{\langle s\rangle}}$, we can see that

$$Z^{\downarrow k} = \sum_{A\in\mathcal{A}} \mathbb{1}_{T_A > t_{\langle s\rangle}}, \tag{2.3.4}$$

where $\mathcal{A}$ denotes the set of all (ordered) k-tuples of pairwise distinct elements of Γ. Note that $\mathcal{A}$ depends on k but we do not indicate this in the notation for ease of readability, and since $k \geq 2$ is completely fixed throughout the proof. (We made here another small abuse of notation, where we identify the sets $A \in \mathcal{A}$ with a subset of Γ.)

We can hence rewrite the expectation as

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right] = \sum_{A\in\mathcal{A}} \mathbb{P}_\pi\left(T_A > t_{\langle s\rangle}\right). \tag{2.3.5}$$

Now, let us consider the quantity q_A defined in (2.3.3) for $A \in \mathcal{A}$. Applying Proposition 2.2.9, we get the approximation

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right](1+o(1)) = \sum_{A\in\mathcal{A}} \exp\left(-\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A}(\log(n)+s)\right) = \sum_{A\in\mathcal{A}} n^{-q_A} e^{-q_A s}. \tag{2.3.6}$$

Let us first give some intuition on the rest of the proof. Our aim will be to get a good approximation of q_A. If all the points of A are "far away" from each other, then we expect that hitting A occurs on average k times faster as hitting a single point. This suggests that q_A is approximately equal to k.

Suppose for a moment that we were given the following uniform bound: for all $A \in \mathcal{A}$,

$$q_A = k + o\left(\frac{1}{\log n}\right). \tag{2.3.7}$$

Then

$$n^{-q_A} = n^{-k}(1+o(1))$$

and since $|\mathcal{A}| = n^k(1+o(1))$, we would immediately get

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right] = (1+o(1)) \sum_{A\in\mathcal{A}} n^{-q_A} e^{-q_A s} = (1+o(1)) n^k n^{-k} e^{-ks} = (1+o(1)) e^{-ks}, \tag{2.3.8}$$

which is exactly what we want. Therefore, for the proof of the main theorem it would suffice to prove (2.3.7). Of course, such an approximation is wrong if some of the points of A are too close to one another. The rest of the proof will develop arguments to show that the sets A for which (2.3.7) fails are not too numerous (and of course we will also need to control q_A in those cases). For now, it suffices to point out that as soon as (2.3.7) holds, we do not need to study further the properties of A.

Remark 2.3.1. Note that for $k = 1$, $q_A = 1$ for every singleton $A \subset \Gamma$, which proves that $\mathbb{E}_\pi Z \xrightarrow[n\to\infty]{} e^{-s}$. In what follows we will therefore assume that $k \geq 2$.

Initially it is not clear whether one needs to worry about the whole spectrum of possibilities for the mutual distances between points of A or if a cruder bound on the *minimum distance* between points of A is sufficient to distinguish between the good and the bad cases. As it turns out, we are rather lucky that this cruder strategy is sufficient. We will therefore introduce the following quantity:

$$\operatorname{mindist}(A) := \min_{x,y \in A,\, x \neq y} d(x,y). \tag{2.3.9}$$

It will also be convenient to introduce for $1 \leq \delta \leq D$ the notation:

$$\mathcal{A}_\delta = \{A \in \mathcal{A} : \operatorname{mindist}(A) \leq \delta\}\,, \tag{2.3.10}$$

as well as the partial sums:

$$S(\delta) = \sum_{A \in \mathcal{A}_\delta} n^{-q_A} e^{-q_A s}. \tag{2.3.11}$$

With these notations, our aim is to show that

$$S(D) \to e^{-ks}.$$

We record here a simple argument to say that if we assume (2.3.7) for a class of sets A with $\operatorname{mindist}(A) \lll D$ and that the remaining contribution is negligible, then we have the desired conclusion. Although intuitively clear, this argument will appear several times in the next subsections, so we prefer to state it here once and for all.

Lemma 2.3.1. *Fix a sequence $\delta = \delta(n)$ such that $1 \leq \delta \lll D$. Suppose that*

$$S(\delta) = o(1) \tag{2.3.12}$$

and that (2.3.7) is fulfilled uniformly for all $A \notin \mathcal{A}_\delta$. Then

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right] = e^{-ks}(1+o(1)). \tag{2.3.13}$$

Proof. Let $V(r)$ denote the volume of a ball of radius r. Observe that for $r \geq 1$,

$$V(r) \leq \frac{3r}{D} n \tag{2.3.14}$$

because we can fit $\lfloor D/(2r+1) \rfloor$ disjoint balls of radius r in Γ. Consequently, since $\delta = o(n)$, it follows that $V(\delta) = o(n)$ and therefore

$$|\mathcal{A}_\delta| = o(n^k). \tag{2.3.15}$$

We therefore have

$$\begin{aligned} S(D) &= S(\delta) + \sum_{A \in \mathcal{A} \setminus \mathcal{A}_\delta} n^{-q_A} e^{-q_A s} \\ &= o(1) + n^k(1-o(1)) n^{-k} e^{-ks}(1+o(1)) \\ &= e^{-ks}(1+o(1)), \end{aligned}$$

and (2.3.6) allows to conclude. ∎

We also remark that we always trivially have

$$q_A \leq k + o(1), \tag{2.3.16}$$

by using (2.2.75) and a union bound $\mathbb{P}_\pi(T_A < t^*) \leq k\mathbb{P}_\pi(T_o < t^*)$ on the numerator. Such an *upper bound* is useful since q_A appears not only in the term n^{-q_A} but also in the term $e^{-q_A s}$, while $s \in \mathbb{R}$ can be negative.

2.3.2 Organisation of the section

The rest of the section is organised as follows.

- In Section 2.3.3 we simplify a set A into a modified set in which points are sufficiently far away from each other, controlled by a threshold δ which we can choose arbitrarily. We call this the δ-*skeleton* of the set A. We then derive estimates for the probabilities to hit the skeleton within a suitably chosen timescale (and thus on q_A) which will be useful throughout.
- In Section 2.3.4 we show that if $\mathrm{mindist}(A) \leq (\log n)^{1/2}$ (the "microscopic case") then $q_A \geq k-1+\beta'$ for some $\beta' > 0$. Such a bound is sufficient to neglect the overall contribution of such sets to (2.3.6).
- In Section 2.3.5, we make the strong assumption that $n \geq D^5$ (roughly speaking, Γ is at least five-dimensional, so that we are in the "high-dimensional case"). In that case, we complement the results of the previous section by sharper estimates on q_A when $\mathrm{mindist}(A) \geq (\log n)^{1/2}$, and show directly that (2.3.7) holds. This concludes the proof of the main theorem in the high-dimensional case.
- In Section 2.3.6, we suppose $(\log n)^2 \lll f(n)$ and $D^5 \leq n$. This is the "intermediate regime" where Γ is no more than five-dimensional but we assume slightly more than the optimal condition $f(n) \ggg \log n$. In that case, we introduce an intermediary scale, the "mesoscopic scale", where $(\log n)^{1/2} \leq \mathrm{mindist}(A) \leq D^{1/2}$, and show in Proposition 2.3.7 that $q_A \geq k + o(1)$ and therefore the overall contribution to (2.3.6) of mesoscopic sets is negligible. The remaining (macroscopic) scales are handled in Proposition 2.3.8.
- Finally, we handle the most delicate case where $f(n)$ is only assumed to be $\ggg \log n$ in Sections 2.3.7 and 2.3.8. This is done using an elaborate bootstrap argument which will be detailed later.

2.3.3 Skeleton of a set

We will need to group the points of $A \in \mathcal{A}_\delta$ into subsets of points which are close to one another. Let us define an equivalence relation which will allow us to realize such partitions in a convenient way.

Definition 2.3.1. Let $A \subset \Gamma$ and $1 \leq \delta \leq D$. We say that two points $x, y \in A$ are (A,δ)-linked if there exist an integer $r \geq 2$ and a sequence $x = x_1, ..., x_r = y$ of points in A such that for all $1 \leq i \leq r-1$, $d(x_i, x_{i+1}) \leq \delta$. We will write A under the form

$$A = A_1 \sqcup ... \sqcup A_\ell, \tag{2.3.17}$$

where the A_i are the (A,δ)-connected components of A, and $\ell_\delta(A)$ is the number of such components.

Remark 2.3.2. It is worth noting that some points can be closer to points in different connected components than to some points in their own components. For example on $\mathbb{Z}$, take $\delta = 2$ and $A = \{0, 2, 4, 7\}$. The partition is then $\{\{0,2,4\}, \{7\}\}$, though 4 is closer to 7 than to 0.

Let $A \in \mathcal{A}_\delta$. Then, since at least two points of A are in the same component (we assume without loss of generality that $|A_1| \geq 2$), we have $\ell_\delta(A) \leq k-1$. Let now $t^* = t^*(n)$ be (a sequence of times) such that $D^2 \lll t^* \lll n$. From Corollary 2.2.4 we have in particular

$$q_A = \frac{\mathbb{P}_\pi (T_A \leq t^*)}{\mathbb{P}_\pi (T_o \leq t^*)} (1 + o(1)), \tag{2.3.18}$$

where the $o(1)$ is explicit and is of the form $O(n/(t^*f(n))) + O(t^*/n)$.

Let γ be the set of points visited by a random walk starting from the uniform distribution π and run for some time t^*, which is independent of our underlying walk. With this notation and Fubini's theorem we can rewrite $\mathbb{P}_\pi(T_o \leq t^*) = \mathbb{P}(o \in \gamma)$ and $\mathbb{P}_\pi(T_A \leq t^*) = \mathbb{P}(A \cap \gamma \neq \varnothing) = \mathbb{P}(\cup_{i=1}^{\ell_\delta(A)} \{A_i \cap \gamma \neq \varnothing\})$. It is then natural to define and study, for $1 \leq i \leq \ell_\delta(A)$, the events

$$E_i = \{A_i \cap \gamma \neq \varnothing\}\,. \tag{2.3.19}$$

What we want to show is that as the sets A_i are quite far from one another, the event E_i are essentially independent.

Proposition 2.3.2. With the same notations as above, we have

$$\frac{\mathbb{P}(E_i \cap E_j)}{\mathbb{P}_\pi\left(T_o < t^*\right)} \lesssim \max_{\substack{x,y\in\Gamma\\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t^*). \tag{2.3.20}$$

Proof. First observe that for all $x, y \in \Gamma$, if T_x is the hitting of x by γ, then

$$\begin{aligned}\mathbb{P}(\{x,y\} \subset \gamma, T_x < T_y) &= \mathbb{P}(x \in \gamma)\mathbb{P}(T_x < T_y < t^* \mid T_x < t^*)\\ &\leq \mathbb{P}(x \in \gamma)\mathbb{P}_x(T_y < t^*)\end{aligned}$$

by the Markov property. On the other hand, $\mathbb{P}_x(T_y < t^*) = \mathbb{P}_y(T_x < t^*)$ by symmetry, so altogether,

$$\mathbb{P}(\{x,y\} \subset \gamma) \leq 2\mathbb{P}(o \in \gamma)\mathbb{P}_x(T_y < t^*).$$

Therefore,

$$\begin{aligned}\mathbb{P}(E_i \cap E_j) &= \mathbb{P}\left(\bigcup_{x\in A_i, y\in A_j} \{\{x,y\} \subset \gamma\}\right) &(2.3.21)\\ &\leq |A_i|\,|A_j| \max_{\substack{x,y\in\Gamma\\ d(x,y)\geq\delta}} \mathbb{P}(\{x,y\} \subset \gamma) &(2.3.22)\\ &\leq 2k^2\mathbb{P}(o \in \gamma) \max_{\substack{x,y\in\Gamma\\ d(x,y)\geq\delta}} \mathbb{P}_x(T_y < t^*). &(2.3.23)\end{aligned}$$

The proof is concluded by noting that $\max_{\substack{x,y\in\Gamma\\ d(x,y)\geq\delta}} \mathbb{P}_x(T_y < t^*) = \max_{\substack{x,y\in\Gamma\\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t^*)$. ☐

Let us conclude this subsection with a useful lower bound on q_A.

Proposition 2.3.3.

$$q_A - \sum_{i=1}^{\ell_\delta(A)} \frac{\mathbb{P}(E_i)}{\mathbb{P}_\pi(T_o < t^*)} \gtrsim -\left(\max_{\substack{x,y\in\Gamma\\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t^*) + \frac{n}{t^*f(n)} + \frac{t^*}{n}\right). \tag{2.3.24}$$

In particular, if $\delta < \mathrm{mindist}(A)$ and $D^2 \log n \lll t^* \lll n/\log n$ we have

$$q_A - k \gtrsim -\max_{\substack{x,y\in\Gamma\\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t^*) + o\left(\frac{1}{\log n}\right). \tag{2.3.25}$$

Proof. From the Bonferroni inequality, we have

$$\mathbb{P}_\pi\left(T_A < t^*\right) = \mathbb{P}\left(\bigcup_{i=1}^{\ell_\delta(A)} E_i\right) \geq \sum_{i=1}^{\ell_\delta(A)} \mathbb{P}(E_i) - \sum_{1\leq i<j\leq \ell_\delta(A)} \mathbb{P}(E_i \cap E_j). \tag{2.3.26}$$

Dividing on both sides by $\mathbb{P}_\pi\left(T_o < t^*\right)$, we obtain

$$\frac{\mathbb{P}_\pi(T_A < t^*)}{\mathbb{P}_\pi(T_o < t^*)} - \sum_{i=1}^{\ell_\delta(A)} \frac{\mathbb{P}(E_i)}{\mathbb{P}_\pi(T_o < t^*)} \gtrsim - \sum_{1\leq i<j\leq \ell_\delta(A)} \frac{\mathbb{P}(E_i \cap E_j)}{\mathbb{P}_\pi(T_o < t^*)}. \tag{2.3.27}$$

The result then follows from Corollary 2.2.4 and Proposition 2.3.2. □

2.3.4 Microscopic scale

The aim of this subsection is to show that the contribution of the $A \in \mathcal{A}$ which have two very close points is negligible. More precisely, we set $\delta_{\text{micro}} = (\log n)^{1/2}$, and we want to show that

$$S(\delta_{\text{micro}}) = o(1). \tag{2.3.28}$$

This will be proved in Proposition 2.3.5. First, we combine Proposition 2.3.3 with Corollary 2.2.3 and Proposition 2.2.8 to get a lower bound on q_A.

Proposition 2.3.4. There exists $\beta' > 0$ independent of n such that for all n sufficiently large, we have, uniformly over all $A \in \mathcal{A}_{\delta_{\text{micro}}}$,

$$q_A \geq \ell_{\delta_{\text{micro}}}(A) + \beta'. \tag{2.3.29}$$

Proof. To lighten notations, we will write δ for δ_{micro} in this proof. Let $t^* = t^*(n)$ such that $D^2 \lll t^* \lll \min(n, D^3)$. Applying the lower bound on q_A proven in Proposition 2.3.3, together with Corollary 2.2.3, we have,

$$q_A - \sum_{i=1}^{\ell_\delta(A)} \frac{\mathbb{P}(E_i)}{\mathbb{P}_\pi(T_o < t^*)} \geq o(1). \tag{2.3.30}$$

Let x, y be distinct points of A_1 such that $d(x,y) \leq \delta$ and E_1' be the event $\{x \in \gamma\} \cup \{y \in \gamma\}$. Then we have by symmetry,

$$\begin{aligned}\mathbb{P}(E_1') &= \mathbb{P}(x \in \gamma) + \mathbb{P}(y \in \gamma) - \mathbb{P}(\{x,y\} \subset \gamma) \\ &= 2\mathbb{P}(o \in \gamma) - \mathbb{P}_\pi(T_x < t^*, T_y < t^*)\end{aligned} \tag{2.3.31}$$

Now observe that

$$\begin{aligned}\mathbb{P}_\pi(T_x < t^*, T_y < t^*) &\leq \mathbb{P}_\pi(T_{\{x,y\}} < t^*)(1-\beta) \\ &= \mathbb{P}(E_1')(1-\beta)\end{aligned} \tag{2.3.32}$$

where β is the constant appearing in Proposition 2.2.8, $T_{\{x,y\}}$ is the hitting time of the set $\{x, y\}$ and we have used the strong Markov property at this time (together with Proposition 2.2.8). Combining (2.3.31) and (2.3.32) together we get

$$\mathbb{P}(E_1')(2-\beta) \geq 2\mathbb{P}(o \in \gamma),$$

so that

$$\mathbb{P}(E_1) \geq \mathbb{P}(E_1') \geq \frac{2}{2-\beta}\mathbb{P}_\pi(T_o < t^*).$$

Therefore we have proved that

$$\frac{\mathbb{P}(E_1)}{\mathbb{P}_\pi(T_o < t^*)} \geq \frac{2}{2-\beta}. \tag{2.3.33}$$

For the $\ell_\delta(A) - 1$ other E_i's, a very crude bound is sufficient. As the sets A_i are non-empty, we have

$$\frac{\mathbb{P}(E_i)}{\mathbb{P}_\pi(T_o < t^*)} \geq 1, \tag{2.3.34}$$

and we finally get, putting everything together,

$$q_A \geq \frac{2}{2-\beta} + (\ell_\delta(A) - 1) + o(1) = \ell_\delta(A) + \beta' + o(1). \tag{2.3.35}$$

for some $\beta' > 0$ independent of n. □

We are now able to conclude this subsection with the aforementioned result.

Proposition 2.3.5.

$$S(\delta_{\text{micro}}) = o(1). \tag{2.3.36}$$

Proof. Let $1 \leq \ell \leq k$. As points in the same component are at distance at most $k\delta_{\text{micro}}$ from one another, the number of sets $A \in \mathcal{A}_{\delta_{\text{micro}}}$ with ℓ connected components is

$$N(\ell) := |\{A \in \mathcal{A}_{\delta_{\text{micro}}} : \ell_\delta(A) = \ell\}| \lesssim n^\ell V\left(k\delta_{\text{micro}}\right)^k \leq n^\ell \left((d+1)^{k\delta_{\text{micro}}}\right)^k. \tag{2.3.37}$$

As $\delta_{\text{micro}} = o(\log n)$, we deduce that

$$N(\ell) \lesssim n^\ell (d+1)^{k^2\delta_{\text{micro}}} = n^{\ell + o(1)}. \tag{2.3.38}$$

Combining this with Proposition 2.3.4 gives finally (recalling from (2.3.16) that $1 \leq q_A \leq k + o(1)$)

$$S(\delta_{\text{micro}}) \leq \sum_{\ell=1}^{k-1} N(\ell) n^{-(\ell+\beta')} e^{q_A|s|} \lesssim n^{-\beta'} = o(1), \tag{2.3.39}$$

which concludes the proof. □

2.3.5 Convergence in the high-dimensional case

In this subsection, we treat the case $n \geq D^5$ where the diameter is very small compared to the volume. As the growth is faster, the walk is "more transient", and we can use cruder bounds to conclude quickly. The point of including this calculation is that it is quick and illustrates the ideas developed so far usefully, while preparing the ground for the more sophisticated arguments below. Getting the high-dimensional case out of the way from the start is also useful to simplify some arguments later on.

We will show that if $\text{mindist}(A) > \delta_{\text{micro}}$, then (2.3.7) is satisfied under the high-dimensional assumption $n \geq D^5$.

Proposition 2.3.6. Assume that $n \geq D^5$. Then we have

$$\mathbb{E}_\pi \left[Z^{\downarrow k} \right] \to e^{-ks}. \tag{2.3.40}$$

Proof. Let $t^* = D^2(\log n)^{1.1}$. Then, from Proposition 2.2.6, (b), and as $D \gtrsim \log n$, we have uniformly over all $x, y \in \Gamma$ such that $d(x,y) \geq \delta_{\text{micro}} = (\log n)^{1/2}$,

$$\mathbb{P}_x(T_y < t^*) \lesssim \frac{1}{\delta_{\text{micro}}^3} + \frac{t^*}{D^5} = o\left(\frac{1}{\log n}\right). \tag{2.3.41}$$

Combining this with Proposition 2.3.3, and noting that $D^2 \log n \lll t^* \lll n/\log n$, we have, uniformly over all $A \in \mathcal{A} \backslash \mathcal{A}_{\delta_{\text{micro}}}$,

$$q_A \geq k + o\left(\frac{1}{\log n}\right). \tag{2.3.42}$$

Then, since from Proposition 2.3.5 we know that $S(\delta_{\text{micro}}) = o(1)$ the result follows from Proposition 2.3.1. □

2.3.6 Convergence in the case of moderate polynomial growth

In this section, we consider the case

$$D^2(\log n)^2 \lll n \leq D^5. \tag{2.3.43}$$

This time, Condition (2.3.7) does not necessarily hold over all A such that $\text{mindist}(A) > \delta_{\text{micro}}$, and we need to consider another scale, which we call mesoscopic. To this purpose we set

$$\delta_{\text{meso}} := D^{1/2}. \tag{2.3.44}$$

We will prove that Condition (2.3.7) is satisfied if $\text{mindist}(A) \geq \delta_{\text{meso}}$, but first we need to reach this scale and prove that $S(\delta_{\text{meso}}) = o(1)$. This will be our first task. In fact, this estimate will also be needed to start the bootstrap argument in Section 2.3.8 for the proof of Theorem 2.1.1 in its full generality. Therefore we state it under the condition $D^2 \log n \lll n \leq n^5$, so it can be used there as well.

Proposition 2.3.7. Assume that $D^2 \log n \lll n \leq D^5$. Then

$$S(\delta_{\text{meso}}) = o(1). \tag{2.3.45}$$

Proof. We already know from Proposition 2.3.5 that $S(\delta_{\text{micro}}) = o(1)$. Therefore, we only have to study the contribution of the sets A such that $\delta_{\text{micro}} < \text{mindist}(A) \leq \delta_{\text{meso}}$. We know from Proposition 2.3.3 (see in particular (2.3.25)), Corollary 2.2.3, and (2.3.16), that for such sets we have $q_A = k + o(1)$. Moreover, from the volume bound $V(r) \leq \frac{3r}{D} n$ (see (2.3.14)), and as $n \leq D^5$,

$$V(\delta_{\text{meso}}) \lesssim \frac{n}{D^{1/2}} \leq n^{9/10}. \tag{2.3.46}$$

Setting

$$\mathcal{A}_{\text{meso}} = \{A \in \mathcal{A} : \delta_{\text{micro}} < \text{mindist}(A) \leq \delta_{\text{meso}}\}, \tag{2.3.47}$$

we have

$$|\mathcal{A}_{\text{meso}}| \lesssim n^{k-1} V(\delta_{\text{meso}}) \lesssim n^{k-1/10}, \tag{2.3.48}$$

and we finally deduce that

$$\sum_{A \in \mathcal{A}_{\text{meso}}} n^{-q_A} e^{-q_A s} \lesssim |\mathcal{A}_{\text{meso}}| n^{-k+o(1)} \leq n^{-1/10+o(1)} = o(1). \tag{2.3.49}$$

Since $S(\delta_{\text{micro}}) = o(1)$, this concludes the proof of Proposition 2.3.7. □

Proposition 2.3.8. Assume that $D^2(\log n)^2 \lll n \leq D^5$. Then we have

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right] \to e^{-ks}. \tag{2.3.50}$$

Proof. Let $t^* = \min\left(D^2\sqrt{f(n)}, D^2(\log n)^{1.1}\right)$, so that $D^2 \log n \lll t^* \lll n/\log n$. Then, from Proposition 2.2.6, (a), we have uniformly over all $x, y \in \Gamma$ such that $d(x,y) \geq \delta_{\text{meso}} = D^{1/2}$, recalling that $D \geq n^{1/5}$ and $f(n) \ggg (\log n)^2$,

$$\mathbb{P}_x(T_y < t^*) \lesssim \frac{1}{\delta_{\text{meso}}} + \frac{\log D}{f(n)} + \frac{t^*}{n} + \frac{t^*}{D^3} = o\left(\frac{1}{\log n}\right). \tag{2.3.51}$$

Combining this with Proposition 2.3.3 (once again specifically (2.3.25)), and noting that $D^2 \log n \lll t^* \lll n/\log n$, we have, uniformly over all $A \in \mathcal{A}\backslash\mathcal{A}_{\delta_{\text{meso}}}$,

$$q_A \geq k + o\left(\frac{1}{\log n}\right). \tag{2.3.52}$$

Furthermore, from Proposition 2.3.7 we know that $S(\delta_{\text{meso}}) = o(1)$. The result thus follows from Proposition 2.3.1. □

2.3.7 Bootstrap argument

The basis of the bootstrap argument is the following estimate. We first postpone its proof and explain how this implies the desired estimate. Recall that $\delta_{\text{meso}} = D^{1/2}$.

Proposition 2.3.9. There exists a constant K such that for all $\delta_{\text{meso}} \leq \delta \leq D/2$ and $A \in \mathcal{A}$ such that $\delta \leq \text{mindist}(A)$, we have

$$q_A \geq k - K\frac{\log(D/\delta)}{f(n)}.$$

Moreover, the constant K depends only on k and the degree bound d.

Remark 2.3.3. In reality this proposition will only be used with $\delta = o(D)$, so the restriction $\delta \leq D/2$ is not important.

Proof. We postpone the proof until Section 2.3.8. □

We now explain how Proposition 2.3.9 implies the desired estimate. Recall that

$$\mathbb{E}_\pi\left[Z^{\downarrow k}\right](1 + o(1)) = \sum_{A\in\mathcal{A}} n^{-q_A} e^{-q_A s} = S(D), \tag{2.3.53}$$

and recall that by (2.3.16) $q_A \leq k + o(1)$, uniformly over all $A \in \mathcal{A}$, and by (2.3.52) $q_A \geq k + o(1)$, uniformly over all $A \in \mathcal{A}$, such that $\text{mindist}(A) \geq D^{1/2}$. (The upper bound comes from , and see e.g. (2.3.52) for the lower bound). As $s \in \mathbb{R}$ is fixed, it follows that $e^{-q_A s} \leq e^{(k+1)|s|} = O(1)$, so the only term we have to study is n^{-q_A}. Let us also write

$$b(n) := f(n)/\log n, \tag{2.3.54}$$

which tends to infinity by the hypothesis $n/(D^2 \log n) \ggg 1$.

Lemma 2.3.2. *For all $\delta_{meso} \leq \delta \leq \delta' \leq D/2$, we have (with K the constant from Proposition 2.3.9),*

$$S(\delta') - S(\delta) \lesssim \frac{\delta'}{D}\left(\frac{D}{\delta}\right)^{K/b(n)}, \tag{2.3.55}$$

where the implicit constant depends only on k, s.

Proof. From Proposition 2.3.9, we have for $A \in \mathcal{A}_{\delta'} \backslash \mathcal{A}_\delta$,

$$n^{-q_A} \leq n^{-k} n^{K \log(D/\delta)/f(n)} = n^{-k}(D/\delta)^{K(\log n)/f(n)} = n^{-k}(D/\delta)^{K/b(n)}. \tag{2.3.56}$$

Moreover, from the volume bound $V(r) \leq \frac{3r}{D} n$ for $r \geq 1$,

$$|\mathcal{A}_{\delta'} \backslash \mathcal{A}_\delta| \leq |\mathcal{A}_{\delta'}| \leq \binom{k}{2} V(\delta') n^{k-1} \leq \binom{k}{2} \frac{3\delta'}{D} n n^{k-1} = \binom{k}{2} \frac{3\delta'}{D} n^k \tag{2.3.57}$$

For δ sufficiently large, since $e^{-q_A s} \leq e^{(k+1)|s|}$ (since $q_A = k + o(1)$), we finally obtain

$$S(\delta') - S(\delta) = \sum_{A \in \mathcal{A}_{\delta'} \backslash \mathcal{A}_\delta} n^{-q_A} e^{-q_A s} \leq \left(3e^{(k+1)|s|}\binom{k}{2}\right) \frac{\delta'}{D}\left(\frac{D}{\delta}\right)^{K/b(n)}, \tag{2.3.58}$$

as desired. □

Lemma 2.3.2 allows us to increase the value of δ iteratively in such a way that the contribution to (2.3.6) is negligible (except for macroscopic scales). More precisely, set

$$J = J(n) := \left\lfloor 4\frac{\log\log n}{\log(b(n))} - 1 \right\rfloor, \tag{2.3.59}$$

In particular, we have

$$4\frac{\log\log n}{\log(b(n))} - 2 \leq J \leq 4\frac{\log\log n}{\log(b(n))} - 1. \tag{2.3.60}$$

Furthermore, we define a sequence of scales $(\delta_j)_{1 \leq j \leq J(n)}$ by setting for all $1 \leq j \leq J(n)$,

$$\delta_j := D\exp\left(-\frac{\log n}{b(n)^{j/4}}\right) = Dn^{-1/b(n)^{j/4}}. \tag{2.3.61}$$

Note that $\delta_1 \geq \delta_{\text{meso}}$, and that from (2.3.60), we have

$$b(n)^{J/4} \geq b(n)^{(\log\log n)/\log(b(n)) - 1/2} = \frac{\log n}{\sqrt{b(n)}},$$

so

$$\delta_J \geq Dn^{-\sqrt{b(n)}/\log n} = De^{-\sqrt{b(n)}}. \tag{2.3.62}$$

Plugging this into the estimate from Proposition 2.3.9, we see that $q_A = k+$ error, where the error (if $\text{mindist}(A) \geq \delta_J$) is at most

$$K\frac{\log(D/\delta_J)}{f(n)} \leq \frac{K\sqrt{b(n)}}{f(n)} = \frac{K}{(\log n)\sqrt{b(n)}} = o\left(\frac{1}{\log n}\right). \tag{2.3.63}$$

We already know that such an estimate suffices to imply the desired result (see (2.3.7)). Therefore, it will suffice for our purposes to prove the following proposition:

Proposition 2.3.10. We have $S(\delta_J) = o(1)$.

Proof. We already know from Proposition 2.3.7 that $S(\delta_{\text{meso}})$ tends to zero. Furthermore, using Lemma 2.3.2, we see that we also have $S(\delta_1) - \delta_{\text{meso}} \to 0$.

Now observe that for $1 \leq j \leq J-1$, from Lemma 2.3.2 and some elementary computations,

$$S(\delta_{j+1}) - S(\delta_j) \lesssim n^{-\phi}, \tag{2.3.64}$$

where as before the implicit constant is uniform on j, depending only on k, s, d, and

$$\phi := \frac{1}{b(n)^{\frac{j+1}{4}}} - \frac{K}{b(n)^{\frac{j}{4}+1}}.$$

Hence, for n large enough (which we assume in the following), we have $\phi \geq \frac{1}{2b(n)^{(j+1)/4}}$. Then,

$$S(\delta_J) = o(1) + \sum_{j=1}^{J-1} (S(\delta_{j+1}) - S(\delta_j)) \lesssim o(1) + \sum_{j=1}^{J-1} n^{-1/\left(2b(n)^{(j+1)/4}\right)}. \tag{2.3.65}$$

Note that from 2.3.60, we have

$$b(n)^{J/4} \leq (\log n)/b(n)^{1/4}. \tag{2.3.66}$$

Making the change of variables $i = J - j$, we therefore have

$$\sum_{j=1}^{J-1} n^{-1/\left(2b(n)^{(j+1)/4}\right)} = \sum_{i=1}^{J-1} n^{-1/\left(2b(n)^{J/4-(i-1)/4}\right)} \leq \sum_{i=1}^{J-1} e^{-b(n)^{i/4}/2} = \sum_{i=1}^{J-1} u_i, \tag{2.3.67}$$

with $u_i := \exp(-b(n)^{i/4}/2)$. Note also that for n sufficiently large, we have for all $1 \leq i \leq J-1$ that

$$\frac{u_{i+1}}{u_i} \leq 1/e. \tag{2.3.68}$$

The sum $\sum_{i=1}^{J(n)-1} u_i$ being subgeometric, it is of the same order of magnitude as its first term, u_1, which tends to 0 since $b(n) \to \infty$. This concludes the proof of Proposition 2.3.10. $\square$

2.3.8 Proof of Proposition 2.3.9

Let $\delta_{\text{meso}} \leq \delta \leq D/2$ and $A \in \mathcal{A}$ such that $\delta < \text{mindist}(A)$. To estimate q_A we start from the definition of q_A as

$$q_A = \frac{\mathbb{E}_\pi(T_o)}{\mathbb{E}_\pi(T_A)}.$$

We will rely on the following well-known identity relating the expected hitting time of a vertex to the "fundamental matrix" of the chain. For this it is useful to introduce the discrete time lazy random walk $(X_t^\#, t = 0, 1, \dots)$ on Γ, which is the Markov chain with transition matrix $K = (I + P)/2$ where P is the transition matrix of simple random walk on Γ. We note that if $T_o^\#$ and $T_A^\#$ denote its hitting times of o and A respectively, then by conditioning on the jump chain we easily get

$$\mathbb{E}_\pi(T_o) = (1/2)\mathbb{E}_\pi(T_o^\#) \quad \text{and} \quad \mathbb{E}_\pi(T_A) = (1/2)\mathbb{E}_\pi(T_A^\#),$$

so that $q_A = \mathbb{E}_\pi(T_o^\#)/\mathbb{E}_\pi(T_A^\#)$. These expected hitting times can then be evaluated through the following identity (see [LP17, Proposition 10.26] and [AF, Lemma 2.11]).

Lemma 2.3.3. *Let X be a discrete-time irreducible and aperiodic Markov chain on a finite state space S with invariant distribution $\pi = (\pi_i)_{i \in S}$. For $j \in S$, let T_j denote the hitting time of j by X and let $p_m(x, y)$ denote its m-step transition probabilities. Then*

$$\mathbb{E}_\pi(T_j) = \frac{1}{\pi(j)} \sum_{m=0}^{\infty} [p_m(j, j) - \pi(j)].$$

While this is stated for the hitting time of a single state, we can also use this lemma to compute $\mathbb{E}_\pi(T_A^\#)$. One should however be a little careful. Indeed, as highlighted in [AF] (see Section 2.3.3) one cannot naively hope that the last formula holds with replacing $p_m(j, j)$ by $\mathbb{P}_{\pi_A}[X_m \in A]$ in Lemma 2.3.3, where π_A denotes the uniform distribution on A. Instead, we will apply this lemma to the lazy random walk on the following **weighted collapsed graph** Γ_A defined as follows. The vertices of Γ_A are obtained by replacing all the vertices in A by a single vertex, which we denote (with a small abuse of notation) by A. The edges of Γ_A are those induced by the edges of Γ and this wiring. Note that Γ_A may be a multigraph (there can be several edges between A and a vertex x if x had several neighbours in A in the original graph Γ) and may contain loops if A contained adjacent vertices in the original graph Γ. To the edges of this (multi)graph Γ_A we associate weights w as follows: let e be an edge of Γ_A. Then

- if at least one endpoint of e is distinct from A we set $w(e) = 1$
- if instead e is a loop from A to A then we set $w(e) = 2$.

Let $\tilde{p}_m^\#(x, y)$ denote the m-step transition probabilities of the random walk in discrete time $(\tilde{X}_m^\#, m = 0, 1, \ldots)$ on this graph, with laziness parameter $1/2$ at each vertex (including at A). In other words, this is the m-step transition probability of the Markov chain $\tilde{X}^\#$ whose transition matrix is by definition $\widetilde{K} = (I + \widetilde{P})/2$, where $\widetilde{P}(x, y) = n(x, y)w(x, y)/\sum_z n(x, z)w(x, z)$, and $n(x, y)$ is the number of edges in the multigraph Γ_A that lead from x to y. An elementary computation shows that the Markov chain $\tilde{X}^\#$ is reversible w.r.t. the distribution $\tilde{\pi}$ given by $\tilde{\pi}(x) = 1/n$ for $x \in \Gamma_A$ with $x \neq A$, and $\tilde{\pi}(A) = k/n$. In particular, $\tilde{\pi}$ is the invariant distribution of $\tilde{X}^\#$. In other words, $\tilde{\pi} = \pi$, with a small abuse of notation (this is one of the reasons for giving weight 2 to edges internal to A).

Note that until hitting A, the Markov chain $\tilde{X}^\#$ on Γ_A coincides with the original lazy random walk $X^\#$ on Γ, so that $\mathbb{E}_\pi(T_A^\#) = \mathbb{E}_\pi(\tilde{T}_A^\#)$, where $\tilde{T}_A^\#$ is the hitting time of (the vertex) A by $\tilde{X}^\#$.

Therefore, by Lemma 2.3.3,

$$\mathbb{E}_\pi(T_A^\#) = \frac{1}{\pi(A)} \sum_{m=0}^{\infty} [\tilde{p}_m^\#(A, A) - \pi(A)].$$

Since $\pi(A)/\pi(o) = k$, we can rewrite q_A under the discrete form

$$q_A = k \frac{\displaystyle\sum_{m=0}^{\infty} [p_m^\#(o, o) - \pi(o)]}{\displaystyle\sum_{m=0}^{\infty} [\tilde{p}_m^\#(A, A) - \pi(A)]} \tag{2.3.69}$$

and in order to prove Proposition 2.3.9 it suffices to show that

$$\frac{\displaystyle\sum_{m=0}^{\infty} [p_m^\#(o, o) - \pi(o)]}{\displaystyle\sum_{m=0}^{\infty} [\tilde{p}_m^\#(A, A) - \pi(A)]} \geq 1 - C \frac{\log(D/\delta)}{f(n)} \tag{2.3.70}$$

for some constant C depending only on k, s, d.

We start with the numerator, and observe that since the chain $X^{\#}$ is lazy (hence all its eigenvalues are nonnegative), $p^{\#}_m(o,o) \geq \pi(o)$ for all m, therefore:

$$\sum_{m=0}^{\infty}[p^{\#}_m(o,o) - \pi(o)] \geq \sum_{m=0}^{\delta^2}[p^{\#}_m(o,o) - \frac{1}{n}] \geq \left(\sum_{m=0}^{\delta^2} p^{\#}_m(o,o)\right) - \frac{1}{f(n)},$$

where in the last inequality we used $\delta^2 \leq D^2 = n/f(n)$. The bulk of the proof of (2.3.70) is therefore to give an upper bound to the denominator. We divide the proof into the following two steps.

Lemma 2.3.4. *We have, uniformly over sets A of size k such that* $\mathrm{mindist}(A) \geq \delta \geq \delta_{meso}$*,*

$$\sum_{m=0}^{\delta^2} \tilde{p}^{\#}_m(A,A) \leq \sum_{m=0}^{\delta^2} p^{\#}_m(o,o) + O(1/f(n)).$$

Lemma 2.3.5. *We have, uniformly over sets A of size k such that* $\mathrm{mindist}(A) \geq \delta \geq \delta_{meso}$*,*

$$\sum_{m=\delta^2}^{2D^2} \tilde{p}^{\#}_m(A,A) \lesssim \frac{\log(D/\delta)}{f(n)}.$$

Before proving these lemmas, we first explain how they imply (2.3.70). By Corollary 3.27 in [AF] we observe that the spectral gap of the collapsed chain is greater than that of the original chain and therefore (as already remarked in (2.2.16)) the relaxation time of the lazy version of the collapsed chain is at most $2dD^2$. From Lemma 2.3.5 and the Poincaré inequality it therefore follows that

$$\sum_{m=D^2}^{\infty}[\tilde{p}^{\#}_m(A,A) - \tilde{\pi}(A)] \lesssim \frac{\log(D/\delta)}{f(n)}. \tag{2.3.71}$$

Indeed the Poincaré inequality (for the collapsed chain) implies that the summand decays geometrically fast for $t \gtrsim D^2$ since the relaxation time is $\lesssim D^2$. Hence the integral from D^2 to infinity is bounded by a geometric sum which is itself bounded by its first term up to a constant. This first term is at most the right hand side of Lemma 2.3.5.

Let us explain this argument in more detail. First recall that by spectral decomposition (since $\tilde{K}$ is reversible, see for instance [LP17, Lemma 12.2]) for every $m = 0, 1, \ldots$ and

$$\frac{\tilde{p}^{\#}_m(x,y)}{\tilde{\pi}(y)} = 1 + \sum_{j=2}^{n} f_j(x) f_j(y) \lambda_j^m, \tag{2.3.72}$$

where $0 \leq \lambda_n \leq \ldots \leq \lambda_2 < \lambda_1 = 1$ are the eigenvalues of $\widetilde{K}$ (which are nonnegative since the chain is lazy) and $f_1, ..., f_n$ are the associated eigenfunctions. We immediately deduce, that for all $t \geq 0$ and a vertex x,

$$\frac{\tilde{p}^{\#}_{t+s}(x,x)}{\tilde{\pi}(x)} - 1 = \sum_{j=2}^{n} f_j(x)^2 \lambda_j^{t+s} \leq \left(\frac{\tilde{p}^{\#}_t(x,x)}{\tilde{\pi}(x)} - 1\right)\lambda_2^s \leq \left(\frac{\tilde{p}^{\#}_t(x,x)}{\tilde{\pi}(x)} - 1\right) e^{-s/\widetilde{t_{\mathrm{rel}}}},$$

where $\widetilde{t_{\mathrm{rel}}} = 1/(1-\lambda_2)$. In other words,

$$\tilde{p}^{\#}_{t+s}(x,x) - \tilde{\pi}(x) \leq e^{-s/\widetilde{t_{\mathrm{rel}}}}(\tilde{p}^{\#}_t(x,x) - \tilde{\pi}(x)). \tag{2.3.73}$$

We can sum this inequality over intervals of length roughly $\widetilde{t_{\mathrm{rel}}}$ to deduce that the sum from $m\widetilde{t_{\mathrm{rel}}}$ to $(m+1)\widetilde{t_{\mathrm{rel}}}$ of $\tilde{p}_t^{\#}(x,x) - \tilde{\pi}(x)$ decays exponentially in m. That is, taking $x = A$,

$$\begin{aligned}
\sum_{t=D^2}^{\infty} [\tilde{p}_t^{\#}(A,A) - \tilde{\pi}(A)] &= \sum_{m=1}^{\infty} \sum_{t=mD^2}^{(m+1)D^2-1} [\tilde{p}_t^{\#}(A,A) - \tilde{\pi}(A)] \\
&\leq \sum_{m=1}^{\infty} \left(\sum_{t=D^2}^{2D^2} e^{-(m-1)D^2/\widetilde{t_{\mathrm{rel}}}} [\tilde{p}_t^{\#}(A,A) - \tilde{\pi}(A)] \right) \\
&\lesssim \sum_{t=D^2}^{2D^2} [\tilde{p}_t^{\#}(A,A) - \tilde{\pi}(A)],
\end{aligned}$$

where in the last line we used that $2D^2/\widetilde{t_{\mathrm{rel}}} \geq 1/d$, as it is easy to check that $\widetilde{t_{\mathrm{rel}}} \leq 2\, t_{\mathrm{rel}} \leq 2dD^2$ (the inequality $\widetilde{t_{\mathrm{rel}}} \leq 2\, t_{\mathrm{rel}}$ follows easily from comparing the Dirichlet forms associated to $\widetilde{K}$ and the original matrix K, see, e.g., Corollary 4.1 in [Ber16], or [LP17, Lemma 13.7]).

Therefore, Lemmas 2.3.4 and 2.3.5 together imply (2.3.70).

It remains to prove Lemmas 2.3.4 and 2.3.5.

Proof of Lemma 2.3.4. A key idea for this lemma will be a probabilistic construction of $\tilde{X}^{\#}$ in terms of excursions away from o of the original lazy walk $X^{\#}$ on the vertex-transitive graph Γ. We explain this construction first. Let $(e_i, i \geq 1)$ be a sequence of i.i.d. *excursions* of the walk $X^{\#}$ (on Γ) from o to o: that is, each e_i is a lazy random walk on Γ starting at o, and killed upon returning to o for the first time (let ζ_i be this time). Consider also an independent and i.i.d. sequence $(a_i)_{i\geq 1}$ such that a_i is uniformly distributed on A for each i. Roughly speaking, we will use a_i to translate the excursion e_i: that is, for a vertex v of Γ, fix ϕ_v a graph isomorphism such that $\phi_v(o) = v$. We then set

$$\tilde{e}_i := \phi_{a_i}(e_i),$$

which defines an independent sequence of excursions based at a_i (instead of o) respectively. The idea is then to concatenate the excursions $\tilde{e}_i$ to form the Markov chain $\tilde{X}^{\#}$ on the collapsed graph Γ_A. However, we need to pay attention to the fact that during the excursion $\tilde{e}_i$ it is also possible to hit A through a different point than merely a_i. Let us define

$$\tau_i := \inf \left\{ t = 1, 2, \ldots : \tilde{e}_i(t) \in A \setminus \{a_i\} \right\}$$

with the usual convention that $\inf \varnothing = \infty$. To obtain the Markov chain $\tilde{X}^{\#}$ on Γ_A, we simply concatenate the truncated excursions $\eta_i := (\tilde{e}_i(t), 0 \leq t \leq \zeta_i \wedge \tau_i - 1)$. We leave it to the reader to check by the Markov property the following claim:

Claim: the concatenation of $(\eta_i)_{i\geq 1}$ is a realisation of $(\tilde{X}_t^{\#}, t = 0, 1, \ldots)$ on Γ_A.

When the stopping time τ_i is infinite, this corresponds to the chain $\tilde{X}^{\#}$ returning to A in a way which we think of as “simple”, corresponding to an excursion from a point $a_i \in A$ to itself, when the walk is seen in Γ; as we will see below the expected number of returns to A in this fashion will be simply controlled by the heat kernel $p_t^{\#}(o,o)$ on the base graph, for which we have good estimates. We will therefore be especially interested in controlling returns to A that occur in the opposite case when $\tau_i < \zeta_i$, i.e., when the excursion seen in Γ corresponds to a path from a point $a_i \in A$ to a distinct point $b_i \in A$. We will refer to such an event as a **thread transition** (from A to A).

Let τ denote the first time that there is such a thread transition. The main argument will be to show that

$$\mathbb{P}\left(\tau \leq \delta^2\right) = O(1/f(n)). \tag{2.3.74}$$

Let us first explain why (2.3.74) implies the conclusion of the lemma. Let $\tilde{L}_A^{\#}(m)) = \sum_{i=0}^{m-1} \mathbf{1}\{\tilde{X}_i^{\#} = A\}$ and $L_o^{\#}(m)) = \sum_{i=0}^{m-1} \mathbf{1}\{X_i^{\#} = o\}$. We observe that

$$\mathbb{E}_A(\tilde{L}_A^{\#}(\tau \wedge \delta^2)) = \mathbb{E}_o(L_o^{\#}(\tau \wedge \delta^2)) \leq \mathbb{E}_o(L_o^{\#}(\delta^2)), \tag{2.3.75}$$

since before τ the only returns to A by $\tilde{X}^{\#}$ correspond to the lifetimes ζ_i of the excursions $\tilde{e}_i$ or, equivalently, of e_i. When we concatenate the excursions e_i we get a random walk $X^{\#}$ on the original graph Γ, and so these returns correspond to $X^{\#}$ hitting o. Therefore, applying the strong Markov property of $\tilde{X}^{\#}$ at time τ, we get

$$\begin{aligned}
\mathbb{E}_A(\tilde{L}_A^{\#}(\delta^2)) &= \mathbb{E}_A(\tilde{L}_A^{\#}(\tau \wedge \delta^2)) + \mathbb{E}_A\left(\tilde{L}_A^{\#}(\delta^2) - \tilde{L}_A^{\#}(\tau \wedge \delta^2)\right) \\
&\leq \mathbb{E}_o(L_o^{\#}(\delta^2)) + \mathbb{P}(\tau < \delta^2)\mathbb{E}_A(\tilde{L}_A^{\#}(\delta^2)) \\
&\leq \mathbb{E}_o(L_o^{\#}(\delta^2)) + O(1/f(n))\mathbb{E}_A(\tilde{L}_A^{\#}(\delta^2))
\end{aligned}$$

where the second inequality follows by (2.3.74), and the first by averaging over τ and noting that $\tilde{L}_A^{\#}(\cdot)$ is non-decreasing. Therefore, putting the last term on the right hand side in the left hand side, we obtain

$$\mathbb{E}_A(\tilde{L}_A^{\#}(\delta^2)) \leq (1 + O(1/f(n)))\mathbb{E}_o(L_o^{\#}(\delta^2)) = \mathbb{E}_o(L_o^{\#}(\delta^2)) + O(1/f(n)),$$

recalling that $\mathbb{E}_o(L_o^{\#}(\delta^2)) = O(1)$, using continuous time estimates (see, e.g., Proposition 2.2.7) and using basic properties of Poisson processes. This is the required bound for Lemma 2.3.4.

It remains to prove (2.3.74). For this we will again use the construction above in terms of excursions and use the following crude bound.

Consider the finite set

$$B := \bigcup_{a \in A} \phi_a^{-1}(A \setminus \{a\}). \tag{2.3.76}$$

In words, B represents "all the ways to see the rest of A, viewed from an arbitrary point in A". Note that B is a set of cardinality at most $k(k-1) \leq k^2$.

Note also that $d(o, b) \geq \delta$ for all $b \in B$ (since $\operatorname{mindist}(A) \geq \delta$). Furthermore, by Markov's inequality,

$$\begin{aligned}
\mathbb{P}(\tau < \delta^2) &\leq \mathbb{P}_o(T_B^{\#} < \delta^2) \leq \mathbb{E}_o(L_B^{\#}(\delta^2)) \\
&\leq \sum_{y \in B} \sum_{t=0}^{\delta^2} p_t^{\#}(o, y).
\end{aligned}$$

We bound the expected discrete local time by the expected continuous local time in order to use our off-diagonal estimates on the heat kernel of Proposition 2.2.6. More precisely, note that if N is a Poisson random variable with mean $10\delta^2$ (corresponding to the number of jumps in continuous time which occurred by time $t = 10\delta^2$), then for some absolute constant $c > 0$

$$\begin{aligned}
\sum_{m=0}^{\delta^2} p_m^{\#}(o, y) &\leq \int_0^{10\delta^2} p_t(o, y) \mathrm{d}t + \delta^2 \mathbb{P}(N \leq 2\delta^2) \\
&\leq O(1/f(n)) + \delta^2 \exp(-c\delta^2),
\end{aligned}$$

where for the first we used Proposition 2.2.6 and the fact that $d(o, y) \geq \delta$, and for the second term, we used standard Chernoff bounds for Poisson random variables. Since $\delta \geq \delta_{\text{meso}} \geq \sqrt{D}$, we clearly have that the second term is negligible compared to the first.

This concludes the proof of (2.3.74) and therefore also of Lemma 2.3.4. □

Proof of Lemma 2.3.5. We wish to argue as in Proposition 2.2.1, but we cannot directly apply this result since the graph Γ_A of the collapsed chain is not vertex-transitive. Nevertheless, it is easy to see that the same bounds apply to this chain.

First observe that the conductance profile ϕ_A of the collapsed chain is larger than that of the original chain: that is, $\phi(u) \lesssim \phi_A(u)$.

Indeed, to every set $\tilde{S} \subset \Gamma_A$ corresponds naturally an *uncollapsed* set S of vertices in Γ (this set S is nothing but $\tilde{S} \setminus \{\{A\}\} \cup A$), and since A is sparse (recall that $\text{mindist}(A) \geq \delta_{meso}$), those sets satisfy $Q_A(\tilde{S}, \tilde{S}^c) = Q(S, S^c)$. Conversely, given any set S of vertices in Γ such that $S \supset A$ or $S \cap A = \varnothing$ (i.e., S does not contain just a portion of A) then we see that S corresponds to a set $\tilde{S}$ of vertices in Γ_A such that $Q_A(\tilde{S}, \tilde{S}^c) = Q(S, S^c)$ and thus $\Phi(S) = \Phi_A(\tilde{S})$.

We deduce that for $1/n \leq u \leq 1/2$,

$$\begin{aligned}
\phi(u) &= \inf\left\{\Phi(S) : \pi(S) \leq u\right\} \\
&\leq \inf\left\{\Phi(S) : \pi(S) \leq u \text{ and } [S \supset A \text{ or } S \cap A = \varnothing]\right\} \\
&= \inf\left\{\Phi_A(\tilde{S}) : \pi(\tilde{S}) \leq u\right\} \\
&= \phi_A(u).
\end{aligned}$$

Combining this inequality and (2.2.7), we see that Proposition 2.2.1 still applies here, even without vertex-transitivity, and that only the constants differ (there is for instance an extra factor k in (2.2.4), as $\pi(A) = k\pi(o)$). This also holds in discrete-time, as noted in Remark 2.2.3.

Therefore, reasoning as in Corollary 2.2.1, we also have the following bounds: If $f(n) < D$, then for every $x \in \Gamma_A$, for $1 \leq t \leq f(n)^2$, $\tilde{p}_t^{\#}(x, x) \lesssim t^{-3/2}$, and for every $f(n)^2 \leq t \leq D^2$,

$$\tilde{p}_t^{\#}(x, x) \lesssim \frac{1}{t f(n)}. \tag{2.3.77}$$

As mentioned above, the only difference is that now the implicit constants depend also on k. Using this, we immediately deduce (recalling that $\delta^2 \geq f(n)^2$) that

$$\sum_{t=\delta^2}^{D^2} \tilde{p}_t^{\#}(A, A) \lesssim \sum_{t=\delta^2}^{D^2} \frac{1}{t f(n)} \lesssim \frac{\log(D/\delta)}{f(n)}. \tag{2.3.78}$$

This completes the proof of Lemma 2.3.5 and thus also of (2.3.70). In turn, the proof of Theorem 2.1.1 is complete. □

2.4 Law of the uncovered set

The goal of this section is to show that the information on the law of the cover time can be supplemented by a precise description of the law of the uncovered sets before the cover time. From the results in Section 2.3 and in particular from (2.3.1), we know that at a time $t_{\langle s \rangle} = \mathbb{E}_\pi T_o(s + \log n)$, the *size* of the uncovered set converges to a Poisson random variable with parameter e^{-s}. We will then turn to describe the geometry of this uncovered set: roughly, we aim to show that the uncovered points are approximately uniformly chosen from the vertex set of the graph.

There are two different ways to express this idea. The first one is to consider the fixed time $t_{\langle s\rangle}$ and condition on the size $Z_s = |U(t_{\langle s\rangle})| = k$ of the uncovered set at this time, and show that the law of the set $U(t_{\langle s\rangle})$ itself is close (in the total variation sense) to a uniformly chosen set of size k. This is carried out in Section 2.4.1. Another way is to consider the stopping time τ_k which is the first time at which the size of uncovered set is equal to k (so $\tau_0 = \tau_{\text{cov}}$), and prove the same approximate uniformity. This is carried out in Section 2.4.2.

As we will see, a stronger form of convergence (namely, convergence in total variation) for the cover time itself will follow relatively quickly from these results. This will be explained in Section 2.4.3.

2.4.1 Convergence to a product measure

In this subsection, we prove the first form of uniformity for the uncovered set mentioned above. The statement is as follows. We recall that from the proof in Section 2.3, in all growth cases, there exists $\delta_* = o(D)$ such that $q_A \geq k - o\left(\frac{1}{\log n}\right)$ uniformly over sets A of size k such that $\operatorname{mindist}(A) > \delta_*$. (Indeed, δ_* can be taken to be δ_{micro} in the high dimensional case, δ_{meso} in the intermediate regime, and δ_{macro} in the most difficult regime, the low-dimensional case). For convenience, let $\mathcal{A}_*$ denote the subsets $A \subset \Gamma$ (i.e., A is a set of vertices of Γ) such that $\operatorname{mindist}(A) > \delta_*$.

Theorem 2.4.1. Let $k \geq 1$ and $s \in \mathbb{R}$. Uniformly over sets $A \in \mathcal{A}_*$ of size k,

$$\mathbb{P}(U(t_{\langle s\rangle}) = A | Z_s = k) = \frac{k!}{n^k}(1 + o(1)). \tag{2.4.1}$$

In particular, if Unif_k denotes the uniform law on subsets of size k, and if $U(t_{\langle s\rangle}\,|k)$ denotes the law of $U(t_{\langle s\rangle})$ conditionally given $Z_s = k$, then

$$\mathrm{d}_{\text{TV}}(U(t_{\langle s\rangle}\,|k); \text{Unif}_k) \to 0$$

as $n \to \infty$.

Remark 2.4.1. It is not hard to see (using the thinning property of Poisson random measures) that an equivalent formulation of Theorem 2.4.1 is as follows: without any conditioning, the law of $U(t_{\langle s\rangle})$ is close in the total variation sense to the law of the set obtained by tossing an independent coin for each vertex, with probability of heads given by e^{-s}/n. In other words, the law of $U(t_{\langle s\rangle})$ is close to that of a product measure $\mu_s^{\otimes\Gamma}$ where the product is over all vertices of the graph, and the law μ_s is Bernoulli with parameter e^{-s}/n. Thus,

$$\mathrm{d}_{\text{TV}}(U(t_{\langle s\rangle}), \mu_s^{\otimes\Gamma}) \to 0,$$

where, as is standard, if X is a random variable and μ a law, $\mathrm{d}_{\text{TV}}(X, \mu)$ denotes the total variation between the law of X and μ.

We start the proof of Theorem 2.4.1 with the following simple lemma.

Lemma 2.4.1. *Let $k \geq 1$ and $s \in \mathbb{R}$. Then we can neglect sets that are not sufficiently well separated:*

$$\sum_{A \in \mathcal{A}_*^c, |A|=k} \mathbb{P}(U(t_{\langle s\rangle}) = A) = o(1). \tag{2.4.2}$$

Proof. This follows from the trivial bound

$$\mathbb{P}(U(t_{\langle s\rangle}) = A) \le \mathbb{P}(T_A > t_{\langle s\rangle}) \tag{2.4.3}$$

and bounds obtained previously. Indeed, by Proposition 2.2.9 we have

$$\mathbb{P}(T_A > t_{\langle s\rangle}) = n^{-q_A} e^{-q_A s}(1+o(1)), \tag{2.4.4}$$

and by choice of δ_*, as shown in Section 2.3, for all growth cases, $S(\delta_*) = o(1)$. This implies that

$$\sum_{A\in\mathcal{A}_*^c |A|=k,} \mathbb{P}(U(t_{\langle s\rangle}) = A) \le S(\delta_*)(1+o(1)) = o(1), \tag{2.4.5}$$

as desired. □

The proof of Theorem 2.4.1 will therefore follow from the following lemma and Lemma 2.4.1.

Lemma 2.4.2. *Let $k \ge 1$ and $s \in \mathbb{R}$. Then, uniformly over sets $A \in \mathcal{A}_*$ of size k,*

$$\mathbb{P}(U(t_{\langle s\rangle}) = A) = \frac{e^{-e^{-s}} e^{-ks}}{n^k}(1+o(1)). \tag{2.4.6}$$

Proof. We cannot directly apply here the moment method (or factorial moment method) as we did in Section 2.3 but we note that there is a relatively simple way to use the work done in this section nevertheless, by exploiting instead the Bonferroni inequalities.

Let us fix a set A as in the lemma, and observe that we can rewrite

$$\begin{aligned}\mathbb{P}(U(t_{\langle s\rangle}) = A) &= \mathbb{P}(A \text{ is not touched but all the points in } \Gamma\backslash A \text{ are}) &(2.4.7)\\ &= \mathbb{P}\left(\left\{T_A > t_{\langle s\rangle}\right\} \cap \bigcap_{x\in\Gamma\backslash A} \left\{T_x \le t_{\langle s\rangle}\right\}\right). &(2.4.8)\end{aligned}$$

For $X \subset \Gamma\backslash A$, set $E_X := \left\{T_{A\cup X} > t_{\langle s\rangle}\right\}$, with the standard abuse of notations when X is a singleton. Then

$$\begin{aligned}\mathbb{P}(U(t_{\langle s\rangle}) \ne A) &= \mathbb{P}\left(\left\{T_A \le t_{\langle s\rangle}\right\} \cup \bigcup_{x\in\Gamma\backslash A} \left\{T_x > t_{\langle s\rangle}\right\}\right) &(2.4.9)\\ &= \mathbb{P}\left(T_A \le t_{\langle s\rangle}\right) + \mathbb{P}\left(\left\{T_A > t_{\langle s\rangle}\right\} \cap \bigcup_{x\in\Gamma\backslash A} \left\{T_x > t_{\langle s\rangle}\right\}\right) &(2.4.10)\\ &= \mathbb{P}\left(T_A \le t_{\langle s\rangle}\right) + \mathbb{P}\left(\bigcup_{x\in\Gamma\backslash A} E_x\right). &(2.4.11)\end{aligned}$$

By Proposition 2.2.9, uniformly over sets $A \in \mathcal{A}^*$ of size k,

$$\mathbb{P}(T_A > t_{\langle s\rangle}) = (1+o(1))\frac{e^{-ks}}{n^k}. \tag{2.4.12}$$

Consequently, we have

$$\mathbb{P}(U(t_{\langle s\rangle}) = A) = (1+o(1))\frac{e^{-ks}}{n^k} - \mathbb{P}\left(\bigcup_{x\in\Gamma\backslash A} E_x\right). \tag{2.4.13}$$

Observe that $E_X \cap E_Y = E_{X\cup Y}$. Note also that if $\operatorname{mindist}(A \cup X) \geq \delta_*$ and $|X| = j$, then $q_{A\cup X} \geq k + j - o(1/\log n)$. Therefore,

$$n^k \sum_{X\subset\Gamma\setminus A, |X|=j, A\cup X\in\mathcal{A}_*} \mathbb{P}\left(E_X\right) \to \frac{e^{-(k+j)s}}{j!}. \tag{2.4.14}$$

Choose $J = J(k, \varepsilon, s)$ such that

$$\sum_{j=J+1}^{\infty} \frac{e^{(k+j)s}}{(k+j)!} \leq \varepsilon.$$

Having chosen J, proceeding as in the proof of Lemma 2.4.1, we can ignore sets $A \cup X \in \mathcal{A}_*^c$ and write for every $1 \leq j \leq J + 1$,

$$n^k \sum_{X\subset\Gamma\setminus A, |X|=j, A\cup X\in\mathcal{A}_*^c} \mathbb{P}\left(E_X\right) \leq \frac{\varepsilon}{J}. \tag{2.4.15}$$

We deduce from the Bonferroni inequalities that for n large enough, and separating according to whether $A \cup X \in \mathcal{A}_*$ or not,

$$\begin{aligned} n^k \left| \mathbb{P}\left(\bigcup_{x\in\Gamma\setminus A} E_x \right) - \sum_{j=1}^{J} (-1)^{j+1} \sum_{X\subset\Gamma\setminus A, |X|=j} \mathbb{P}\left(E_X\right) \right| &\leq \sum_{X\subset\Gamma\setminus A: |X|=J+1} \mathbb{P}(E_X) && (2.4.16)\\ &\leq \frac{\varepsilon}{J} + \frac{e^{(k+J+1)s}}{(k+J+1)!} + \varepsilon && (2.4.17)\\ &\leq 3\varepsilon. && (2.4.18) \end{aligned}$$

Consequently,

$$n^k \left| \mathbb{P}\left(\bigcup_{x\in\Gamma\setminus A} E_x \right) - \sum_{j=1}^{J} (-1)^{j+1} \sum_{X\subset\Gamma\setminus A, |X|=j, A\cup X \mathcal{A}_*} \mathbb{P}\left(E_X\right) \right| \leq 3\varepsilon + J\varepsilon/J = 4\varepsilon. \tag{2.4.19}$$

By (2.4.14), we may assume that n is large enough that for $1 \leq j \leq J$,

$$\left| n^k \left(\sum_{X\subset\Gamma\setminus A, |X|=j, A\cup X\in\mathcal{A}_*} \mathbb{P}\left(E_X\right) \right) - \frac{e^{-(k+j)s}}{j!} \right| \leq \frac{\varepsilon}{J}. \tag{2.4.20}$$

Combining this with the definition of J and by another triangle inequality, we obtain, as

$$e^{-ks} - \sum_{j=1}^{\infty} (-1)^{j+1} \frac{e^{-(k+j)s}}{j!} = e^{-ks} e^{-e^{-s}}, \tag{2.4.21}$$

that for n large enough, we have

$$\left| n^k \mathbb{P}(U(t_{\langle s\rangle}) = A) - e^{-ks} e^{-e^{-s}} \right| \leq 6\varepsilon, \tag{2.4.22}$$

which concludes the proof. $\square$

2.4.2 Convergence of the last k points

In this section we are interested in the first time at which the uncovered set has size k, that is:

$$\tau_k := \inf\{t \geq 0 : |\{X_u, u \leq t\}^c| = k\}.$$

We want to show that the distribution of $U(\tau_k)$ is close to the uniform distribution Unif_k over sets of size k.

Theorem 2.4.2. As $n \to \infty$, we have

$$\mathrm{d_{TV}}(U(\tau_k), \mathrm{Unif}_k) \to 0. \tag{2.4.23}$$

We will need two ingredients.

Our first lemma improves on Theorem 2.4.1 by showing that the position of the walk at time $t_{\langle s\rangle}$ is approximately independent of $U(t_{\langle s\rangle})$ and is distributed (approximately again) according to π.

Lemma 2.4.3. *Let $s \in \mathbb{R}$. Recall the product law $\mu_s^{\otimes\Gamma}$ from Remark 2.4.1. Then*

$$d_{TV}[(U(t_{\langle s\rangle}), X_{t_{\langle s\rangle}}); (\mu_s^{\otimes\Gamma} \otimes \pi)] \to 0.$$

Proof. Recall that $Z_s = |U(t_{\langle s\rangle})|$. Fix $s' < s$ in such a way that

$$D^2 \lll t_{\langle s\rangle} - t_{\langle s'\rangle} \lll n = D^2 f(n). \tag{2.4.24}$$

The upper bound of (2.4.24) implies that $s - s' = o(1)$ and so that $\mathbb{E}(Z_{s'} - Z_s) \to 0$. As moreover $Z_s \leq Z_{s'}$, we deduce that

$$\mathbb{P}(Z_s \neq Z_{s'}) = \mathbb{P}(Z_{s'} - Z_s \geq 1) \leq \mathbb{E}(Z_{s'} - Z_s) \to 0. \tag{2.4.25}$$

Furthermore, we already know from Theorem 2.4.1 that $\mathrm{d_{TV}}(U(t_{\langle s'\rangle}); \mu_{s'}^{\otimes\Gamma}) \to 0$. But since $|s - s'| \to 0$, we also have $\mathrm{d_{TV}}(U(t_{\langle s'\rangle}), \mu_s^{\otimes\Gamma}) \to 0$.

To conclude, let us compare $\mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s\rangle}) = A)$ with our target $\pi(x)\lambda(A)$, where we set $\lambda(A) = e^{-e^{-s}} e^{-|A|s}/n^{|A|}$, i.e., $\lambda = \mu_s^{\otimes\Gamma}$. Then

$$\begin{aligned}
\left|\mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s\rangle}) = A) - \pi(x)\lambda(A)\right| &\leq \left|\mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s\rangle}) = A) - \mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s'\rangle}) = A)\right| \\
&\quad + \left|\mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s'\rangle}) = A) - \pi(x)\lambda(A)\right| \\
&\leq \mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s'\rangle}) = A; Z_s \neq Z_{s'}) \\
&\quad + \left|\mathbb{P}(U(t_{\langle s'\rangle}) = A)\mathbb{P}(X_{t_{\langle s\rangle}} = x | U(t_{\langle s'\rangle}) = A) - \pi(x)\lambda(A)\right|
\end{aligned}$$

Summing over all x, A, for the first term in the right hand side above, we get

$$\sum_x \sum_A \mathbb{P}(X_{t_{\langle s\rangle}} = x, U(t_{\langle s'\rangle}) = A; Z_s \neq Z_{s'}) = \mathbb{P}(Z_s \neq Z_{s'}),$$

which converges to zero by (2.4.25). The second term, on the other hand, can be written by Theorem 2.4.1 (or more precisely Remark 2.4.1) as

$$\left|\lambda(A)(1 + o(1))\mathbb{P}(X_{t_{\langle s\rangle}} = x | U(t_{\langle s'\rangle}) = A) - \pi(x)\lambda(A)\right|$$

But observe that by choice of s' compared to s, in the interval from $t_{\langle s' \rangle}$ to $t_{\langle s \rangle}$ the walk has had time to mix. More specifically, the lower bound of (2.4.24), $D^2 \lll t_{\langle s \rangle} - t_{\langle s' \rangle} =: t$, implies, through [LP17, Proposition 4.15], (2.2.60), and Corollary 2.2.1, that uniformly over $x, y \in \Gamma$,

$$|p_t(x,y) - 1/n| \leq p_t(o,o) - 1/n \leq \exp\left(-\frac{t - D^2}{dD^2}\right) p_{D^2}(o,o) = o\left(\frac{1}{n}\right), \tag{2.4.26}$$

i.e. that $p_t(x,y) = \pi(y)(1 + o(1))$.

Thus, by averaging over $X_{t_{\langle s' \rangle}}$ we have $\mathbb{P}(X_{t_{\langle s \rangle}} = x | U(t_{\langle s' \rangle}) = A) = \pi(x)(1 + o(1))$, and we conclude that this second term is equal to $\lambda(A)\pi(x)o(1)$. Summing over all x and A, we see that the sum is $o(1)$ and we deduce that

$$\sum_x \sum_A \left|\mathbb{P}(X_{t_{\langle s \rangle}} = x, U(t_{\langle s \rangle}) = A) - \pi(x)\lambda(A)\right| \to 0$$

which is the desired result. □

Lemma 2.4.4. *Let $k \geq 2$ and $A \in \mathcal{A}_*$ of size k, and let $x \in A$. Then the probability that x is the first point of A to be touched by the walk starting from uniformity is $1/k + o(1)$:*

$$\mathbb{P}_\pi(T_A = T_x) = \frac{1}{k} + o(1). \tag{2.4.27}$$

Proof. By going in discrete time it would be possible to prove this via a reversibility argument and some simple estimates. We opt however for the following idea using the skeleton of Section 2.3.3. Let α be the quasi-stationary distribution associated to A, and let

$$p := \mathbb{P}_\alpha(T_A = T_x). \tag{2.4.28}$$

Remark that as $\mathrm{d}_{\mathrm{TV}}(\pi, \alpha) = o(1)$ (either by Aldous–Brown, or more precisely by Theorem 2.5.2), we have $\mathbb{P}_\pi(T_A = T_x) = p + o(1)$.

Moreover, we have by the memoryless property of the hitting time of A under the quasi-stationary distribution, that for every $t > 0$,

$$\begin{aligned} p &= \mathbb{P}_\alpha(T_A = T_x | T_A \leq t)\mathbb{P}(T_A \leq t) + \mathbb{P}_\alpha(T_A = T_x | T_A > t)\mathbb{P}(T_A > t) \\ &= \mathbb{P}_\alpha(T_A = T_x | T_A \leq t)\mathbb{P}(T_A \leq t) + p\mathbb{P}(T_A > t), \end{aligned}$$

so simplifying this identity and doing the cancellations,

$$p = \mathbb{P}_\alpha(T_A = T_x | T_A \leq t) = \frac{\mathbb{P}_\alpha(T_A = T_x, T_A \leq t)}{\mathbb{P}_\alpha(T_A \leq t)}. \tag{2.4.29}$$

From now on, fix t such that $D^2 \lll t \lll \min(n, D^3)$.

For the numerator, let us use again the uniform distribution instead of the quasistatonary distribution. From the Cauchy-Schwarz inequality and Theorem 2.5.2 we have:

$$|\mathbb{P}_\alpha(T_A = T_x, T_A \leq t) - \mathbb{P}_\pi(T_A = T_x, T_A \leq t)| \leq \mathrm{d}_{\mathrm{TV}}(\pi, \alpha) = O(1/n). \tag{2.4.30}$$

To estimate $\mathbb{P}_\pi(T_A = T_x, T_A \leq t)$ we make use of our results in Subsection 2.3.3. Recall the notion of skeleton, and note that under the assumption $A \in \mathcal{A}_*$, the set A has k δ_*-components, which are all singletons. We can write for instance $A = \{a_1, ..., a_k\}$, and we have $E_i = \{a_i \in \gamma\}$, where γ is a path starting from uniformity and of time-length t. In particular, assuming without loss of generality that $x = a_1$, we have a trivial upper bound

$$\mathbb{P}_\pi(T_A = T_{a_1}, T_A \leq t) \leq \mathbb{P}_\pi(T_{a_1} \leq t). \tag{2.4.31}$$

For the lower bound, observe that

$$\begin{aligned}\{T_A = T_{a_1}, T_A \le t\} &\supset \{T_{a_1} \le t \text{ and } T_{a_i} > t \text{ for all } 2 \le i \le k\} \\ &= \{T_{a_1} \le t\} \setminus \bigcup_{2 \le i \le k} \{T_{a_1} \le t \text{ and } T_{a_i} \le t\}.\end{aligned}$$

Hence, as for each $2 \le i \le k$ we have from Proposition 2.3.2 and Corollary 2.2.3 that

$$\mathbb{P}_\pi(T_{a_1} \le t \text{ and } T_{a_i} \le t) = \mathbb{P}(E_1 \cap E_i) \le \mathbb{P}_\pi(T_o \le t) \max_{\substack{x,y \in \Gamma \\ d(x,y) \ge \delta_*}} \mathbb{P}_x(T_y < t) = \mathbb{P}_\pi(T_o \le t) \cdot o(1). \tag{2.4.32}$$

which finally proves that

$$\mathbb{P}_\pi(T_A = T_{a_1}, T_A \le t) = \mathbb{P}_\pi(T_o \le t)(1 + o(1)). \tag{2.4.33}$$

As $\mathbb{P}_\pi(T_o \le t) \gtrsim t/n \ggg 1/n$ (for instance by considering again the quasi-stationary distribution of o), we have also using (2.4.30)

$$\mathbb{P}_\alpha(T_A = T_x, T_A \le t) = \mathbb{P}_\pi(T_o \le t)(1 + o(1)). \tag{2.4.34}$$

We deduce that

$$p = \frac{\mathbb{P}_\pi(T_o \le t)}{\mathbb{P}_\alpha(T_A \le t)}(1 + o(1)). \tag{2.4.35}$$

Note that this obviously does not depend on the choice of x in A, and so must be $1/k + o(1)$. Alternatively, by Corollary 2.2.4 and the fact that π and α are close, the quotient on the right hand side is $(1 + o(1))/q_A$, so it is indeed equal to $1/k + o(1)$. □

Lemma 2.4.5. *Let $A \in \mathcal{A}_*$ of size k, and let y be such that $d(y, A) \ge \delta_*$. Let $x \in A$. Then starting from y, the probability that x is the first point of A to be touched by the walk is $1/k + o(1)$:*

$$\mathbb{P}_y(T_A = T_x) = \frac{1}{k} + o(1). \tag{2.4.36}$$

Proof. This follows simply from the previous lemma (Lemma 2.4.4) if we allow a burn-in period during which the walk can mix but is unlikely to touch A. Let t^* be such that $D^2 \lll t^* \lll \min(n, D^3)$. Note that by Proposition 2.2.3 (see in particular (2.2.56)), $\mathbb{P}_y(T_A \le t^*) \to 0$. Furthermore, since $t_{\text{mix}} \lesssim D^2 \lll t^*$, the distribution of the walk at time t^* is $o(1)$ away (in the total variation sense) from π. Hence

$$\begin{aligned}\mathbb{P}_y(T_A = T_x) &= \mathbb{P}_y(T_A = T_x; T_A \le t^*) + \mathbb{P}_y(T_A = T_x; T_A > t^*) \\ &= o(1) + \mathbb{P}_\pi(T_A = T_x) + o(1) \\ &= 1/k + o(1),\end{aligned}$$

as desired. □

We can now prove the main result of this subsection.

Proof of Theorem 2.4.2. Let $k \ge 2$ and $\varepsilon > 0$. Let $s = s(k, \varepsilon) \in \mathbb{R}$ be such that for n large enough,

$$\mathbb{P}(\tau_k > t_{\langle s \rangle}) \ge 1 - \varepsilon. \tag{2.4.37}$$

Let also $K = K(s, k, \varepsilon) \geq k + 1$ be such that for all sufficiently large n we have that,

$$\mathbb{P}(\tau_K \leq t_{\langle s \rangle}) \geq 1 - \varepsilon. \tag{2.4.38}$$

In particular, with probability at least $1 - 2\varepsilon$, there are between $k + 1$ and K uncovered points at time $t_{\langle s \rangle}$:

$$\mathbb{P}(k + 1 \leq Z_s \leq K) = \mathbb{P}(\tau_K \leq t_{\langle s \rangle} < \tau_k) \geq 1 - 2\varepsilon. \tag{2.4.39}$$

Fix j such that $k + 1 \leq j \leq K$, and condition on the event $Z_s = j$. Let us now describe the evolution of $U(t_{\langle s \rangle})$, up to events of probability $o(1)$, in order to get an approximation of $U(\tau_k)$ in the total variation sense. Conditionally on $\{Z_s = j\}$, we know by Theorem 2.4.1 that $U(t_{\langle s \rangle})$ is a uniformly chosen set of size j, Unif_j, which we may assume is in $\mathcal{A}_*$. Furthermore, by Lemma 2.4.3, the position of the walk at time $t_{\langle s \rangle}$ is uniformly distributed on Γ, independently from $U(t_{\langle s \rangle})$. (Again these descriptions refer in reality to approximation in the total variation sense.) By Lemma 2.4.4, the next point that is removed from $U(t_{\langle s \rangle})$ is therefore uniformly chosen among $U(t_{\langle s \rangle})$. Applying next Lemma 2.4.5 $j - 1 - k$ times successively, from this point onwards, at each successive stage until time τ_k, points are removed uniformly at random. Since $U(t_{\langle s \rangle})$ was uniformly distributed among sets of size j initially, it therefore follows that (still under the conditional law given $\{Z_s = j\}$) that the law of $U(\tau_k)$ is (close to, in the total variation sense) Unif_k. Since this is true for every $k + 1 \leq j \leq K$, we deduce

$$\mathrm{d_{TV}}(U(\tau_k), \mathrm{Unif}_k) \leq 2\varepsilon + o(1).$$

The result follows. □

2.4.3 Convergence in total variation of cover time

We illustrate the results above by strengthening the mode of convergence for the rescaled cover time: namely the distribution ν_n of the random variable $Y_n = \frac{\tau_{\mathrm{cov}}}{\mathbb{E}_\pi T_o} - \log n$, converges in total variation to a standard Gumbel distribution ν:

Theorem 2.4.3. As n tends to infinity, we have

$$\mathrm{d_{TV}}(\nu_n, \nu) \to 0. \tag{2.4.40}$$

Proof. We need several ingredients. First, we saw in Theorem 2.4.2 that for each $k \geq 1$, $U(\tau_k)$ is approximately uniform, in particular, at time τ_1, the walk is with probability $1 - o(1)$ macroscopically far from the last point x_0 to be visited, and hence, with the same arguments as for the convergence of the k last points, the walk with probability $1 - o(1)$ get mixed again before touching x_0. Recalling moreover that the distance in total variation between the uniform distribution and the quasi-stationary distribution associated to x_0 is $o(1)$, we have that

$$\mathrm{d_{TV}}\left(\frac{\tau_0 - \tau_1}{\mathbb{E}_\pi(T_o)}, \mathrm{Exp}(1)\right) \to 0.$$

Let ν_{cont} denote the law of $\frac{\tau_1}{\mathbb{E}_\pi(T_o)} + X$, where X is an independent exponential random variable with mean 1. Note that ν_{cont} is obtained by convolution with an exponential law and so has a 1-Lipschitz density with respect to Lebesgue measure. Furthermore, since $\tau_{\mathrm{cov}} = \tau_0 = \tau_1 + (\tau_0 - \tau_1)$,

$$\mathrm{d_{TV}}\left(\frac{\tau_{\mathrm{cov}}}{\mathbb{E}_\pi(T_o)}, \nu_{\mathrm{cont}}\right) \to 0. \tag{2.4.41}$$

Second, we have already shown that for each fixed $s \in \mathbb{R}$,

$$F_n(s) := \mathbb{P}(Y_n \leq s) = \mathbb{P}(Z_s = 0) \to e^{-e^{-s}} =: F(s). \tag{2.4.42}$$

Let now $\varepsilon > 0$ and let us fix $S = S(\varepsilon)$ large enough such that for n large enough, we have $F_n(S) \geq 1 - \varepsilon$ and $F_n(-S) \leq \varepsilon$. In particular, for such n's, we have

$$\nu_n([-S, S]) = \mathbb{P}(Y_n \in [-S, S]) \geq 1 - 2\varepsilon. \tag{2.4.43}$$

Let f_Y be the density of Y (which is just the density of the Gumbel law) and let f_n denote the density of ν_{cont} after translating by $\log n$. Since $\mathrm{d}_{\text{TV}}(\nu_n, f_n) \to 0$ it suffices to show that

$$\mathrm{d}_{\text{TV}}(f_n, f_Y) = (1/2) \int_{s \in \mathbb{R}} |f_n(s) - f_Y(s)| \mathrm{d}s \to 0. \tag{2.4.44}$$

Observe that f_n is 1-Lipschitz over $[-S, S]$, since it is a convolution of some given law with an exponential law. It is also pointwise bounded at, say $s = 0$ (indeed, since f_n is Lipschitz, it cannot be large at any point without its integral being large, which is not possible by (2.4.41)). It is therefore uniformly equicontinuous, and by the Ascoli–Arzéla theorem has subsequential uniform limits. However, the limit can only be f_Y, again by (2.4.41). Thus f_n converges to f_Y uniformly over $[-S, S]$, and hence also in the L^1 sense over $[-S, S]$. This proves (2.4.44). □

2.5 Refining the Aldous–Brown approximation

In this section we provide a refinement of Aldous and Brown's result about hitting time from stationarity ([AB92], our equation (2.2.67)). We believe this refinement is of interest in its own right. Indeed, in recent years there has been much interest in understanding the quasi-stationary distribution, and the rate of convergence to it for the chain conditioned on not hitting the corresponding set, e.g., [DM09, DM15, DHESC21]. The results below should be useful in that context too.

We first explain why such a refinement is needed here. In the examples studied in §2.7 we have that $t_{\text{rel}} \log n \asymp \mathbb{E}_\pi[T_A]$. As always, denote the quasi-stationary distribution of A by α (suppressing the dependence of α on A from the notation). Then by (2.2.67) we get that

$$\mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A] = O(\mathbb{E}_\pi[T_A]/\log n).$$

Since we consider $\mathbb{P}_\pi[T_A > t]$ for times $t \asymp \mathbb{E}_\pi[T_A] \log n$, in order to replace $\mathbb{E}_\alpha[T_A]$ with $\mathbb{E}_\pi[T_A]$ in the term $\exp(-t/\mathbb{E}_\alpha[T_A])$ in (2.2.67), it is crucial for us that $\mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A] = o(\mathbb{E}_\alpha[T_A]/\log n)$. Indeed, for $t = (1 \pm o(1))\mathbb{E}_\pi[T_A] \log n$, if $\mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A] \gtrsim \mathbb{E}_\alpha[T_A]/\log n$, then

$$\frac{\exp(-t/\mathbb{E}_\alpha[T_A])}{\exp(-t/\mathbb{E}_\pi[T_A])} = \exp\left(\frac{t\,(\mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A])}{\mathbb{E}_\alpha[T_A]\mathbb{E}_\pi[T_A]}\right) \gtrsim 1,$$

whereas we need this ratio to be $1 + o(1)$ for our purposes. However, when $t_{\text{rel}} \log n \asymp \mathbb{E}_\pi[T_A]$ (2.2.67) only implies that $\mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A] = O(\mathbb{E}_\alpha[T_A]/\log n)$. A refinement of (2.2.67) is therefore required.

Consider a reversible Markov chain with a finite state space V. Denote the transition matrix of the chain by P and its stationary distribution by π. Let $\varnothing \neq A \subsetneq V$. Throughout this section we assume without repeating this assumption that for all $a, b \in A^c$ we have that $\mathbb{P}_a[T_b < T_A] > 0$. (The next remark explains how can this condition be relaxed.)

For any non-empty $B \subset V$, we write π_B for the distribution of π conditioned on B. That is, $\pi_B(\cdot) := \frac{\pi(\cdot)1_{\{\cdot \in B\}}}{\pi(B)}$.

Remark 2.5.1. If the above condition fails, one can consider equivalent classes $B_1, \dots, B_\ell$ of the equivalence relation (on B) $a \sim b$ iff $\mathbb{P}_a[T_b < T_A] > 0$. By reversibility this relation is indeed symmetric, whereas transitivity follows by the strong Markov property. In this case

$$\mathbb{P}_{\pi_{A^c}}[T_A > t] = \sum_{i=1}^{\ell} \frac{\pi(B_i)}{\pi(A^c)} \mathbb{P}_{\pi_{B_i}}[T_{B_i^c} > t].$$

The theory we develop below can be applied to $\mathbb{P}_{\pi_{B_i}}[T_{B_i^c} > t]$ for all i.

Before stating our refinement of Aldous-Brown we need to introduce some additional notation. We denote by $\alpha/\pi : V \to \mathbb{R}$ the Radon-Nikodym derivative of α w.r.t. π. That is $(\alpha/\pi)(a) := \alpha(a)/\pi(a)$ for all $a \in V$ (with the convention that $\alpha(a) = 0$ if $a \in A$). For $f, g : V \to \mathbb{R}$ we define the inner-product $\langle f, g \rangle_\pi := \mathbb{E}_\pi[fg]$ (where $\mathbb{E}_\pi[h] := \sum_{x \in V} \pi(x)h(x)$), L^2 norm $\|f\|_2 := \sqrt{\langle f, f \rangle_\pi}$ and variance w.r.t. π, $\mathrm{Var}_\pi f := \|f - \mathbb{E}_\pi[f]\|_2^2$. Lastly, we denote the ε total variation mixing time of this chain by $t_{\mathrm{mix}}(\varepsilon)$ and its relaxation-time by t_{rel}. Below we consider the continuous-time setup. We first state all of our results for this section.

Theorem 2.5.1 (Refinement of Aldous-Brown)**.** In the above setup and notation we have that for all $t \geq 0$

$$0 \leq \mathbb{P}_\pi[T_A > t] - \frac{1}{\|\alpha/\pi\|_2^2} \exp\left(-\frac{t}{\mathbb{E}_\alpha[T_A]}\right) \leq \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)\right) \exp\left(-\frac{t}{t_{\mathrm{rel}}}\right). \tag{2.5.1}$$

Consequently, integrating over $t > 0$,

$$0 \leq \mathbb{E}_\pi[T_A] - \frac{1}{\|\alpha/\pi\|_2^2} \mathbb{E}_\alpha[T_A] \leq \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)\right) t_{\mathrm{rel}}, \tag{2.5.2}$$

and in particular

$$0 \leq \mathbb{E}_\alpha[T_A] - \mathbb{E}_\pi[T_A] \leq \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2}\right) \mathbb{E}_\alpha[T_A]. \tag{2.5.3}$$

The L^2 norm (with respect to π) of a signed measure ρ is defined as $\|\rho\|_{2,\pi} := \|\rho/\pi\|_2$. Observe that $\|\alpha/\pi\|_2^2 - 1 = \|\alpha/\pi - 1\|_2^2$ is exactly $\|\alpha - \pi\|_{2,\pi}^2$. We also observe that by Cauchy–Schwarz

$$\|\alpha/\pi\|_2^2 \pi(B) \geq \mathbb{E}_\pi\left[(\alpha/\pi)\mathbf{1}_B\right]^2 = \|\alpha/\pi\|_1^2 = 1.$$

Hence $\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) = \pi(B) - \frac{1}{\|\alpha/\pi\|_2^2} \geq 0$.

Remark 2.5.2. If instead of $\mathbb{E}_\pi[T_A]$ or $\mathbb{P}_\pi[T_A > t]$ we consider $\mathbb{E}_{\pi_B}[T_A]$ or $\mathbb{P}_{\pi_B}[T_A > t]$, respectively, then the term $\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)$ in (2.5.1) and (2.5.2) is replaced by $\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2}$.

To make use of Theorem 2.5.1 we require an upper bound on $\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)$. The next theorem gives such a bound whenever $t_{\mathrm{rel}} / \mathbb{E}_\alpha[T_A]$ is small.

Denote by P_B the transition matrix of the chain killed upon hitting A (that is, $P_B(x, y) = P(x, y)\mathbf{1}\{x, y \in B\}$). Observe that by the reversibility of P we have that P_B is self-adjoint with respect to the inner-product $\langle f, g \rangle_\pi = \mathbb{E}_\pi[fg]$ on the space

$$C_0(B) := \{f \in \mathbb{R}^V : f(x) = 0 \text{ for all } a \in A\}.$$

Hence P_B has $m := |B|$ real eigenvalues $1 - \lambda = \gamma_1 > \gamma_2 \geq \cdots \geq \gamma_m \geq -|\gamma_1|$, where the last inequality follows by the Perron-Frobenius Theorem, which also asserts that $\gamma_m = -|\gamma_1|$ if and only if the restriction of P_B to B has period 2, and that $\gamma_2 < \gamma_1$, due to our assumption that $\min_{a,b\in B} \mathbb{P}_a[T_b < T_A] > 0$, which is equivalent to the restriction of P_B to B being irreducible. A fairly standard application of the interlacing eigenvalues theorem (see Lemma 2.5.2) yields that

$$1 - \gamma_2 \geq 1/\, t_{\mathrm{rel}}\,. \tag{2.5.4}$$

Denote $c_i := \mathbb{E}_\pi[f_i] = \sum_x \pi(x) f_i(x)$ and $\lambda_i := 1 - \gamma_i$.

We shall show later in this section that the law of T_A under $\mathbb{P}_\pi$ can be written as a mixture $\pi(A)\delta_0 + \sum_{i=1}^m c_i^2 \nu_i$, where for all $i, j \in [m]$ we have that ν_i is the Exponential distribution with parameter λ_i, where for some $f_1, \ldots, f_m : V \to \mathbb{R}$ in $C_0(B)$ we have for all $i, j \in [m]$ that $c_i := \mathbb{E}_\pi[f_i]$, $\mathbb{E}_\pi[f_i f_j] = \mathbf{1}\{i = j\}$, where $\mathbb{E}_\pi[g] := \sum_{x\in V} \pi(x) g(x)$, and that $P_B f_i = (1 - \lambda_i) f_i$. Moreover, we may take $f_1 := \frac{\alpha/\pi}{\|\alpha/\pi\|_2}$, where α is the quasi-stationary distribution on B, which satisfies $\alpha P_B = (1 - \lambda_1)\alpha$. In particular,

$$\begin{gathered}\sum_{i=1}^m c_i^2 = \mathbb{P}_\pi[T_A > 0] = \pi(B) \quad \text{and} \quad c_1^2/\lambda_1 \leq \mathbb{E}_\pi[T_A] = \sum_{i=1}^m c_i^2/\lambda_i \leq \pi(B)/\lambda_1, \\ c_1^2 = 1 - \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} \quad \text{and} \quad p := \sum_{j=2}^m c_j^2 = \pi(B) - c_1^2 = \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A).\end{gathered} \tag{2.5.5}$$

Finally, let us define t_{med} by

$$\sum_{i=2}^m c_i^2 (1 - e^{-\lambda_i t_{\mathrm{med}}})^2 = \frac{1}{2} \sum_{i=2}^m c_i^2.$$

By (2.5.4) we have that $t_{\mathrm{med}} \leq \log(1 + \frac{1}{\sqrt{2}})/\lambda_2 \leq \log(1 + \frac{1}{\sqrt{2}})\, t_{\mathrm{rel}} \leq \frac{1}{\sqrt{2}}\, t_{\mathrm{rel}}$.

Theorem 2.5.2 (Bounding $\|\alpha/\pi\|_2^2 - 1$)**.** For all $\varnothing \neq A \subsetneq V$ we have that

$$\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) = 2 \sum_x \pi(x) \left(\mathbb{P}_x[T_A \leq t_{\mathrm{med}}]\right)^2 - 2\frac{1}{\|\alpha/\pi\|_2^2}(1 - e^{-\lambda_1 t_{\mathrm{med}}})^2. \tag{2.5.6}$$

Note that when $\pi(A)$ is small (2.5.6) offers an improvement over the estimate $\frac{\|\alpha/\pi\|_2^2-1}{\|\alpha/\pi\|_2^2} \leq t_{\mathrm{rel}}/\mathbb{E}_\alpha[T_A]$ of Aldous and Brown. Indeed, if $t_{\mathrm{rel}} = \varepsilon \mathbb{E}_\alpha[T_A]$ for $\varepsilon \leq 1/2$ then by (2.5.1)

$$\sum_x \pi(x) \left(\mathbb{P}_x[T_A \leq t_{\mathrm{rel}}]\right)^2 \leq \mathbb{P}_\pi[T_A \leq t_{\mathrm{rel}}] \leq 2\, t_{\mathrm{rel}}/\mathbb{E}_\alpha[T_A] = 2\varepsilon.$$

Of course the first inequality can be very wasteful. As we shall see in the next theorem, in many cases we have that $\sum_x \pi(x) \left(\mathbb{P}_x[T_A \leq t_{\mathrm{rel}}]\right)^2 \asymp (t_{\mathrm{rel}}/\mathbb{E}_\alpha[T_A])^2$. We now specialize Theorems 2.5.1 and 2.5.2 to the case of vertex-transitive graphs. For the applications in this paper we are mostly interested in the "two-dimensional case" $m = 2$ (in the next theorem m no longer stands for the cardinality of B). In that case the main term of the right hand side is of order $\lesssim 1/R^2$ which corresponds to $\lesssim 1/f(n)^2$ in the notations of Theorem 2.1.1.

Theorem 2.5.3. Let $G = (V, E)$ be a connected vertex-transitive graph of degree d with n vertices, $o \in V$, and $\alpha = \alpha_o$. Then

$$\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \lesssim_{d,k} \frac{D^4}{(\mathbb{E}_\pi T_0)^2} \left(1 + \frac{n}{D^4} \int_0^D \frac{s^3 \mathrm{d}s}{V(s)}\right). \tag{2.5.7}$$

In particular, if $D^m R = n$ where $1 \le R < D$ and $2 \le m \in \mathbb{N}$, we have

$$\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \lesssim_{d,k} \begin{cases} \frac{D^4}{(\mathbb{E}_\pi[T_o])^2} & m = 1,2 \\ \frac{D^4}{(\mathbb{E}_\pi[T_o])^2}\left(1 + \frac{R \log R}{D}\right) & m = 3 \\ \frac{D^4}{(\mathbb{E}_\pi[T_o])^2}\left(R + \log(\frac{D}{R})\right) & m = 4 \\ 1/n & m \ge 5 \end{cases}. \tag{2.5.8}$$

2.5.1 Preparation towards the proofs

We now present some background on quasi-stationary distributions and their relation to hitting times. Let $\varnothing \neq A \subsetneq V$. Denote $B := A^c$. We begin the analysis by recalling (e.g. [AF, Ch. 4]) that the quasi-stationary distribution $\alpha = \alpha_B$ corresponding to the set B satisfies that the first hitting time of A under $\mathbb{P}_\alpha$ has in continuous-time an Exponential distribution whose parameter λ is the smallest eigenvalue of $I_B - P_B$, where $P_B(x,y) := P(x,y)\mathbf{1}\{x,y \in B\}$ and $I_B(a,b) := \mathbf{1}\{a = b \in B\}$. Note that $\lambda > 0$ since P_B is sub-stochastic. The matrix P_B is the transition matrix of the chain killed upon hitting A, and $P_B - I_B$ is the Markov generator corresponding to the rate one continuous-time version of this killed chain. As we shall work in continuous-time, we introduce the notation $P_{t,B} := e^{-t(I_B - P_B)} = e^{-t}\sum_{\ell=0}^{\infty}(P_B)^\ell \frac{t^\ell}{\ell!}$.

We now prove the above claim. It is an immediate consequence of the Perron-Frobenius Theorem that there exists a distribution α such that $\alpha P_B = (1-\lambda)\alpha$ from which it follows that for all $t \ge 0$ we have that

$$\alpha P_{t,B} = \alpha e^{-t(I_B - P_B)} = e^{-t}\sum_{\ell=0}^{\infty}\frac{[(1-\lambda)t]^\ell}{\ell!}\alpha = e^{-\lambda t}\alpha,$$

and so $\mathbb{P}_\alpha[T_A > t] = \alpha P_{t,B}(B) = e^{-\lambda t}$, where we used the fact that for all x, $\alpha P_{t,B}(x) := \sum_b \alpha(b) P_{t,B}(b,x) = \mathbb{P}_\alpha[X_t = x, T_A > t]$. It follows that starting from α the law of T_A is exponential with mean $1/\lambda$, and so $1/\lambda = \mathbb{E}_\alpha[T_A]$.

The identity $\alpha P_{t,B} = e^{-\lambda t}\alpha$ also implies that for all $x \in B$ and all $t \ge 0$

$$\mathbb{P}_\alpha[X_t = x \mid T_A > t] = e^{\lambda t}\mathbb{P}_\alpha[X_t = x, T_A > t] = e^{\lambda t}\alpha P_{t,B}(x) = \alpha(x).$$

The last equation justifies the name "quasi-stationary distribution".

We now present the derivation of the well-known fact that for continuous-time reversible chains, the hitting time of a set A starting from the stationary distribution π is a mixture of exponentials. In other words, this law is completely monotone. The proof we present for this well-known fact shall play a role in our refinement of Aldous-Brown.

It is natural to identify $P_{t,B}$ as an operator with its extension to $C_0(B)$ defined as follows: for $f : V \to \mathbb{R}$ supported on B

$$(P_{t,B}f)(a) := \sum_b P_{t,B}(a,b)f(b) = \mathbb{E}_a[f(X_t)\mathbf{1}\{T_A > t\}].$$

Denote $m = |B|$. Let $f_1, \ldots, f_m$ be an orthonormal (w.r.t. the inner-product $\langle \cdot, \cdot \rangle_\pi$) basis of $C_0(B)$, such that $P_B f_i = \gamma_i f_i$ for all i. By reversibility we can take $f_1 = \frac{\alpha/\pi}{\|\alpha/\pi\|_2}$. Then $1_B = \sum_{i=1}^m c_i f_i$, where $c_i := \mathbb{E}_\pi[f_i]$ and $P_{t,B}1_B = \sum_{i=1}^m c_i e^{-t(1-\gamma_i)} f_i$. Hence, by orthonormality, we have that

$$\mathbb{P}_\pi[T_A > t] = \langle P_{t,B}1_B, 1_B \rangle_\pi = \sum_{i=1}^m c_i^2 e^{-t(1-\gamma_i)} = \frac{1}{\|\alpha/\pi\|_2^2} e^{-t/\mathbb{E}_\alpha[T_A]} + \sum_{i=2}^m c_i^2 e^{-t(1-\gamma_i)} \tag{2.5.9}$$

(similarly, in discrete time we have that $\mathbb{P}_\pi[T_A > \ell] = \langle P_B^\ell 1_B, 1_B\rangle_\pi = \sum_{i=1}^m c_i^2 \gamma_i^\ell$). Taking $t = 0$ we see that $\sum_{i=1}^m c_i^2 = \pi(B)$. Hence $\sum_{i=2}^m c_i^2 = \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)$, and so

$$\frac{1}{\|\alpha/\pi\|_2^2} e^{-t/\mathbb{E}_\alpha[T_A]} \le \mathbb{P}_\pi[T_A > t] \le \frac{1}{\|\alpha/\pi\|_2^2} e^{-t/\mathbb{E}_\alpha[T_A]} + \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)\right) e^{-t(1-\gamma_2)}. \tag{2.5.10}$$

Since $\mathbb{P}_\alpha[T_A > t] = e^{-t(1-\gamma_1)} = e^{-t/\mathbb{E}_\alpha(T_A)}$, the last inequality can also be written as

$$0 \le \mathbb{P}_\pi[T_A > t] - \frac{1}{\|\alpha/\pi\|_2^2}\mathbb{P}_\alpha[T_A > t] \le \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)\right) e^{-t(1-\gamma_2)}. \tag{2.5.11}$$

For the sake of completeness, we recall the following lemma of Aldous and Brown [AB92, Lemma 10], which in conjunction with Lemma 2.5.2 and (2.5.10) immediately implies that for all $t \ge 0$, $\mathbb{P}_\pi[T_A > t] \ge \left(1 - \frac{t_{\mathrm{rel}}}{\mathbb{E}_\alpha[T_A]}\right) e^{-t/\mathbb{E}_\alpha[T_A]}$, which is the non-trivial direction of (2.2.67). We note that while the next lemma is taken from [AB92], the derivation of the (2.2.67) in [AB92] is considerably more complicated than the one presented here.

Lemma 2.5.1. *In the above setup and notation we have that*

$$\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} \le \frac{t_{\mathrm{rel}}}{\mathbb{E}_\alpha[T_A]}.$$

Proof. Since $f_1 = \frac{\alpha/\pi}{\|\alpha/\pi\|_2}$ is supported on B we get that $((I-P)f_1(x))\, f_1(x) = \lambda f_1(x)^2$ for all x. This, together with the fact that $\|f_1\|_2 = 1$ gives that

$$\frac{1}{\mathbb{E}_\alpha[T_A]} = \lambda = \frac{\langle (I-P)f_1, f_1\rangle_\pi}{\|f_1\|_2^2} = \mathrm{Var}_\pi f_1 \frac{\langle (I-P)f_1, f_1\rangle_\pi}{\mathrm{Var}_\pi f_1}. \tag{2.5.12}$$

Using the extremal characterization of the relaxation-time (see, e.g., Theorem 3.1 in [Ber16]), this implies that $\frac{\langle (I-P)f_1, f_1\rangle_\pi}{\mathrm{Var}_\pi f_1} \ge 1/\,t_{\mathrm{rel}}$ and observing that $\mathrm{Var}_\pi f_1 = \frac{\|\alpha/\pi - 1\|_2^2}{\|\alpha/\pi\|_2^2} = \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2}$ conclude the proof. □

We recall the construction of the auxiliary Markov chain in which A is collapsed into a single point. Its state space is $B \cup \{A\}$, where $B = A^c$. Its transitions are given by $K(x,y) = P(x,y)$, $K(x,\{A\}) = P(x,A)$, $K(\{A\},x) = \sum_{a\in A} \pi_A(a)P(a,x)$ for $x, y \in B$ and $K(\{A\},\{A\}) = \sum_{a\in A} \pi_A(a)P(a,A)$. This is a reversible chain w.r.t. the distribution $\hat{\pi}$ given by $\hat{\pi}(\{A\}) = \pi(A)$ and $\hat{\pi}(x) = \pi(x)$ for all $x \in B$. Denote the relaxation time of K by $t_{\mathrm{rel}}(K)$.

Lemma 2.5.2 (Interlacing eigenvalues theorem)**.** *We have that*

$$\frac{1}{1-\gamma_2} \le t_{\mathrm{rel}}(K) \le t_{\mathrm{rel}}\,. \tag{2.5.13}$$

Proof. Denote the second largest eigenvalue of K by $\hat{\lambda}$. By the extremal characterization of the second largest eigenvalue (e.g. [LP17, Remark 13.8]) we get that

$$\frac{1}{t_{\mathrm{rel}}} = \min_{f:\mathrm{Var}_\pi f \neq 0} \frac{\langle (I-P)f, f\rangle_\pi}{\mathrm{Var}_\pi f} \le \min_{f:\mathrm{Var}_\pi f\neq 0, f(a)=f(b) \text{ for all } a,b\in A} \frac{\langle (I-P)f, f\rangle_\pi}{\mathrm{Var}_\pi f}$$

Observe that for f such that $f(a) = f(b)$ for all $a, b \in A$ we have that $\frac{\langle (I-P)f,f\rangle_\pi}{\mathrm{Var}_\pi f} = \frac{\langle (I-K)\hat{f},\hat{f}\rangle_\pi}{\mathrm{Var}_{\hat{\pi}} \hat{f}}$, where $\hat{f}(x) = f(x)$ for all $x \in B$ and $\hat{f}(\{A\}) = f(a)$ for $a \in A$. Hence the r.h.s. of the last display equals $1 - \hat{\lambda} = 1/\,t_{\mathrm{rel}}(K)$. To conclude the proof it remains to show that $\hat{\lambda} \ge \gamma_2$. Observe that P_B is obtained from K by deleting the row and column corresponding to $\{A\}$. It follows from the interlacing eigenvalues theorem that indeed $\hat{\lambda} \ge \gamma_2$, as desired. □

2.5.2 Proofs

Proof of Theorem 2.5.1. This follows at once by combining (2.5.10) and Lemma 2.5.2. $\square$

Proof of Theorem 2.5.2. Denote $t := t_{\mathrm{med}}$. Let $f_1 = \frac{\alpha/\pi}{\|\alpha/\pi\|_2}, f_2, \ldots, f_m$ be an orthonormal basis of $C_0(B)$ such that $P_{B,s} f_i = e^{-\lambda_i s} f_i$ for all $i \in [m]$ and $s \geq 0$. We see that for all $x, y \in B$

$$\begin{aligned}\mathbb{P}_x[X_t = y, T_A > t] &= \langle P_{t,B} 1_y, \frac{1_x}{\pi(x)} \rangle_\pi = \sum_{i=1}^m \pi(y) f_i(y) f_i(x) e^{-t(1-\gamma_i)} \\ &= \alpha(y) \frac{\alpha(x)/\pi(x)}{\|\alpha/\pi\|_2^2} e^{-t(1-\gamma_1)} + \sum_{i=2}^m \pi(y) f_i(y) f_i(x) e^{-t(1-\gamma_i)}.\end{aligned}$$

Summing over y (and recalling that $c_i := \mathbb{E}_\pi f_i$) gives

$$\mathbb{P}_x[T_A > t] = \frac{\alpha(x)/\pi(x)}{\|\alpha/\pi\|_2^2} e^{-t(1-\gamma_1)} + \sum_{i=2}^m c_i f_i(x) e^{-t(1-\gamma_i)}. \tag{2.5.14}$$

Let $h(x) := \mathbb{P}_x[T_A \leq t] - c_1 f_1(x)(1 - e^{-\lambda_1 t})$. By (2.5.14) we have that $h(x) = \sum_{i=2}^m c_i f_i(x)(1 - e^{-\lambda_i t})$ and $\mathbb{P}_x[T_A \leq t] = \sum_{i=1}^m c_i f_i(x)(1 - e^{-\lambda_i t})$, where $c_i := \mathbb{E}_\pi[f_i]$. Hence, by orthonormality of $f_1, \ldots, f_m$, the definition of t_{med} and the fact that $\sum_{i=2}^m c_i^2 = \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A)$ (as explained above (2.5.10)), we get that

$$\sum_x \pi(x) h^2(x) = \sum_{i=2}^m c_i^2 (1 - e^{-\lambda_i t})^2 = \frac{1}{2} \sum_{i=2}^m c_i^2 = \frac{1}{2} \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \right),$$

$$\sum_x \pi(x) (\mathbb{P}_x[T_A \leq t])^2 = \sum_{i=1}^m c_i^2 (1 - e^{-\lambda_i t})^2.$$

Hence,

$$\begin{aligned}\sum_x \pi(x) (\mathbb{P}_x[T_A \leq t])^2 &= \sum_x \pi(x) h^2(x) + c_1^2 (1 - e^{-\lambda_1 t})^2 \\ &= \frac{1}{2} \left(\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \right) + \frac{1}{\|\alpha/\pi\|_2^2} (1 - e^{-\lambda_1 t})^2. \qquad \square\end{aligned} \tag{2.5.15}$$

Proof of Theorem 2.5.3. We first prove (2.5.8). Let A be of size k and $o \in V$. Denote $t := t_{\mathrm{mid}} \leq \frac{1}{\sqrt{2}} t_{\mathrm{rel}}$. By Theorem 2.5.2 our goal is to show that $\frac{\sum_x \pi(x)(\mathbb{P}_x[T_A \leq t])^2}{C(m)(dk)^2}$ is upper bounded by the r.h.s. of (2.5.8). Let $\tau \sim \mathrm{Exp}(\frac{1}{t_1})$ be independent of the random walk, where $t_1 := 2\, t_{\mathrm{rel}}$. Then $\mathbb{P}(\tau \geq t)^2 \geq \exp(-\frac{1}{\sqrt{2}}) \geq \frac{2}{5}$ and so by transitivity (used to argue that $\mathbb{P}_\pi[T_a \leq \tau]$ is independent of a),

$$\begin{aligned}\frac{2}{5} \sum_x \pi(x) (\mathbb{P}_x[T_A \leq t])^2 &\leq \sum_x \pi(x) (\mathbb{P}_x[T_A \leq \tau])^2 \leq \sum_x \pi(x) \left(\sum_{a \in A} \mathbb{P}_x[T_a \leq \tau] \right)^2 \\ &\leq k^2 \sum_x \pi(x) (\mathbb{P}_x[T_o \leq \tau])^2,\end{aligned} \tag{2.5.16}$$

by the Cauchy–Schwarz inequality. Let $N = L_o(\tau)$ be the time the chain spent at o by time τ. Note that $\mathbb{P}_x[X_s = y, \tau > s] = P_s(x,y) e^{-\frac{s}{t_1}}$. Hence $\mathrm{E}_x[N] = \int_0^\infty P_s(x,o) e^{-\frac{s}{t_1}} \mathrm{d}s$. By the memory property of the exponential distribution $\kappa := \mathrm{E}_x[N \mid N > 0] = \mathrm{E}_o[N] = \int_0^\infty P_s(o,o)e^{-\frac{s}{t_1}}\mathrm{d}s$. Using

$$\mathbb{P}_x[T_o \le \tau] = \frac{\mathrm{E}_x[N]}{\mathrm{E}_x[N \mid N > 0]},$$

and reversibility, and the change of variables $s + r = u$, we get that

$$\begin{aligned}
\sum_x \pi(x)(\mathbb{P}_x[T_o \le \tau])^2 &= \kappa^{-2} \sum_x \pi(x) \left(\int_0^\infty P_s(x,o)e^{-\frac{s}{t_1}}\mathrm{d}s \right)^2 \\
&= \kappa^{-2} \sum_x \pi(o) \left(\int_0^\infty P_s(o,x)e^{-\frac{s}{t_1}}\mathrm{d}s \right) \left(\int_0^\infty P_r(x,o)e^{-\frac{r}{t_1}}\mathrm{d}r \right) \qquad (2.5.17)\\
&= (\kappa n)^{-2} \int_0^\infty nuP_u(o,o)e^{-\frac{u}{t_1}}\mathrm{d}u \\
&=: J.
\end{aligned}$$

Let $f_1, \ldots, f_n$ be an orthonormal basis of $\mathbb{R}^V$ with respect to $\langle \cdot, \cdot \rangle_\pi$, such that $P_s f_i = e^{-\beta_i s} f_i$ for all i. We now argue that

$$\int_0^\infty nsP_s(o,o)e^{-\frac{s}{t_1}}\mathrm{d}s \le t_1^2 + \sum_{i=2}^n \beta_i^{-2}. \qquad (2.5.18)$$

Indeed, $\int_0^\infty nsP_s(o,o)e^{-\frac{s}{t_1}}\mathrm{d}s = t_1^2 + \int_0^\infty ns\left(P_s(o,o) - \pi(o)\right)e^{-\frac{s}{t_1}}\mathrm{d}s$ and so by transitivity, we get that

$$\begin{aligned}
\int_0^\infty nsP_s(o,o)e^{-\frac{s}{t_1}}\mathrm{d}s - t_1^2 &= \int_0^\infty \sum_x s\left(P_s(x,x) - \pi(x)\right)e^{-\frac{s}{t_1}}\mathrm{d}s = \sum_{i=2}^n \int_0^\infty se^{-\beta_i s - \frac{s}{t_1}}\mathrm{d}s \\
&= \sum_{i=2}^n \frac{1}{(\beta_i + 1/t_1)^2} \le \sum_{i=2}^n \beta_i^{-2}.
\end{aligned}$$

Similarly, using the fact that $\beta_i + \frac{1}{t_1} \le \frac{3}{2}\beta_i$ for $i \ge 2$ (by the choice of t_1)

$$\begin{aligned}
n\kappa - t_1 &= \int_0^\infty \sum_x \left(P_s(x,x) - \pi(x)\right)e^{-\frac{s}{t_1}}\mathrm{d}s \\
&= \sum_{i=1}^n \int_0^\infty e^{-s(\beta_i + 1/t_1)}\mathrm{d}s \ge \sum_{i=2}^n \frac{1}{\beta_i + \frac{1}{t_1}} \ge \frac{2}{3} \sum_{i=2}^n \frac{1}{\beta_i} = \frac{2}{3}\mathbb{E}_\pi[T_o],
\end{aligned}$$

where we used the eigentime identity $\sum_{i=2}^n \beta_i^{-1} = \sum_x \pi(x)\mathbb{E}_\pi[T_x] = \mathbb{E}_\pi[T_o]$ (see e.g. [AF, p. 117]), where the last equality holds by transitivity.

We now bound $\frac{1}{n}\sum_{i=2}^n \beta_i^{-2}$ from above. Let $\mu = (1/n)\sum_{i=1}^n \delta_{\beta_i}$ be the spectral measure and recall that $\beta_1 = 0$. It will be convenient to also consider $\nu = \mu - \delta_{\{0\}}/n$. Let $X \sim \mu$. Then, since $\mu((0,r]) = \mu([0,r]) - 1/n = \nu([0,r])$,

$$\frac{1}{n}\sum_{i=2}^n \beta_i^{-2} = \mathbb{E}[X^{-2}\mathbf{1}\{X > 0\}] = \int_0^{\beta_2^{-2}} \mathbb{P}[s \le X^{-2} < \infty]\mathrm{d}s \le 1/4 + \int_{1/4}^{\beta_2^{-2}} \mathbb{P}[s \le X^{-2} < \infty]\mathrm{d}s$$

$$= \frac{1}{4} + \int_{\beta_2}^2 \frac{2}{r^3}\mathbb{P}[0 < X \le r]\mathrm{d}r \quad (\text{change of variables } r = \frac{1}{\sqrt{s}})$$

$$= \frac{1}{4} + \int_{\beta_2}^{2} \frac{2}{r^3}\nu([0,r])\mathrm{d}r \le \frac{1}{4} + \int_{\frac{1}{dD^2}}^{2} \frac{2}{r^3}\nu([0,r])\mathrm{d}r,$$

where we used the fact that by transitivity $\beta_2 \ge \frac{1}{dD^2}$ (see e.g. [LOG18, Theorem 1.7]).

Denote the number of vertices in a ball of radius $\lceil \rho \rceil$ by $V(\rho)$. Theorem 1.7 in [LOG18] (with their paramater α set to $\alpha = 1/2$) asserts that

$$\nu([0,r]) \le \frac{4}{V(\sqrt{(2dr)^{-1}})}.$$

Substituting this bound, and using the change of variables $s = \sqrt{(2dr)^{-1}}$ yields

$$\int_{\frac{1}{dD^2}}^{2} \frac{2}{r^3}\nu([0,r])\mathrm{d}r \le 64d^2 \int_{\frac{1}{\sqrt{d}}}^{D/\sqrt{2}} \frac{s^3 \mathrm{d}s}{V(s)}.$$

Putting everything together we get that

$$J \le \frac{t_1^2 + 1/4 + 64d^2 n \int_{\frac{1}{\sqrt{d}}}^{D/\sqrt{2}} \frac{s^3\mathrm{d}s}{V(s)}}{\left(t_1 + \frac{2}{3}\mathbb{E}_\pi[T_0]\right)^2}.$$

Recall that from (2.5.15), (2.5.16), and (2.5.17), $\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \le 5k^2 J$. Since $t_1 \le 2dD^2$ (and $D \ge 1$), we deduce that

$$\begin{aligned}
\frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} - \pi(A) \le 5k^2 J &\le \frac{45k^2}{4} \frac{5d^2D^4 + 64d^2 n \int_{\frac{1}{\sqrt{d}}}^{D/\sqrt{2}} \frac{s^3\mathrm{d}s}{V(s)}}{(\mathbb{E}_\pi T_0)^2} \\
&\le 720(kd)^2 \frac{D^4 + n\int_0^D \frac{s^3\mathrm{d}s}{V(s)}}{(\mathbb{E}_\pi T_0)^2} \\
&= 720(kd)^2 \frac{D^4}{(\mathbb{E}_\pi T_0)^2}\left(1 + \frac{n}{D^4}\int_0^D \frac{s^3\mathrm{d}s}{V(s)}\right).
\end{aligned} \tag{2.5.19}$$

The proof of (2.5.8) is concluded by substituting the bounds on $V(s)$ of Tessera and Tointon [TT21, Corollary 1.5]

$$V(s) \ge \begin{cases} c(m)s^{m+1} & s \le R \\ c(m)Rs^m & s > R \end{cases}, \tag{2.5.20}$$

and using for $m \ge 5$ the fact that $\mathbb{E}_\pi[T_o] = \sum_{i=2}^n \frac{1}{\beta_i} \ge \frac{1}{2}(n-1) \ge \frac{n}{4}$. Note that replacing $c(m)$ by $\min_{1\le m\le 5} c(m)$ removes the dependence in m of the constant. $\square$

2.5.3 Alternative proof for the low dimensional case

In Section 2.3, we proved that if $f(n) \ggg \log n$, then the factorial moments $\mathbb{E}_\pi Z^{\downarrow k}$ converge to e^{-ks}. A difficulty was that the approximation (2.2.76) of q_A did not hold if $f(n) \lesssim (\log n)^2$. To prove that the result holds under the sharp diameter condition $n \ggg D^2 \log n$, we had to approximate directly the expected hitting times $\mathbb{E}_\pi T_A$ using auxiliary chains. The reason for this is that the approximation appearing in the results of Aldous–Brown, which we recalled in (2.2.67), is not precise enough. We will show as a consequence of Theorem 2.5.3 that (2.2.76) holds even in the sharp case $n \ggg D^2 \log n$, and we will give a simpler proof of Proposition 2.3.9. We will first prove an approximation under the more general assumption $n \gtrsim D^2$, which will be useful in Section 2.7 where we will have $n \asymp D^2 \log n$, and then specify it to the case $n \ggg D^2 \log n$. The first proposition we prove refines Lemma 2.2.5 when t is not too large.

Proposition 2.5.4. Assume that $1 \leq f(n) < D$. Then we have uniformly over $A \subset \Gamma$ of size k and $t \geq 0$ that

$$0 \leq \mathbb{P}_\alpha[T_A > t] - \mathbb{P}_\pi[T_A > t] = O\left(\frac{1}{f(n)^2}\right), \tag{2.5.21}$$

where the implicit constant depends on d and k.

Proof. We have, from Theorem 2.5.3, noting that $f(n) = R$ and that we are in the case $m = 2$,

$$\frac{\frac{\|\alpha/\pi\|_2^2-1}{\|\alpha/\pi\|_2^2} - \pi(A)}{C(2)(dk)^2} \leq \frac{d^2D^4 + n(D^2 - f(n)^2)}{f(n)(\mathbb{E}_\pi[T_o])^2} \leq \frac{nD^2(1 + d^2/f(n))}{f(n)(\mathbb{E}_\pi[T_o])^2} \leq \frac{(1+d^2)\, n^2}{f(n)^2(\mathbb{E}_\pi T_o)^2}. \tag{2.5.22}$$

We deduce from Theorem 2.5.1, that (uniformly) for every $t \geq 0$,

$$0 \leq \mathbb{P}_\pi[T_A > t] - \frac{1}{\|\alpha/\pi\|_2^2}\mathbb{P}_\alpha[T_A > t] \lesssim C(d,k)\left(\frac{n}{f(n)\mathbb{E}_\pi T_o}\right)^2 \exp\left(-\frac{t}{t_{\text{rel}}}\right). \tag{2.5.23}$$

Note also that from (2.5.22) (for the right hand side), we have

$$0 \leq \frac{\|\alpha/\pi\|_2^2 - 1}{\|\alpha/\pi\|_2^2} \lesssim \left(\frac{n}{f(n)\mathbb{E}_\pi T_o}\right)^2, \tag{2.5.24}$$

so

$$\frac{1}{\|\alpha/\pi\|_2^2} = 1 - O\left(\frac{n}{f(n)\mathbb{E}_\pi T_o}\right)^2. \tag{2.5.25}$$

We deduce finally, as $e^{-t/\,t_{\text{rel}}} \leq 1$ and using that $\mathbb{E}_\pi T_o \geq n/(8e)$ (see the proof of the lower bound of Proposition 2.2.5) that uniformly over every $t \geq 0$,

$$0 \leq \mathbb{P}_\alpha[T_A > t] - \mathbb{P}_\pi[T_A > t] \lesssim \left(\frac{n}{f(n)\mathbb{E}_\pi T_o}\right)^2 \lesssim \frac{1}{f(n)^2}. \tag{2.5.26}$$

where the implicit constant depends on d and k. $\square$

Corollary 2.5.1. *Assume that $1 \leq f(n) < D$. Then we have at time $t = \mathbb{E}_\alpha T_A/f(n)$, uniformly over $A \subset \Gamma$ of size k, that*

$$q_A = \frac{\mathbb{P}_\pi(T_A \leq t)}{\mathbb{P}_\pi(T_o \leq t)}\left(1 + O\left(\frac{1}{f(n)}\right)\right). \tag{2.5.27}$$

In particular, if $f(n) \ggg \log n$, we have at time $t = D^2$, that

$$q_A = \frac{\mathbb{P}_\pi(T_A \leq t)}{\mathbb{P}_\pi(T_o \leq t)}\left(1 + o\left(\frac{1}{\log n}\right)\right). \tag{2.5.28}$$

Moreover, the last estimation still holds if we take $t = n/\sqrt{f(n)\log n}$ instead of $t = D^2$.

Proof. We know from Proposition 2.5.8 that $\mathbb{E}_\alpha T_A = \mathbb{E}_\pi T_A \left(1 + O(1/f(n)^2)\right)$, and from the proof of Proposition 2.2.5 that $n/(8ek) \leq \mathbb{E}_\pi T_A$, so $\mathbb{E}_\alpha T_A \gtrsim n$. Consequently, we have for $t \leq n$ that

$$\mathbb{P}_\alpha[T_A > t] = \exp(-t/\mathbb{E}_\alpha T_A) \gtrsim 1. \tag{2.5.29}$$

We deduce from Proposition 2.5.4 that for $t \leq n$,

$$\mathbb{P}_\pi \left(T_A > t\right) = \exp\left(-\frac{t}{\mathbb{E}_\alpha T_A}\right)\left(1 + O\left(\frac{1}{f(n)^2}\right)\right).$$

Assume now moreover that $t \lll n$, then we can follow the proof of Proposition 2.2.10. The only difference is that we do not upper bound $\mathbb{E}_\alpha T_A$ by n as they are not necessarily of the same order of magnitude. First, we have

$$\mathbb{E}_\pi T_A = \frac{t}{\mathbb{P}_\pi\left(T_A \leq t\right)}\left(1 + O\left(\frac{\mathbb{E}_\alpha T_A}{tf(n)^2}\right) + O\left(\frac{t}{\mathbb{E}_\alpha(T_A)}\right)\right). \tag{2.5.30}$$

Taking quotients, we get

$$q_A = \frac{\mathbb{P}_\pi(T_A \leq t)}{\mathbb{P}_\pi(T_o \leq t)}\left(1 + O\left(\frac{\mathbb{E}_\alpha T_A}{tf(n)^2}\right) + O\left(\frac{t}{\mathbb{E}_\alpha(T_A)}\right)\right), \tag{2.5.31}$$

and we finally have at time $t = \mathbb{E}_\alpha T_A/f(n)$,

$$q_A = \frac{\mathbb{P}_\pi(T_A \leq t)}{\mathbb{P}_\pi(T_o \leq t)}\left(1 + O\left(\frac{1}{f(n)}\right)\right). \tag{2.5.32}$$

For the last point, if $f(n) \ggg \log n$, we know from Proposition 2.2.5 (and Proposition 2.2.4) that $\mathbb{E}_\alpha T_A \asymp n$. At time $t = n/\sqrt{f(n)\log n}$, we have both $n/(tf(n)^2) = o(1/\log n)$ and $t/n = o(1/\log n)$ so the error in (2.5.30) is $o(1/\log n)$. Taking quotients, we get the desired result. $\square$

This allows us to strengthen the results of Section 2.3.3. In particular, the second part of Proposition 2.3.3 becomes:

Proposition 2.5.5. Assume that $n \ggg D^2 \log n$ and $\delta < \operatorname{mindist}(A)$. Then at time $t = n/\sqrt{f(n)\log n}$, we have

$$q_A - k \gtrsim -\max_{\substack{x,y\in\Gamma \\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t) + o\left(\frac{1}{\log n}\right).$$

We can now give a simpler proof of Proposition 2.3.9.

Alternative proof of Proposition 2.3.9: Let δ such that $\delta_{\text{meso}} \leq \delta \leq D/2$ and $A \subset \Gamma$ such that $|A| = k$, and $\operatorname{mindist}(A) \geq \delta$. Then from the intermediate step of (2.2.54), and Lemma 2.2.2 we have immediately, at $t = n/\sqrt{f(n)\log n}$,

$$\begin{aligned}\max_{\substack{x,y\in\Gamma \\ d(x,y)=\delta}} \mathbb{P}_x(T_y < t) \lesssim \mathbb{E}_x L_y(t) &\lesssim \frac{1}{\delta} + \frac{\log(D/\delta)}{f(n)} + \frac{t}{n} + \frac{t}{D^3} + \frac{1}{f(n)} \\ &= \frac{\log(D/\delta)}{f(n)} + o\left(\frac{1}{\log n}\right). \qquad \square\end{aligned}$$

2.6 When the diameter condition fails

2.6.1 Green function estimates

In this section $\Gamma = (V, E)$ is a vertex-transitive graph of degree d. Denote the transition matrix of simple random walk on Γ by P and the time t transition probabilities for the rate 1 continuous-time simple random walk on Γ by $P_t = e^{-t(I-P)}$ (so that we have $P_t(x,y) = p_t(x,y)$). Recall that for $x, y \in V$ and $t \geq 0$,

$$h_t(x,y) = p_t(x,y) - \frac{1}{n}. \tag{2.6.1}$$

We start by adapting a gradient inequality, due to Diaconis and Saloff-Coste [DSC94] for Cayley graphs, to vertex-transitive graphs. The main difference is that we need to use the mass transport principle.

Proposition 2.6.1. Let $\Gamma = (V, E)$ be a finite vertex-transitive graph of degree d. Then for all $s, t \geq 0$ and $o, y, z \in V$ we have that

$$|h_{t+s}(o,y) - h_{t+s}(o,z)| \leq d(y,z)\sqrt{\frac{d}{2es}}h_t(o,o). \tag{2.6.2}$$

Proof. By the triangle inequality it suffices to prove the inequality when y and z are neighbours. Let then $y, z \in V$ such that $y \sim z$. First observe that

$$\begin{aligned} h_{t+s}(o,y) - h_{t+s}(o,z) &= p_{t+s}(o,y) - p_{t+s}(o,z) \\ &= \sum_{w \in V} p_{t/2}(o,w)(p_{t/2+s}(w,y) - p_{t/2+s}(w,z)) \\ &= \sum_{w \in V} h_{t/2}(o,w)(h_{t/2+s}(w,y) - h_{t/2+s}(w,z)) \end{aligned}$$

because, by reversibility, $\sum_{w \in V}(h_{t/2+s}(w,y) - h_{t/2+s}(w,z)) = 0$. By reversibility again, we have

$$\sum_{w \in V} h_{t/2}(o,w)^2 = \left(\sum_{w \in V} p_{t/2}(o,w)p_{t/2}(w,o)\right) - \frac{1}{n} = p_t(o,o) - \frac{1}{n} = h_t(o,o). \tag{2.6.3}$$

It follows from the triangle inequality and the Cauchy–Schwarz inequality that

$$\begin{aligned} |h_{t+s}(o,y) - h_{t+s}(o,z)|^2 &\leq \left(\sum_w h_{t/2}(o,w)^2\right)\left(\sum_w |h_{t/2+s}(w,y) - h_{t/2+s}(w,y)|^2\right) \\ &= h_t(o,o)\sum_w |p_{t/2+s}(y,w) - p_{t/2+s}(z,w)|^2 \\ &\leq h_t(o,o)\sum_w \sum_{y' \,:\, y' \sim y} |p_{t/2+s}(y,w) - p_{t/2+s}(y',w)|^2 \\ &= h_t(o,o)\sum_w F(y,w), \end{aligned} \tag{2.6.4}$$

where

$$F : (a,b) \mapsto \sum_{x \,:\, x \sim a} \left|p_{t/2+s}(a,b) - p_{t/2+s}(x,b)\right|^2.$$

Observe that for any $\gamma \in \mathrm{Aut}(\Gamma)$, and $a, b \in V$, we have $F(\gamma(a), \gamma(b)) = F(a,b)$. Moreover $\mathrm{Aut}(\Gamma)$ is a discrete group of automorphisms, so by [LP16b][Corollary 8.9] is unimodular. Since

Γ is vertex-transitive, $\mathrm{Aut}(\Gamma)$ is (by definition) transitive, so we can apply the mass transport principle. By [LP16b][Equation (8.4)], we hence have, for all $y \in V$, that

$$\sum_{w \in V} F(y,w) = \sum_{w \in V} F(w,y) = \sum_{w \in V} \sum_{w' \,:\, w' \sim w} |p_{t/2+s}(w,y) - p_{t/2+s}(w',y)|^2. \tag{2.6.5}$$

Denote $P_{\mathrm{L}} := \frac{1}{2}(I+P)$. For $f,g : V \to \mathbb{R}$ denote $\langle f,g\rangle = \pi(o)\sum_{v\in V} f(v)g(v)$ and $\|f\|_2^2 := \langle f,f\rangle$. Since $I - P_{\mathrm{L}}$ is self-adjoint and $\langle (I-P_{\mathrm{L}})f,f\rangle \geq 0$ for all $f : V \to \mathbb{R}$, we may consider $K := \sqrt{I-P_{\mathrm{L}}}$ which is a self-adjoint operator satisfying $K^2 = I - P_{\mathrm{L}}$, and $\langle (I-P_{\mathrm{L}})f,g\rangle = \langle Kf,Kg\rangle$ for all $f,g : V \to \mathbb{R}$. Noting that $P_s = e^{-2s(I-P_{\mathrm{L}})}$ we have that $KP_s = q(I-P_{\mathrm{L}})$, where $q : [0,1] \to [0,1]$ is defined by $q(u) = \sqrt{u}e^{-2su}$, notation which we also extend (slightly abusing notation) to matrices. Since the spectral measure of $I - P_{\mathrm{L}}$ is supported on $[0,1]$, it is standard that $\|KP_s\|_2^2 = \|q(I-P_{\mathrm{L}})\|_2^2 \leq \max_{u\in[0,1]} q(u)^2 = \frac{1}{4es}$, where $\|KP_s\|_2 := \sup_{f\in\mathbb{R}^V : \|f\|_2=1} \|KP_sf\|_2$ is the operator norm of KP_s.

For $r \geq 0$ and $w \in V$, denote $f_r(w) = h_r(w,y)$. By reversibility, the sum on the right hand side of (2.6.5) equals $2d$ times (see e.g., [LP17, Lemma 13.6])

$$\frac{1}{\pi(o)}\langle (I-P_{\mathrm{L}})\, f_{t/2+s}, f_{t/2+s}\rangle = \frac{1}{\pi(o)}\langle K^2 P_s f_{t/2}, P_s f_{t/2}\rangle \leq \frac{1}{\pi(o)}\|KP_s\|_2^2 \|f_{t/2}\|_2^2 \leq \frac{h_t(o,o)}{4es},$$

where in the last inequality we used that $\|KP_s\|_2^2 \leq \frac{1}{4es}$ and (2.6.3).

It follows that the right hand side of (2.6.5) is less than $\frac{dh_t(o,o)}{2es}$, which in conjunction with (2.6.4) concludes the proof. □

We now prove that at times t proportional to D^2 the distribution of the walk is still far from being uniform.

Proposition 2.6.2. Assume that $n \leq D^3$. There exists a constant $0 < a = a(d) < 1$ such that for all $t \leq aD^2$,

$$nh_t(o,o) \geq 1/2. \tag{2.6.6}$$

Proof. Recall the off-diagonal heat kernel bound (2.2.48). Note that the constant C in (2.2.47), as well as the constant c and the implicit constant in (2.2.48) (call it C' for this proof), depend only on the graph degree, and the growth of the function g defined in (2.2.47). We can hence take the same constants for all (connected) finite vertex-transitive graphs of degree d such that $D^3 \geq n$.

First assume that $D^2 \leq n < D^3$, and let x such that $d(o,x) = D$. We have for every $D \leq t \leq D^2$,

$$p_t(o,x) \leq C'\frac{1}{g(t)}\exp\left(-c\frac{D^2}{t}\right) = C'\frac{C}{\min(f(n)t, t^{3/2})}\exp\left(-c\frac{D^2}{t}\right). \tag{2.6.7}$$

Let $0 < a < 1$. At time $t = aD^2$ we have

$$\min(f(n)t, t^{3/2}) = \min(an, a^{3/2}D^3) \geq a^{3/2}n. \tag{2.6.8}$$

It follows that

$$p_t(o,x) \leq \frac{CC'}{a^{3/2}}\frac{1}{n}\exp\left(-c\frac{1}{a}\right). \tag{2.6.9}$$

Fixing $a = a(c,C,C')$ small enough, we hence have at time $t = aD^2$ that $np_t(o,x) \leq 1/2$, i.e.

$$nh_t(o,x) \leq -\frac{1}{2}. \tag{2.6.10}$$

Finally, we have $h_t(o,o) = \max_{z,w\in\Gamma} |h_t(z,w)|$ by [LP17, Proposition 4.15] and vertex-transitivity. We conclude that

$$nh_t(o,o) \geq \frac{1}{2}. \tag{2.6.11}$$

The proof for $D \leq n < D^2$ is similar.

□

We define the Green function between two points $x, y \in V$ by

$$G(x,y) = \int_0^\infty h_t(x,y)\mathrm{d}t. \tag{2.6.12}$$

The following proposition collects some useful identities involving Green functions.

Proposition 2.6.3. Let Γ be any finite connected vertex-transitive graph. We have the following identities.

1. Fix $o \in \Gamma$. We have
$$\mathbb{E}_\pi T_o = nG(o,o) \geq n - (1 + \log n). \tag{2.6.13}$$

2. For every $x, y \in \Gamma$,
$$\mathbb{E}_x T_y = n\left(G(o,o) - G(x,y)\right). \tag{2.6.14}$$

3. For any subset set $A = \{x,y\}$ of size 2 of Γ,
$$q_A = \frac{2}{1 + \frac{G(x,y)}{G(o,o)}}. \tag{2.6.15}$$

Proof. The equalities in the first two points are stated in [AF] in a more general framework, as Lemma 2.11 and Lemma 2.12, respectively, and Section 2.2.3 explains why they also hold in continuous time. We moreover have, lower bounding the probability to be at o by the probability to have never jumped:

$$G(o,o) \geq \int_0^{\log n} \left[e^{-t} - \frac{1}{n}\right] dt = 1 - \frac{1+\log n}{n}, \tag{2.6.16}$$

proving the ineqiality from the first point.

For the third point, let us denote with tildes the collapsed chain where A is reduced to a point which we denote $\tilde{A}$ (which, if x and y are neighbours, has a loop at $\tilde{A}$), and jumps at rate 1. Then by [AF], Lemma 2.11, and transitivity of the original graph Γ, we have

$$\mathbb{E}_\pi T_A = \mathbb{E}_{\tilde{\pi}} T_{\tilde{A}} = \frac{1}{\tilde{\pi}(\tilde{A})}\tilde{G}(\tilde{A},\tilde{A}) = \frac{2}{n}(G(o,o) + G(x,y)), \tag{2.6.17}$$

using the same idea as in the proof of Lemma 2.3.4 (the difference being that the set A here has size two, so the argument is much simpler, and we are in continuous time with the walk jumping out of every vertex – including $\tilde{A}$ – at rate 1).

Therefore,

$$q_A = \frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\pi T_A} = \frac{nG(o,o)}{G(o,o) + G(x,y)} = \frac{2}{1 + \frac{G(x,y)}{G(o,o)}}, \tag{2.6.18}$$

as desired. □

We first show that the tail of the integral defining the Green function above decreases exponentially fast (this is similar to arguments already used in Section 2.3.7 although this was written in discrete time).

Lemma 2.6.1. *There exist constants c_1, c_2 depending on d such that uniformly over all real $b \geq 1$, we have, uniformly over all finite (connected) vertex-transitive graphs of degree d such that $n \leq D^5$,*

$$\int_{bD^2}^{\infty} h_t(o,o)dt \leq c_1 e^{-c_2 b} \frac{D^2}{n}. \tag{2.6.19}$$

Proof. By spectral estimates, we have for all $t, s \geq 0$.

$$h_{t+s}(o,o) \leq e^{-s/\,t_{\mathrm{rel}}} h_t(o,o). \tag{2.6.20}$$

We deduce, splitting the integral into parts of length D^2, that

$$\sum_{i=1}^{\infty} \int_{iD^2}^{(i+1)D^2} h_t(o,o)dt \leq \left(\int_{D^2}^{2D^2} h_t(o,o)dt \right) \sum_{j=0}^{\infty} e^{-jD^2/\,t_{\mathrm{rel}}}. \tag{2.6.21}$$

By (2.2.8) and since $t \mapsto h_t(o,o)$ is decreasing, there exists a constant $c_0 = c_0(d)$ such that

$$\int_{D^2}^{2D^2} h_t(o,o)dt \leq h_{D^2}(o,o) \cdot D^2 \leq c_0 \frac{D^2}{n}. \tag{2.6.22}$$

Finally, by Lemma 2.2.1, $t_{\mathrm{rel}} \leq dD^2$, hence $e^{-D^2/\,t_{\mathrm{rel}}} \leq e^{-1/d}$. We deduce

$$\sum_{j=0}^{\infty} e^{-jD^2/\,t_{\mathrm{rel}}} \leq \frac{1}{1-e^{-1/d}}. \tag{2.6.23}$$

We have proved the claim for integers b. The claim for real b follows immediately, since $b \mapsto \int_{bD^2}^{\infty} h_t(o,o)dt$ is decreasing. $\square$

Proposition 2.6.4. Let (Γ) be a sequence of finite (connected) vertex-transitive graphs of fixed degree d, such that $|\Gamma| = n \leq D^3$. There exist positive constants $\eta = \eta(d), C = C(d)$ such that for all $x \in B(o, \eta D)$,

$$G(o,x) \geq C \frac{D^2}{n}. \tag{2.6.24}$$

Proof. Let $\varepsilon > 0$, to be chosen later. By Lemma 2.6.1, we can chose $b = b(d,\varepsilon) \geq 1$ large enough such that

$$\int_{bD^2}^{\infty} h_t(o,o)dt \leq \varepsilon \frac{D^2}{n}. \tag{2.6.25}$$

We hence have

$$G(o,x) = \int_0^{\varepsilon D^2} h_t(o,x)dt + \int_{\varepsilon D^2}^{bD^2} h_t(o,x)dt + \int_{bD^2}^{\infty} h_t(o,x)dt.$$

Since for every t, $h_t(o,x) \geq -1/n$, and $h_t(o,x) \geq -h_t(o,o)$, we deduce that

$$G(o,x) \geq \int_{\varepsilon D^2}^{bD^2} h_t(o,x)dt - 2\varepsilon \frac{D^2}{n}.$$

By Proposition 2.6.1, we have

$$\int_{\varepsilon D^2}^{bD^2} h_t(o,x)dt \geq \int_{\varepsilon D^2}^{bD^2} h_t(o,o)dt - d(o,x)\int_{\varepsilon D^2}^{bD^2} \sqrt{\frac{d}{et}} h_{t/2}(o,o)dt. \tag{2.6.26}$$

By Proposition 2.6.2, we have, assuming without loss of generality that $\varepsilon \leq a/2$, that

$$\int_{\varepsilon D^2}^{bD^2} h_t(o,o)dt \geq \int_{aD^2/2}^{aD^2} h_t(o,o)dt \geq \frac{a}{4}\frac{D^2}{n}. \tag{2.6.27}$$

Finally, from (2.2.13) there exists a constant $C = C(d,\varepsilon)$ such that $h_{\varepsilon D^2/2} \leq C(d,\varepsilon)/n$, so we have, setting $C' = Cb\sqrt{\frac{d}{e\varepsilon}}$, for every $x \in B(o, \frac{\varepsilon}{C'}D)$, that

$$\begin{aligned} d(o,x)\int_{\varepsilon D^2}^{bD^2} \sqrt{\frac{d}{et}} h_{t/2}(o,o)dt &\leq d(o,x)(b-\varepsilon)D^2\sqrt{\frac{d}{e\varepsilon D^2}} h_{\varepsilon D^2/2}(o,o) \\ &\leq C'\frac{d(o,x)}{D}\frac{D^2}{n} \leq \varepsilon\frac{D^2}{n}. \end{aligned} \tag{2.6.28}$$

The arguments above show that taking $\varepsilon = a/24$, we have, for some $\eta > 0$ (depending only on d), that for every $x \in B(o, \eta D)$,

$$G(o,x) \geq \left(\frac{a}{4} - 3\varepsilon\right)\frac{D^2}{n} \geq \frac{a}{8}\frac{D^2}{n}, \tag{2.6.29}$$

which concludes the proof. □

Lemma 2.6.2. *Let* (Γ) *be a sequence of finite (connected) vertex-transitive graphs of fixed degree* d*, such that* $|\Gamma| = n \leq D^3$*. There exists a constant* $C = C(d) > 0$ *such that for all* $x \in \Gamma$*,*

$$G(o,x) \geq -C\frac{D^2}{n}. \tag{2.6.30}$$

Proof. We have

$$G(o,y) = \int_0^\infty h_t(o,x)dt \geq \int_0^{D^2}\left(-\frac{1}{n}\right)dt - \int_{D^2}^\infty h_t(o,o)dt. \tag{2.6.31}$$

Applying Lemma 2.6.1 to the second integral on the right hand side concludes the proof. □

Let us set, for every $x \in \Gamma$ and $c > 0$,

$$S_{o,c} := \left\{z \in \Gamma : G(o,z) \leq -c\frac{D^2}{n}\right\}. \tag{2.6.32}$$

Corollary 2.6.1. *There exist constants* $c = c(d) > 0$ *and* $c' = c'(d) > 0$ *such that for all finite (connected) vertex-transitive graphs of degree* d *such that* $n = |\Gamma| \leq D^3$*, we have*

$$|S_{o,c}| \geq c'n. \tag{2.6.33}$$

Proof. Let us call for this proof C_1 the constant from Proposition 2.6.4 and C_2 the constant from Lemma 2.6.2. Also, by [TT20][Proposition 5.6], there is a constant $C_3 = C_3(d, \eta)$ such that $V(\eta D) \geq C_3 n$. We hence have

$$\begin{aligned} 0 = \sum_{x\in\Gamma} G(o,x) &= \sum_{\substack{x\in\Gamma \\ G(o,x)\geq 0}} G(o,x) + \sum_{\substack{x\in\Gamma \\ 0>G(o,x)>-\frac{C_1C_3}{2}\frac{D^2}{n}}} G(o,x) + \sum_{\substack{x\in\Gamma \\ -\frac{C_1C_3}{2}\frac{D^2}{n}\geq G(o,x)}} G(o,x) \\ &=: A_1 + A_2 + A_3. \end{aligned}$$

The first term can be lower bounded using Proposition 2.6.4:

$$A_1 \geq \sum_{x\in B(o,\eta D)} G(o,x) \geq C_1 \frac{D^2}{n} V(\eta D) \geq C_1 C_3 D^2. \tag{2.6.34}$$

Since the second sum is over at most n terms, we have the raw bound

$$A_2 \geq -\frac{C_1C_3}{2} D^2. \tag{2.6.35}$$

Finally, since by Lemma 2.6.2 for every x, $G(o,x) \geq -C_2 \frac{D^2}{n}$,

$$A_3 \geq -|S_{o,c}|\, C_2 \frac{D^2}{n} \tag{2.6.36}$$

where $c = C_1C_3/2$.

All in all, dividing by D^2, we have proved that

$$0 \geq C_1C_3 - \frac{C_1C_3}{2} - \frac{C_2}{n}|S_{o,c}| = c - \frac{C_2}{n}|S_{o,c}|, \tag{2.6.37}$$

i.e., setting $c := \frac{C_1C_3}{2}$ and $c' := c/C_2 = \frac{C_1C_3}{2C_2}$, that

$$|S_{o,c}| \geq c' n, \tag{2.6.38}$$

as desired. $\square$

Roughly speaking, one should think of the set $S_{o,c}$ as the set of points which are "far away" from o. However, it turns out that the variations of the Green function are *macroscopically continuous*, and hence that the set $S_{o,c}$ contains a ball of size $\asymp n$. This is the content of Proposition 2.6.5 and Proposition 2.6.6.

Proposition 2.6.5. Let (Γ) be a sequence of finite (connected) vertex-transitive graphs of fixed degree d, such that $|\Gamma| = n \leq D^3$. Let $0 < \rho < 1$. For every $\varepsilon > 0$, there exists a constant $\delta = \delta(\rho, \varepsilon, d) > 0$ such that for every $x, y, z \in \Gamma$ such that $d(o,x) \geq \rho D$, $d(x,y) \leq \delta D$, and $d(o,z) \leq \delta D$, we have

$$|G(o,x) - G(z,y)| \leq \varepsilon \frac{D^2}{n}. \tag{2.6.39}$$

Proof. Let $\varepsilon > 0$. We first assume that $z = o$ and $D^2 \leq n \leq D^3$. Let $a, b > 0$ to be fixed, and let $x, y \in B(o, \rho D)^c$. Denote for this proof, for $0 < t_1 < t_2 \leq \infty$, $I(t_1,t_2) := \int_{t_1}^{t_2} |h_t(o,x) - h_t(o,y)| dt$. By the triangle inequality,

$$|G(o,x) - G(o,y)| \leq I(o,D) + I(D,aD^2) + I(aD^2,bD^2) + I(bD^2,\infty). \tag{2.6.40}$$

By the Carne-Varopoulos inequality, we have $I(o, D) = o(D^2/n)$. Now observe that for all t, $|h_t(o, x) - h_t(o, y)| = |p_t(o, x) - p_t(o, y)| \leq \max(p_t(o, x), p_t(o, y))$. To bound $p_t(o, x)$ (and $p_t(o, y)$) we proceed as in the proof of Proposition 2.6.2 to get (2.6.9). (Now $d(o, x) \geq \rho D$ instead of having $d(o, x) = D$.) Let c as in (2.6.9). Observing that the bound on $p_t(o, x)$ in the integral is maximised at $t = aD^2$, we have

$$I(D, aD^2) \lesssim \int_D^{aD^2} \frac{1}{a^{3/2}} \frac{1}{n} \exp\left(-c\frac{\rho^2}{a}\right) dt \leq \frac{e^{-c\rho^2/a}}{a^{1/2}} \frac{D^2}{n}, \tag{2.6.41}$$

which is less than $\varepsilon D^2/n$ if we fix $a > 0$ small enough. Let us also fix $b > 0$ large enough so that (2.6.25) holds. Then by the triangle inequality (and since for all t, $h_t(o, o) = \max_x |h_t(o, x)|$), we have $I(bD^2, \infty) \leq 2\varepsilon D^2/n$. Finally, by Proposition 2.6.1 with $s = aD^2/2$ and making a change of variables, we have

$$I(aD^2, bD^2) \leq \sqrt{d/e}\frac{d(x, y)}{D} \int_{aD^2/2}^{(b-a/2)D^2} h_t(o, o)dt \lesssim_{a,d} \frac{d(x, y)}{D} \frac{D^2}{n}, \tag{2.6.42}$$

so for $\delta > 0$ fixed small enough, if $d(x, y) \leq \delta D$, we also have $I(aD^2, bD^2) \leq \varepsilon D^2/n$. Putting everything together, we have proved that for $a, \delta > 0$ fixed as above and x, y such that $d(x, y) \leq \delta D$, we have $|G(o, x) - G(o, y)| \leq (4\varepsilon + o(1))\frac{D^2}{n}$. We can bound similarly $|G(o, y) - G(z, y)|$, and hence by the triangle inequality we have $|G(o, x) - G(z, y)| \leq (8\varepsilon + o(1))\frac{D^2}{n}$, concluding the proof when $D^2 \leq n \leq D^3$. The proof for $D \leq n \leq D^2$ is analog. □

Proposition 2.6.6. There exist constants $c, \rho > 0$ (depending on d) such that for all finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \leq D^3$, there exist two (disjoint) balls S_1 and S_2 of radius ρD such that for every $x \in S_1$ and $y \in S_2$,

$$G(x, y) \leq -c\frac{D^2}{n}. \tag{2.6.43}$$

In particular, for such c, ρ, $S_{o,c}$ contains a ball of radius ρD.

Proof. By Corollary 2.6.1, there exists $c > 0$ and $x \in \Gamma$ such that $G(o, x) \leq -cD^2/n$. Therefore, $d(o, x) \geq \eta D$ where η is as in Proposition 2.6.4. Therefore, by Proposition 2.6.5 (where the role of ρ there is played by η here) we can find $\delta > 0$ such that for every $y \in B(x, \delta D)$, and $z \in B(o, \delta D)$

$$|G(z, y) - G(o, x)| \leq \frac{c}{2}\frac{D^2}{n}. \tag{2.6.44}$$

It follows that for such y, z, we have $G(z, y) \leq -\frac{c}{2}\frac{D^2}{n}$, concluding the proof. □

We will also need the following bound on $\mathrm{t_{hit}}$.

Proposition 2.6.7. Let $C > 0$. Let (Γ) be a collection of finite (connected) vertex-transitive of fixed degree d, such that $n = |\Gamma| \leq CD^2 \log n$. Then

$$t_{\langle \mathrm{hit} \rangle} \lesssim_{d,C} D^2 \log n. \tag{2.6.45}$$

Proof. Recall that we have $t_{\langle \mathrm{hit} \rangle} = nG(o, o)$ from Lemma 2.11 in [AF]. We split the proof into two cases, depending on the diameter of Γ.

First, if $D^2 \le n \le CD^2 \log n$, we write $n = D^2 R$. Integrating the bound on $p_t(o,o)$ from Proposition 2.2.1, and bounding the tail of the integral with Proposition 2.6.1, we have

$$G(o,o) = \int_0^\infty h_t(o,o)dt \lesssim_d 1 + \frac{1}{R} + \frac{\log(D/R)}{R} + \frac{D^2}{n} \lesssim_{d,C} \frac{\log n}{R} = \frac{1}{n} D^2 \log n, \tag{2.6.46}$$

as desired.

Suppose now that $D \le n \le D^2$, and let H such that $n = DH$. Proceeding as above, we get

$$G(o,o) \lesssim_d 1 + \log H + \frac{D}{H} + \frac{D^2}{n} \lesssim \log n + \frac{D}{H}. \tag{2.6.47}$$

We hence have

$$nG(o,o) \lesssim DH \log n + D^2 \lesssim D^2 \log n, \tag{2.6.48}$$

which concludes the proof. □

2.6.2 Proof that the diameter condition is necessary in Theorem 2.1.1 and Theorem 2.1.2

In this section, we prove that when the diameter condition (DC) does not hold, then the renormalised cover time $\frac{\tau_{\text{cov}}}{t_{\text{hit}}} - \log|\Gamma|$ does not have asymptotic Gumbel fluctuations. We first prove that Proposition 2.2.3 still holds when the average hitting time is replaced by the quasistationary hitting time.

Proposition 2.6.8. Uniformly over all finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \le D^5$, we have

$$t_{\text{hit}} - \mathbb{E}_\alpha T_o \asymp_d D^2, \tag{2.6.49}$$

where α is the quasistationary distribution associated to o.

To prove Proposition 2.6.8 we will need a few lemmas. For $1 \le r \le D$, set $B_r := B(o,r)$, and $\beta(r) = 1/\mathbb{E}_{\alpha_{B_r^c}} T_{B_r^c}$.

Lemma 2.6.3. *For every $2 \le \ell < D$ and $t \ge 0$, we have, writing $r := \lfloor \ell/2 \rfloor$, that*

$$\mathbb{P}_o[T_{B(o,\ell)^c} > t] \ge \max_{v \in B(o,r)} \mathbb{P}_v[T_{B(o,r)^c} > t] \ge e^{-t\beta(r)}, \tag{2.6.50}$$

$$P_t(o,o) \ge \frac{\mathbb{P}_o[X_t \in B(o,\ell)]}{V(\ell)} \ge \frac{\mathbb{P}_o[T_{B(o,\ell)^c} > t]}{V(\ell)} \ge \frac{e^{-t\beta(r)}}{V(\ell)}, \tag{2.6.51}$$

$$\beta(\ell) \le \frac{4V(\ell)}{V(r)} \ell^{-2}. \tag{2.6.52}$$

Proof. Let $v \in B(o,r)$. The first inequality in (2.6.50) follows by noting that if $X_0 = v$ then $d(X_t, X_0) \le 2r \le \ell$ for all $t < T_{B(o,r)^c}$ and so $T_{B(o,r)^c} \le T_{B(v,\ell)^c}$. It follows by transitivity that $\mathbb{P}_v[T_{B(o,r)^c} > t] \le \mathbb{P}_v[T_{B(v,\ell)^c} > t] = \mathbb{P}_o[T_{B(o,\ell)^c} > t]$ for all $t \ge 0$, as desired.

Let $A = B(o,r)^c$, and $\alpha = \alpha_A$ be the quasi-stationary distribution associated to A. The second inequality in (2.6.50) follows from the fact that

$$\max_{v \in B(o,r)} \mathbb{P}_v[T_{B(o,r)^c} > t] \ge \mathbb{P}_{\alpha_r}[T_{B(o,r)^c} > t] = e^{-t\beta(r)}.$$

The first inequality in (2.6.51) follows from the fact that by transitivity $p_t(o,o) = \max_{x \in \Gamma} p_t(o,x)$, while the second is trivial and the third is exactly (2.6.50).

We now prove (2.6.52). For a distribution ν on V and $g, g' : V \to \mathbb{R}$ we write $\langle g, g' \rangle_\nu := \mathbb{E}_\nu[gg'] = \sum_{x \in V} \nu(x) g(x) g'(x)$. Let $\mathcal{C}_\ell$ be the collection of all $g : V \to \mathbb{R}$ whose support $\{x \in V : g(x) \neq 0\}$ is non-empty and contained in $B(o, \ell)$, and denote by π_ℓ the uniform distribution on $B(o, \ell)$. Observe that for all $g \in \mathcal{C}_\ell$ we have that

$$\begin{aligned} \langle (I - P_{B(o,\ell)}) g, g \rangle_{\pi_\ell} &= \langle (I - P) g, g \rangle_{\pi_\ell} = \langle (I - P) g, g \rangle_\pi / \pi(B(o, \ell)) \\ &= \frac{1}{2\pi(B(o,\ell))} \mathbb{E}_\pi \left[(g(X_0) - g(X_1))^2 \right] \\ &= \mathbb{E}_{\pi_\ell} \left[(g(X_0) - g(X_1))^2 \, \frac{1 + \mathbf{1}\{X_1 \notin B(o,\ell)\}}{2} \right] \end{aligned} \tag{2.6.53}$$

(c.f. [LP17, Lemma 13.6] for the third equality). Since $\beta(\ell)$ is the spectral gap of the chain killed at $B(o, \ell)^c$, we have by the Courant-Fischer characterization of $\beta(\ell)$ that

$$\beta(\ell) = \min_{g \in \mathcal{C}_\ell} \left\{ \frac{\langle (I - P_{B(o,\ell)}) g, g \rangle_{\pi_\ell}}{\mathbb{E}_{\pi_\ell}[g^2]} \right\}. \tag{2.6.54}$$

Using (2.6.53), and since the test function $f : x \mapsto d(x, B(o,\ell)^c)$ is 1-Lipschitz we get that $\langle (I - P_{B(o,\ell)}) f, f \rangle_{\pi_\ell} \leq 1$. Since moreover $f \in \mathcal{C}_\ell$, we have

$$\frac{1}{\beta(\ell)} \geq \frac{\mathbb{E}_{\pi_\ell}[f^2]}{\langle (I - P_{B(o,\ell)}) f, f \rangle_{\pi_\ell}} \geq \mathbb{E}_{\pi_\ell}[f^2] \geq (\ell - r)^2 \pi_\ell(B(o, r)) \geq \frac{\ell^2 V(r)}{4V(\ell)},$$

where we used the fact that $f(x) \geq \ell - r$ for all $x \in B(o, r)$. This concludes the proof. ☐

Lemma 2.6.4. *Let (Γ) be a sequence of finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \leq D^5$, and let $a, b > 0$. There exists $C = C(a, b, d) > 0$ such that*

$$\mathbb{P}_o(T_{B^c_{aD}} > bD^2) \geq C. \tag{2.6.55}$$

Proof. Let $\ell := 2\lfloor aD/2 \rfloor$, $r = \ell/2$, and $t = bD^2$, so that ℓ is even and $\mathbb{P}_o(T_{B^c_{aD}} > bD^2) \geq \mathbb{P}_o(T_{B^c_\ell} > t)$. By [TT21, Corollary 1.5] (which we recalled in (2.5.20)), we have $V(r) \gtrsim n$. Plugging this into (2.6.52), we deduce that $\beta(r) \lesssim r^{-2}$, and therefore that $e^{-\beta(r)t} \gtrsim 1$. Plugging this into (2.6.50) concludes the proof. ☐

Lemma 2.6.5. *Let (Γ) be a sequence of finite (connected) vertex-transitive graphs of degree d such that $n = |\Gamma| \leq D^5$, and let $\varepsilon > 0$. There exists $c > 0$ such that*

$$\alpha_o(B_{\varepsilon D}) \geq c(\varepsilon). \tag{2.6.56}$$

Proof. The intuition of this proof is that if a walk starts more than εD away from o, there should be a positive probability that, after time of order D^2, it finds itself in the ball $B_{\varepsilon D}$ (indeed, mixing times are hitting times of large sets, see [PS15]) but has not touched o yet (since starting from the quasistationary distribution α_o and conditioning on hitting o, the distribution at any given later time remains α_o).

Suppose $x \in B^c_{\varepsilon D}$. Then

$$\mathbb{P}_x(T_{B_{\varepsilon D/2}} \leq t^{(\infty)}_{\text{mix}}(1/2)) \geq \mathbb{P}_x(X_{t^{(\infty)}_{\text{mix}}(1/2)} \in B_{\varepsilon D/2}) \geq \frac{1}{2} \frac{V(\varepsilon D/2)}{n} \gtrsim 1. \tag{2.6.57}$$

Furthermore, by Lemma 2.6.4 (recalling that $t_{\text{mix}}^{(\infty)}(1/2) \asymp D^2$ by Proposition 2.2.2), given that the walk enters $B_{\varepsilon D/2}$, the conditional probability that it remains in the annulus $B_{3\varepsilon D/4} \backslash B_{\varepsilon D/4}$ until time $t_{\text{mix}}^{(\infty)}(1/2)$ is at least c for some constant c. Therefore, setting $t = t_{\text{mix}}^{(\infty)}(1/2)$,

$$\begin{aligned}
\alpha(B_{\varepsilon D}) &= \mathbb{P}_\alpha(X_t \in B_{\varepsilon D} | T_o \geq t) \\
&\geq \alpha(B_{\varepsilon D}^c) \mathbb{E}_{\alpha|_{B_{\varepsilon D}^c}}(X_t \in B_{\varepsilon D} | T_o \geq t) \\
&\geq \alpha(B_{\varepsilon D}^c) \inf_{x \in B_{\varepsilon D}^c} \mathbb{P}_x(X_t \in B_{\varepsilon D}; T_o \geq t) \\
&\gtrsim \alpha(B_{\varepsilon D}^c).
\end{aligned}$$

Thus $\alpha(B_{\varepsilon D})$ is bounded away from zero, as desired. □

Proof of Proposition 2.6.8. By Proposition 2.6.3, for $x, y \in \Gamma$, we have $\mathbb{E}_x T_y = \mathbb{E}_\pi T_o - nG(x,y)$, so $\mathbb{E}_x T_y \leq \mathbb{E}_\pi T_o$ if and only if $G(x,y) \geq 0$. Combining this with Proposition 2.6.4, there exists $\eta > 0$ such that for every $x \in B(o, \eta D) =: B_\eta$,

$$\mathbb{E}_x T_o \leq \mathbb{E}_\pi T_o. \tag{2.6.58}$$

Therefore, since $\alpha(B_\eta) \gtrsim 1$ by Lemma 2.6.5, and from Proposition 2.2.3, we conclude that

$$\mathrm{t}_{\text{hit}} - \mathbb{E}_\alpha T_o \geq \sum_{x \in B_\eta} \alpha(x) \left(\mathrm{t}_{\text{hit}} - \mathbb{E}_\pi T_o\right) = \alpha(B_{\eta D}) \left(\mathrm{t}_{\text{hit}} - \mathbb{E}_\pi T_o\right) \gtrsim D^2, \tag{2.6.59}$$

as desired. □

Proposition 2.6.9. Let (Γ) be a collection of finite (connected) vertex-transitive of fixed degree d, and assume that as $n = |\Gamma| \to \infty$,

$$D^2 \log n \gtrsim n. \tag{2.6.60}$$

Then there exists a constant $\kappa < 1$ such that for $s \in \mathbb{R}$ large enough, we have at time $t_s = \mathrm{t}_{\text{hit}}((\log n) + s)$

$$\mathbb{P}(\tau_{\text{cov}} > t_s) \leq \kappa \left(1 - e^{-e^{-s}}\right). \tag{2.6.61}$$

In particular $\frac{\tau_{\text{cov}}}{\mathrm{t}_{\text{hit}}} - \log |\Gamma|$ asymptotically does not have Gumbel fluctuations, and for s large enough, $\mathrm{d}_{\text{TV}}(\mathcal{L}(U(t_s)), \mu_s^{\otimes \Gamma})$ does not converge to 0.

Proof. From Proposition 2.6.8, there exists a constant $c > 0$ such that

$$\frac{\mathrm{t}_{\text{hit}}}{\mathbb{E}_\alpha T_o} \geq 1 + c \frac{D^2}{\mathbb{E}_\alpha T_o}. \tag{2.6.62}$$

It follows from Proposition 2.6.7 that there exists a constant $C > 0$ (depending on d, on c and the implicit constant in (2.6.60)), such that

$$\frac{\mathrm{t}_{\text{hit}}}{\mathbb{E}_\alpha T_o} \geq 1 + \frac{C}{\log n}. \tag{2.6.63}$$

From a simple union bound and (2.6.63), we deduce that

$$\mathbb{P}(\tau_{\text{cov}} > t_s) \leq n \mathbb{P}_\pi(T_o > t_s) \leq n \mathbb{P}_\alpha(T_o > t_s) = n \exp\left(-\frac{\mathrm{t}_{\text{hit}}}{\mathbb{E}_\alpha T_o}((\log n) + s)\right) \leq e^{-s} e^{-C}. \tag{2.6.64}$$

Hence, for s larger than some $s_0 = s_0(C)$, we have

$$\mathbb{P}\left(\frac{\tau_{\text{cov}}}{\mathrm{t_{hit}}} - \log|\Gamma| > s\right) = \mathbb{P}(\tau_{\text{cov}} > t_s) \le e^{-C}\mathbb{P}(\chi > s) \le \kappa\left(1 - e^{-e^{-s}}\right), \tag{2.6.65}$$

where $\kappa := \frac{1+e^{-C}}{2} < 1$. This proves that $\frac{\tau_{\text{cov}}}{\mathrm{t_{hit}}} - \log|\Gamma|$ does not have Gumbel fluctuations. Since $\mathbb{P}(\tau_{\text{cov}} > t_s) = \mathbb{P}(U(t_s) \neq \varnothing)$, it also shows that for $s \ge s_0$, we have

$$\mathrm{d_{TV}}(\mathcal{L}(U(t_s)), \mu_s^{\otimes\Gamma}) \ge (1-\kappa)\left(1 - e^{-e^{-s}}\right). \tag{2.6.66}$$

□

2.6.3 Proof that the diameter condition is necessary in Theorem 2.1.3

In this section, we assume that $t_{\text{rel}} = o(\mathrm{t_{hit}})$. We want to show that the law of the uncovered set at time t_s^* is far, in the total variation sense, from the product measure $\mu_s^{\otimes\Gamma}$. Note that this immediately implies Theorem 2.1.3: indeed, since $t_{\text{rel}} \le dD^2$ by (2.2.1) and $\mathrm{t_{hit}} \ge t_{\langle\text{hit}\rangle} = nG(o,o) \ge n(1-o(1/\log n))$ by Proposition 2.6.3, the assumption $D^2 \lll n$ implies that $t_{\text{rel}} \lll \mathrm{t_{hit}}$. Therefore, it suffices to prove the result under the sole assumption $t_{\text{rel}} = o(\mathrm{t_{hit}})$.

To prove that the law of the uncovered set is far from a product measure when the diameter condition fails, we might initially be tempted to show that the uncovered set is "too" clustered, i.e., the probability that two relatively nearby points are uncovered is larger than it should be under the independent scenario. However, this turns out to be very difficult to control as the contribution to the moments of order k of the size of the uncovered set coming from nearby points start exploding when the diameter condition fails. We cannot translate this into estimates about events of positive probability for the uncovered set; roughly speaking, either the Bonferroni inequality goes in the wrong direction, or one would need to keep track of moments of higher order and compare how they blow up.

Instead, we show that points that are sufficiently far apart are negatively correlated. It turns out that this enables us to use the Bonferroni inequality (i.e., union bound) as this is an upper bound.

Proposition 2.6.10. Let $\gamma > 0$, and (Γ) be a collection of finite (connected) vertex-transitive graphs of fixed degree d, such that $n = |\Gamma| \le \gamma D^2 \log n$ and $t_{\text{rel}} = o(\mathrm{t_{hit}})$, and let (for every $s \in \mathbb{R}$) $\mu_s^{\otimes\Gamma}$ denote the product over all the vertices of the graph of the Bernoulli law μ_s with parameter $e^{-s}/|\Gamma|$. Then for every fixed $s \in \mathbb{R}$, there exists a constant $c^* = c^*(s,d,\gamma)$ such that as $|\Gamma| \to \infty$, for $t \in \{t_{\langle s\rangle}, t_s^*, t_s\}$,

$$\mathrm{d_{TV}}(\mathcal{L}(U(t)), \mu_s^{\otimes\Gamma}) \ge c^* + o(1).$$

We start by comparing $t_{\langle s\rangle}$ and t_s^*. Recall that $t_{\langle s\rangle} = t_{\langle\text{hit}\rangle}((\log n) + s)$ and t_s^* is such that $\mathbb{E}|U(t_s^*)| = e^{-s}$.

Lemma 2.6.6. *Let $s \in \mathbb{R}$. For n large enough, we have $\mathbb{E}(U(t_{\langle s\rangle})) \ge e^{-s}$, or in other words $t_{\langle s\rangle} \le t_s^*$.*

Proof. Set $\alpha = \alpha_o$, and let $s \in \mathbb{R}$. We have

$$\mathbb{E}|U(t_{\langle s\rangle})| = n\mathbb{P}_\pi(T_o > t_{\langle s\rangle}) = n\frac{\mathbb{P}_\pi(T_o > t_{\langle s\rangle})}{\mathbb{P}_\alpha(T_o > t_{\langle s\rangle})}e^{-t_{\langle s\rangle}/\mathbb{E}_\alpha T_o} = n\frac{\mathbb{P}_\pi(T_o > t_{\langle s\rangle})}{\mathbb{P}_\alpha(T_o > t_{\langle s\rangle})}e^{-\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\alpha T_o}((\log n)+s)}. \tag{2.6.67}$$

By Theorem 2.5.1, setting $\theta := 1 - \frac{1}{\|\alpha/\pi\|_2^2}$, we have

$$\frac{\mathbb{P}_\pi(T_o > t_{\langle s\rangle})}{\mathbb{P}_\alpha(T_o > t_{\langle s\rangle})} \geq 1 - \theta, \tag{2.6.68}$$

and,

$$\frac{\mathbb{E}_\pi T_o}{\mathbb{E}_\alpha T_o} \leq 1 - \theta + \theta \frac{t_{\text{rel}}}{\mathbb{E}_\alpha T_o}. \tag{2.6.69}$$

Since $t_{\text{rel}} = o(\mathrm{t_{hit}})$ by hypothesis, we have $\frac{t_{\text{rel}}}{\mathbb{E}_\alpha T_o}$ and $\theta \to 0$. Therefore, using the inequality $e^x \geq 1 + x$, valid for all $x \in \mathbb{R}$, we have

$$\mathbb{E}|U(t_{\langle s\rangle})| \geq e^{-s}(1-\theta)(1 + \theta \log n + o(\theta \log n)) = e^{-s}(1 + \theta \log n + o(\theta \log n)), \tag{2.6.70}$$

which concludes the proof. □

We now prove a technical lemma.

Lemma 2.6.7. *Let* (Γ) *as in Proposition 2.6.10, and* c, S_1, S_2 *as in Proposition 2.6.6. There exists a constant* $C > 0$ *such that for every* $(x, y) \in S_1 \times S_2$ *and* $s \in \mathbb{R}$*, we have (as* $|\Gamma| \to \infty$*), setting* $A = \{x, y\}$*,*

$$\mathbb{P}_\pi(T_A > t_{\langle s\rangle}) \leq \frac{e^{-2s}}{n^2} e^{-C + o(1)}. \tag{2.6.71}$$

Proof. Let us fix $s \in \mathbb{R}$. For every $(x, y) \in S_1 \times S_2$, we have by (2.6.15)

$$q_A = \frac{2}{1 + \frac{G(x,y)}{G(o,o)}} \geq \frac{2}{1 - c\frac{D^2}{nG(o,o)}} \geq 2\left(1 + c\frac{D^2}{nG(o,o)}\right). \tag{2.6.72}$$

Moreover, using Equation (1) in [AB92] (which we recalled in Theorem 2.1.4) for the first inequality, and by Theorem 2.1.5, we have uniformly over all $A \subset \Gamma$ of size 2,

$$\mathbb{P}_\pi(T_A > t_{\langle s\rangle}) \leq \mathbb{P}_{\alpha_A}(T_A > t_{\langle s\rangle}) = \exp\left(-q_A\, t_{\langle s\rangle}\left(1 + O\left(\frac{t_{\text{rel}}}{\mathrm{t_{hit}}}\right)^2\right)\right). \tag{2.6.73}$$

It follows, recalling that $\frac{D^2}{nG(o,o)} \asymp \frac{t_{\text{rel}}}{\mathrm{t_{hit}}}$, that there exists a constant $c' > 0$ such that for every $(x, y) \in S_1 \times S_2$,

$$\mathbb{P}_\pi(T_A > t_{\langle s\rangle}) \leq \frac{e^{-2s}}{n^2} \exp\left(-2c'\frac{t_{\text{rel}}}{\mathrm{t_{hit}}}(\log n)\left(1 + \frac{s}{\log n} + O\left(\frac{t_{\text{rel}}}{\mathrm{t_{hit}}}\right)\right)\right). \tag{2.6.74}$$

From Proposition 2.6.7, there exists a constant C such that (recalling that $t_{\text{rel}} \asymp D^2$)

$$2c'\frac{t_{\text{rel}} \log n}{\mathrm{t_{hit}}} \geq C, \tag{2.6.75}$$

which allows us to conclude, using that $t_{\text{rel}} \lll \mathrm{t_{hit}}$, that

$$\mathbb{P}_\pi(T_A > t_{\langle s\rangle}) \leq \frac{e^{-2s}}{n^2} e^{-C + o(1)}, \tag{2.6.76}$$

as desired. □

Proof of Proposition 2.6.10. Let $s \in \mathbb{R}$. By definition of the total variation distance, it is enough to find a subset B of $\mathcal{P}(\Gamma)$ and a constant c^* such that for $t \in \left\{t_{\langle s\rangle}, t_s^*, t_s\right\}$,

$$\mu_s^{\otimes\Gamma}(B) - \mathbb{P}(U(t) \in B) > c^*. \tag{2.6.77}$$

Let c, S_1, S_2 as in Proposition 2.6.6, and $C > 0$ from Lemma 2.6.7. Set $b = b(s) = e^{-s}$ and let $a > 0$ be small enough such that $|S_1| \geq an$ (for all Γ as in the statement of the proposition), and $1 - 2e^{-ab} + e^{-2ab} \geq (ab)^2(1 + e^{-C})/2$. (Such a choice for a is possible since as $a \to 0$, we have $1 - 2e^{-ab} + e^{-2ab} = (ab)^2 + O(a^3)$.) Let S_1' and S_2' be sub-balls of (respectively) S_1 and S_2 such that $|S_1'| = |S_2'| = (a + o(1))n$, and set

$$B := \{A \subset \Gamma : A \cap S_1' \neq \varnothing \text{ and } A \cap S_2' \neq \varnothing\}. \tag{2.6.78}$$

By a union bound and Lemma 2.6.7, we get the following upper bound on $\mathbb{P}(U(t_{\langle s\rangle}) \in B)$

$$\begin{aligned}\mathbb{P}(U(t_{\langle s\rangle}) \in B) \leq \sum_{(x,y)\in S_1'\times S_2'} \mathbb{P}_\pi\left(\{x,y\} \subset U(t_{\langle s\rangle})\right) &\leq \frac{|S_1'||S_2'|}{n^2} e^{-2s}\left(e^{-C} + o(1)\right) \\ &= (ab)^2 e^{-C} + o(1).\end{aligned} \tag{2.6.79}$$

Let us now lower bound $\mu_s^{\otimes\Gamma}(B)$. Let $B_i = \{A \subset \Gamma : A \cap S_i' \neq \varnothing\}$ for $i \in \{1,2\}$. Since $B = B_1 \cap B_2$, we have

$$\mu_s^{\otimes\Gamma}(B) = 1 - \mu_s^{\otimes\Gamma}(B_1^c \cup B_2^c) = 1 - \left(\mu_s^{\otimes\Gamma}(B_1^c) + \mu_s^{\otimes\Gamma}(B_2^c) - \mu_s^{\otimes\Gamma}(B_1^c \cap B_2^c)\right). \tag{2.6.80}$$

Moreover,

$$\mu_s^{\otimes\Gamma}(B_1^c) = \mu_s^{\otimes\Gamma}(B_2^c) = \left(1 - \frac{b}{n}\right)^{(a+o(1))n} = e^{-ab+o(1)}, \tag{2.6.81}$$

and we have similarly $\mu_s^{\otimes\Gamma}(B_1^c \cap B_2^c) = e^{-2ab+o(1)}$. If follows that

$$\mu_s^{\otimes\Gamma}(B) \geq 1 - 2e^{-ab+o(1)} + e^{-2ab+o(1)}, \tag{2.6.82}$$

and hence, by definition of a, that

$$\mu_s^{\otimes\Gamma}(B) - \mathbb{P}(U(t_{\langle s\rangle}) \in B) \geq (ab)^2\frac{1+e^{-C}}{2} - (ab)^2 e^{-C} + o(1) = (ab)^2\frac{1-e^{-C}}{2} + o(1). \tag{2.6.83}$$

Recall that $t_s^* \geq t_{\langle s\rangle}$ (for n large enough) and $t_s \geq t_{\langle s\rangle}$ (for all n), so since $t \mapsto \mathbb{P}_\pi(T_A > t)$ is decreasing for every A, (2.6.79) also holds with $\mathbb{P}(U(t_s^*) \in B)$ or $\mathbb{P}(U(t_s) \in B)$ on the left hand side. Therefore, (2.6.83) also holds with t_s^* or t_s instead of $t_{\langle s\rangle}$, and the proof is complete. □

2.7 A strongly uniformly transient graph without Gumbel fluctuations

In this section, we construct an example of a sequence of vertex-transitive graphs which satisfy the strong uniform transience (SUT) of (2.1.11), but not the diameter condition (DC).

The idea is to consider the following graph: for $m \geq 2$ even, let $\Gamma = C_m \times G$ be the Cartersian product of a cycle C_m of length m, together with an expander Cayley graph G (in fact it will be convenient to take a Ramanujan graph of degree greater than three, see Morgenstern [Mor94] for an explicit construction generalising the famous construction of Lubotzky, Phillips and Sarnak

[LPS88]) of size mh, where $h = h_m \geq 1$, and say $h_m \leq m$, say. (The interesting examples will be those for which $h \lesssim \log m$.)

Thus the total size (number of vertices) of Γ is $n = m^2 h$ and the diameter is $\asymp m$ and for $h \lesssim \log m$, Γ does not satisfy the diameter condition (DC).

Proposition 2.7.1. The graph Γ above satisfies the strong uniform transience (SUT) condition whenever $h_m \to \infty$, no matter how slowly.

Proof. For a vertex $x \in \Gamma$, let us write $x = (\hat{x}, \check{x})$ with $\hat{x} \in C_m$ and $\check{x} \in G$ so that $\hat{x}$ denote the cycle coordinate and $\check{x}$ the Ramanujan coordinate. Since both C_m and G are vertex-transitive, for a continuous-time random walk $X_t = (\hat{X}_t, \check{X}_t) \in \Gamma$, the coordinates $\hat{X}$ and $\check{X}$ are in fact independent continuous time random walks with rates $2/d$ and $1 - 2/d$ respectively (where $d = \deg(x)$ is the total degree on Γ.)

Since the coordinates are independent, we have that for every $t \geq 0$,

$$p_t(x, y) = \hat{p}_t(\hat{x}, \hat{y})\check{p}_t(\check{x}, \check{y}),$$

where $\hat{p}_t$ denote the transition probabilities of the random walk on the cycle C_m (with rate $2/d$), and $\check{p}_t$ denote those on the Ramanujan graph G (with rate $1 - 2/d$).

Furthermore, by Proposition 2.2.2, $t_{\text{mix}} = t_{\text{mix}}(1/4) \lesssim m^2$. It is therefore easy to estimate both $\hat{p}_t$ and $\check{p}_t$ for $t \leq m^2$: namely, on the cycle we know (for instance from Proposition 2.2.1) that

$$\hat{p}_t(\hat{o}, \hat{o}) \lesssim \frac{1}{\sqrt{t+1}};$$

and on the Ramanujan component, by (2.2.29),

$$|\check{p}_t(\check{o}, \check{o}) - \check{\pi}(\check{o})| \leq e^{-\lambda t}$$

where $\lambda > 0$ is the spectral gap on the Ramanujan component (which, by definition, is bounded away from zero), and $\check{\pi}$ is the uniform distribution on G. Since $\check{\pi}(\check{o}) = 1/(mh)$, we deduce that $\check{p}_t(\check{o}, \check{o}) \leq 1/(mh) + e^{-\lambda t}$.

For each fixed $s > 0$

$$\begin{aligned}
\mathbb{E}_o[L_o(t_{\text{mix}}) - L_o(s)] &= \int_s^{t_{\text{mix}}} p_t(o, o) dt \\
&\lesssim \int_s^{t_{\text{mix}}} \frac{1}{\sqrt{t+1}} (\frac{1}{mh} + e^{-\lambda t}) dt \\
&\lesssim \frac{\sqrt{t_{\text{mix}}}}{mh} + \int_s^{\infty} e^{-\lambda t} \frac{1}{\sqrt{t+1}} dt
\end{aligned}$$

Since $t_{\text{mix}} \lesssim m^2$ and $h = h_m \to \infty$, the first term tends to zero. The second term does not depend on m, and is the integral of a function which is clearly integrable. Consequently, the limsup as $s \to \infty$ is zero. Thus

$$\limsup_{s \to \infty} \limsup_{n \to \infty} \mathbb{E}_o(L_o(t_{\text{mix}}) - L(s)) = 0,$$

which shows (2.1.11). □

Chapter 3

Limit profile for random transpositions

We present an improved version of Diaconis–Shahshahani upper bound lemma, which is used to compute the limiting value of the distance to stationarity. We then apply it to the random transposition shuffle.

3.1 Introduction

3.1.1 Main results

Let $\mathfrak{S}_n$ be the symmetric group of indice n and P_n the probability on $\mathfrak{S}_n$ defined by

$$P_n(Id) = \frac{1}{n} \quad \text{and} \quad P_n(\tau) = \frac{2}{n^2} \quad \text{if } \tau \text{ is a transposition.} \tag{3.1.1}$$

This is the random transposition shuffle on $\mathfrak{S}_n$, as studied in a landmark paper of Diaconis and Shahshahani [DS81].

Let also U_n be the uniform probability on $\mathfrak{S}_n$. If E is a set and μ, ν are probabilities on E, we define the total variation distance[1] between μ and ν by the formula

$$\mathrm{d_{TV}}(\mu,\nu) = \frac{1}{2}\mathrm{d}_1(\mu,\nu) = \frac{1}{2}\sum_{x\in E}|\mu(x)-\nu(x)|\,. \tag{3.1.2}$$

In [DS81], Diaconis and Shahshahani showed that this random walk undergoes a cutoff phenomenon at $\frac{1}{2}n\log(n)$, that is, letting $f(n) = \frac{1}{2}n\log(n)$, that for all $0 < \epsilon < 1$,

$$\mathrm{d_{TV}}\left(P_n^{*\lfloor(1-\epsilon)f(n)\rfloor}, U_n\right) \xrightarrow[n\to\infty]{} 1 \quad \text{and} \quad \mathrm{d_{TV}}\left(P_n^{*\lfloor(1+\epsilon)f(n)\rfloor}, U_n\right) \xrightarrow[n\to\infty]{} 0. \tag{3.1.3}$$

Despite a lot of work on mixing times in general and on random transpositions in particular (see references below), obtaining a precise description of the way this transition occurs has remained an open problem, formally asked at an AIM workshop on Markov chains mixing times in 2016 (http://aimpl.org/markovmixing/5/).

[1] In the proofs we will use the L^1 distance, noted d_1, in order not to carry the factor $\frac{1}{2}$.

Problem 5.3. [Nathanaël Berestycki] Can we obtain the cutoff profile for the uniform random transposition walk on S_n?

Our main result is the following.

Theorem 3.1.1. Let $c \in \mathbb{R}$. Then we have

$$\mathrm{d}_{\mathrm{TV}}\left(P_n^{*\lfloor \frac{1}{2} n \log(n)+cn \rfloor}, U_n\right) \xrightarrow[n\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1+e^{-2c}\right), \mathrm{Poiss}(1)\right), \tag{3.1.4}$$

where $\mathrm{Poiss}(a)$ stands for the Poisson law of parameter a.

3.1.2 Limiting profile conjectures

We anticipate the limiting profile

$$\mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1+e^{-c}\right), \mathrm{Poiss}(1)\right),$$

which we obtain in our problem if we replace the time $\lfloor \frac{1}{2} n \log(n) + cn \rfloor$ by a slightly more natural time, $\lfloor \frac{1}{2}(n\log(n)+cn)\rfloor$, to arise for many other mixing time problems on $\mathfrak{S}_n$, namely the problems where the last things to be mixed are the fixed points. It seems to be often the case when the probability P_n is constant over conjugacy classes. For example, using the formulas in [FJR03], one can adapt the present proof for random k-cycles (k fixed) at time $\lfloor \frac{1}{k}(n\log(n)+cn)\rfloor$, and we conjecture that the same limiting profile still holds for random conjugacy classes of size $o(n)$, as studied in [BŞ19], but that it would be technically much harder to adapt the present proof in that case. For this general case, a beautiful formula (Proposition 10.15 in [Mé17]) used in the proof of the Stanley–Féray formula, which allows to compute any reduced character as an expectation, $\chi^\lambda(\mu) = \mathbb{E}[(-1)^{\mathrm{inv}(\sigma_\mu)}]$, might be very useful.

We conjecture that this profile also holds for the random involution walk studied by Megan Bernstein in [Ber18], at time $\lfloor \frac{1}{\log(p)}(\log(n)+c)\rfloor$. For other problems where the limiting profile is known, see [BD92] and [Lac16a].

3.1.3 Links with previous results and idea of the proof

3.1.3.1 Links with previous results

In 1981, Diaconis and Shahshahani showed in [DS81], using representations of the symmetric group, a cutoff[2] at $\lfloor \frac{1}{2} n\log(n)\rfloor$ for the random transposition shuffle, giving asymptotic inequalities at time $\lfloor \frac{1}{2} n\log(n)+cn\rfloor$, $c > 0$ fixed. In 1987, Matthews gave in [Mat88b] a very precise lower bound, which matches with the limiting profile. (However, the upper bound part of Matthews' proof has a flaw, as pointed out in 2017 by Graham White in [Whi19]. In 2019, Graham White presented in [Whi19] an alternative probabilistic proof of the upper bound.) In 2011, Berestycki, Schramm and Zeitouni generalized in [BSZ11] the previous result to the shuffle by random k-cycles, for k fixed as $n \to \infty$, proving a cutoff at $\lfloor \frac{1}{k} n\log(n)\rfloor$, conjectured by Diaconis. Finally, in 2014, Berestycki and Şengül generalized again this result, in [BŞ19], to any conjugacy class whose support is $o(n)$, and without representation theory.

Before giving more details about our proof, let us mention some other interesting related works and sources. For a study of couplings for the random transposition walk, see Blumberg's thesis [Blu11]. For other bounds on the characters of the symmetric group, see [LP02], by Lulov and Pak. Also, an intriguing cutoff at $\frac{3}{4} n\log(n) - \frac{1}{4} n\log\log(n)$ for the random to random card

[2] In fact their lower bound is $1/e$ so it is not exactly a cutoff.

shuffle was proven by Bernstein and Nestoridi in [BN19]. About spacial mixing, see the very recent article [DP22] by Diaconis and Pal. For polynomial bounds on the diameter and mixing time of shuffles with random sets of generators, see [BH05] by Babai and Hayes, and [HSZ15] by Helfgott, Seress and Zuk. Finally, about cutoff on compact quantum groups, see [Fre19] by Freslon.
To learn about mixing times, an excellent source is the book [LP17] by Levin and Peres. Another very good source is the lecture notes by Salez, [Sal20]. About representation theory, the first part of Serre's book, [Ser78], gives nice and clear introduction. To study more specifically the representations of the symmetric group, see the very complete book by Méliot, [Mé17], which contains many recent results.

The proof in [DS81] relies on the so-called *Diaconis–Shahshahani upper bound lemma*, which leads to a sum over irreducible representations which they delicately bound with representation theory and analysis. Actually we can observe that the only place where a lot of information (we lose a factor e in the limit $c \to \infty$ of the limit profile) is lost on the limit profile is at the very begining, when the Cauchy–Schwarz inequality is used in the proof of the upper bound lemma. Section 3.2 presents a remedy to this information loss, improving the upper bound lemma to an approximation lemma (Lemma 3.2.1) which is asymptotically much more precise. Section 3.4.1, which is quite technical, generalizes the asymptotic bounds of Diaconis and Shahshahani to any $c \in \mathbb{R}$.

Another crucial point of our proof is to pack together, in the sums over the irreducible representations $\lambda = (\lambda_1, \dots)$ of $\mathfrak{S}_n$, all the partitions with the same λ_1. More precisely, Section 3.4.2 shows that when $j \in \mathbb{N}^*$ is fixed, we can study the sum over the partitions with λ_1 equal to $n - j$ as a sum over the partitions of the integer j, resulting in explicit manipulable formulas.

To understand where the limiting profile comes from, observe that, thanks to the lower bound of Matthews, the key observable is the number of fixed points. The limit profile is the distance between the asymptotic distribution of the number of fixed points of our walk at time $\lfloor \frac{1}{2} n \log(n) + cn \rfloor$, which is a Poiss$(1+e^{-2c})$ distribution, and that of a permutation taken uniformly at random, that is, Poiss(1).

Theorem 3.1.1 stated above gives support to the following conjecture of Nathanaël Berestycki.

Conjecture 3.1.2. Let τ_n be the first time that all cards have been touched, and let X_{τ_n} be the state of the deck of cards at this (random) time. Then $\mathrm{d_{TV}}(X_{\tau_n}, U_n) \to 0$ as $n \to \infty$.

In other words, the conjecture says that τ_n is a stopping time at which the random permutation is well mixed for all practical purposes. Note that at time $\tau_n - 1$ the permutation contains at least one fixed point, so that $\mathrm{d_{TV}}(X_{\tau_n - 1}, U_n)$ cannot converge to zero. Hence, the conjecture implies that τ_n is in some strong sense optimal for mixing the deck of cards.

Let us now explain in what way Theorem 3.1.1 above is related to this conjecture. For any time t, let G_t be the random graph which contains an edge (i, j) if and only if the corresponding transposition has been applied at least once prior to time t. Then G_t is essentially a realisation of the Erdős–Rényi random graph with parameters n and $p = 1 - \exp(-t/\binom{n}{2})$. It is easy to check that any cycle of the random permutation X_t at time t, considered as a set, is a subset of a connected component of G_t. Hence, it makes sense to consider the cycle structure of the permutation *restricted to any particular connected component of* G_t. Let $\mathfrak{C}_t$ be the largest component of G_t (which is macroscopic if $t \geq cn$ for some $c > 1$, and actually contains all vertices with high probability after time τ_n). $\mathfrak{C}_t$ is called the giant component of G_t. By a famous result of Schramm [Sch05], the distribution of the lengths of the largest cycles of X_t within $\mathfrak{C}_t$, normalised by the total size $|\mathfrak{C}_t|$ of the giant component, converges to a Poisson–Dirichlet distribution (in the sense of finite dimensional distributions). Hence, these largest cycles can be seen to coincide in the limit with the distribution of a uniform permutation on the giant component (see, e.g., [BSZ11]). A stronger version of Schramm's theorem would be the following conjecture (also by Nathanaël Berestycki).

Conjecture 3.1.3. Suppose $t \geq cn/2$ for some $c > 1$. Given $\mathfrak{C}_t$, the distribution of $X_t|_{\mathfrak{C}_t}$, is approximately uniform, in the sense that $\mathrm{d}_{\mathrm{TV}}(X_t|_{\mathfrak{C}_t}, U_{\mathfrak{C}_t}) \to 0$ in probability as $n \to \infty$, where $U_{\mathfrak{C}_t}$ is a uniform permutation on the giant component $\mathfrak{C}_t$.

It is not hard to see that Conjecture 3.1.3 implies Conjecture 3.1.2. Indeed, Conjecture 3.1.3 implies a very precise description of the structure of X_t close to the mixing time: if $t = \lfloor \frac{1}{2} n \log n + cn \rfloor$, then according to this conjecture X_t would consist, if $\tau_n > t$ of a permutation that is approximately uniform on $n-1$ points, plus an extra fixed point; and would otherwise be indistinguishable from a uniform permutation if $\tau_n \geq t$. Such a description would imply that

$$\mathrm{d}_{\mathrm{TV}}(X_t, U_n) = \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Fix}(X_t), \mathrm{Poiss}(1)\right) + o(1), \tag{3.1.5}$$

where $\mathrm{Fix}(X_t)$ is the number of fixed points of X_t. It is furthermore relatively easy to check that $\mathbb{P}(\tau_n > t) \to e^{-2c}$ and hence, still assuming Conjecture 3.1.3, we would deduce

$$\mathrm{d}_{\mathrm{TV}}(X_t, U_n) = \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1 + e^{-2c}\right), \mathrm{Poiss}(1)\right), \tag{3.1.6}$$

where the extra term e^{-2c} in the right hand side accounts precisely for the probability that $\tau_n > t$. Of course, this last display is precisely the content of our Theorem 3.1.1.

3.1.3.2 Organisation of the article

In Section 3.2, we present the improvement of Diaconis–Shahshahani upper bound lemma, using the non-commutative Fourier transform, which brings us back to group representations. In Section 3.3, we will recall some results on the representations of the symmetric group, get precise estimations of the hook-length and Murnagham–Nakayama combinatorial formulas when the size n of our partitions tend to infinity with $n - \lambda_1$ constant, and we will prove some upper bounds useful in the sequel. In Section 3.4, we will prove the announced theorem decomposing approximation by approximation. From now on, k will denote without ambiguity the integer

$$k = k(n, c) = \lfloor \frac{1}{2} n \log(n) + cn \rfloor. \tag{3.1.7}$$

3.1.3.3 Idea of the proof

The algebraic objects $\widehat{\mathfrak{S}_n}$, triv, d_λ, s_λ and ch^λ will be defined at the beginning of Section 3.2. For all $\sigma \in \mathfrak{S}_n$, $\mathrm{Fix}(\sigma)$ will denote the number of fixed points of the permutation σ. For $j \in \mathbb{N}^*$, let us also define the polynomial $T_j(z)$ by the formula $\sum_{i=0}^{j} \binom{z}{j-i} \frac{(-1)^i}{i!}$. The idea is to first fix $c \in \mathbb{R}$, and then to define for all $\epsilon > 0$ an integer $M = M(c, \epsilon)$ such that when n tends to infinity, all the following approximations are true up to ϵ.

Rewriting the sum using the Fourier transform and the improvement of Diaconis–Shahshahani lemma,

$$\mathrm{d}_1\left(P_n^{*k}, U_n\right) = \frac{1}{|\mathfrak{S}_n|} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{\lambda \in \widehat{\mathfrak{S}_n} \setminus \{\mathrm{triv}\}} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) \right| \approx \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{\lambda_1 \geq n-M} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) \right|. \tag{3.1.8}$$

Then, thanks to the polynomial convergence lemma and letting $M \to \infty$, we will get

$$\begin{aligned}\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda_1\geq n-M} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\right| &\approx \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{M} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right)\right| \\ &\approx \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right)\right|.\end{aligned}$$

Finally, letting $n \to \infty$,

$$\begin{aligned}\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right)\right| &= \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|e^{-e^{-2c}}\left(1+e^{-2c}\right)^{\mathrm{Fix}(\sigma)} - 1\right| \\ &\approx \mathbb{E}\left|e^{-e^{-2c}}\left(1+e^{-2c}\right)^{\mathrm{Poiss}(1)} - 1\right| \\ &= \mathrm{d}_1\left(\mathrm{Poiss}\left(1+e^{-2c}\right), \mathrm{Poiss}(1)\right).\end{aligned}$$

3.2 Improvement of Diaconis–Shahshahani upper bound lemma

In this section, we present the improvement of Diaconis–Shahshahani upper bound lemma. We will stay in the framework of finite groups, but this lemma can be used in a wider framework, of compact groups for example. Our aim is to get a better approximation than in [DS81] by not using Cauchy–Schwarz before Fourier.

Let G be a finite group, $\mathbb{C}G$ the group algebra of G and $\widehat{G}$ the set of the irreducible representations of G. We recall that the convolution product on $\mathbb{C}G$ is defined, for all $f_1, f_2 \in \mathbb{C}G$ and $z \in G$, by the formula

$$(f_1 * f_2)(z) = \sum_{xy=z} f_1(x) f_2(y). \tag{3.2.1}$$

We note triv the trivial representation of G and $\widehat{G}^* = \widehat{G}\backslash\{\mathrm{triv}\}$. For $\alpha \in \widehat{G}$, we also name ρ_α the matrix of the representation α, ch^α its character and d_α its dimension. Let us first recall the inversion formula for the non-commutative Fourier transform, well-explained in [Mé17]. For $f : G \to \mathbb{C}$ and $g \in G$, we have

$$f(g) = \sum_{\alpha\in\widehat{G}} \frac{d_\alpha}{|G|}\,\mathrm{Tr}\left(\rho^\alpha(g)^* \widehat{f}(\alpha)\right).$$

We deduce that for all $t \in \mathbb{N}$,

$$\begin{aligned}\mathrm{d}_1\left(P^{*t}, U\right) &= \sum_{g\in G}\left|P^{*t}(g) - U(g)\right| \\ &= \sum_{g\in G}\left|\sum_{\alpha\in\widehat{G}} \frac{d_\alpha}{|G|}\,\mathrm{Tr}\left((\widehat{P^{*t} - U})(\alpha)\rho^\alpha(g)^*\right)\right| \\ &= \sum_{g\in G}\left|\sum_{\alpha\in\widehat{G}^*} \frac{d_\alpha}{|G|}\,\mathrm{Tr}\left(\widehat{P^{*t}}(\alpha)\rho^\alpha(g)^*\right)\right|.\end{aligned}$$

In addition, as P is a function which is constant on every conjugacy class, we know that for each α, by Schur's lemma, $\widehat{P}(\alpha)$ is a homothety, of ratio

$$s_\alpha = \frac{\mathrm{Tr}(\widehat{P}(\alpha))}{d_\alpha}.$$

We hence obtain

$$\mathrm{d}_1\left(P^{*t}, U\right) = \frac{1}{|G|} \sum_{g \in G} \left| \sum_{\alpha \in \widehat{G}^*} d_\alpha s_\alpha^t \overline{\mathrm{ch}^\alpha(g)} \right|.$$

Now, if instead of having a single group G we have an increasing sequence of groups $(G_n)_{n\in\mathbb{N}}$, each group being provided with a probability measure P_n, and if $t = t(n)$ is a well-chosen time depending on n (and possibly on another parameter), we will wish to make n tend to infinity inside our sums, and thus obtain a convergence to an explicit formula which will prove a cutoff or give a limiting profile. The idea of the following lemma is to spot a finite set of irreducible representations which will (asymptotically) have most of the mass, in order to approximate the sum over all irreducible representations by a sum over only finitely many terms, uniformly in n, and then be allowed to make n tend to infinity inside the finite sum.

Lemma 3.2.1 (Approximation lemma)**.** *Let G be a finite group, P a probability on G constant on every conjugacy class, and $S \subset \widehat{G}^*$. Then,*

$$\left| \mathrm{d}_1\left(P^{*t}, U\right) - \frac{1}{|G|} \sum_{g \in G} \left| \sum_{\alpha \in S} d_\alpha s_\alpha^t \overline{\mathrm{ch}^\alpha(g)} \right| \right| \leq \sum_{\alpha \in \widehat{G}^* \setminus S} d_\alpha \left| s_\alpha \right|^t.$$

Proof. Using the fact that $||a| - |b|| \leq |a - b|$ and triangle inequalities,

$$\begin{aligned}
&\left| \mathrm{d}_1\left(P^{*t}, U\right) - \frac{1}{|G|} \sum_{g \in G} \left| \sum_{\alpha \in S} d_\alpha s_\alpha^t \overline{\mathrm{ch}^\alpha(g)} \right| \right| \\
&\quad \leq \frac{1}{|G|} \sum_{g \in G} \left| \sum_{\alpha \in \widehat{G} \setminus S} d_\alpha s_\alpha^t \overline{\mathrm{ch}^\alpha(g)} \right| \\
&\quad \leq \frac{1}{|G|} \sum_{g \in G} \sum_{\alpha \in \widehat{G} \setminus S} d_\alpha \left| s_\alpha \right|^t \left| \mathrm{ch}^\alpha(g) \right| \\
&\quad = \sum_{\alpha \in \widehat{G} \setminus S} d_\alpha \left| s_\alpha \right|^t \frac{1}{|G|} \sum_{g \in G} \left| \mathrm{ch}^\alpha(g) \right|. \quad (*)
\end{aligned}$$

Now, for every irreducible character α, by Cauchy–Schwarz inequality and orthonormality of the characters,

$$\frac{1}{|G|} \sum_{g \in G} \left| \mathrm{ch}^\alpha(g) \right| \leq \frac{1}{|G|} \sqrt{|G| \sum_{g \in G} \left| \mathrm{ch}^\alpha(g) \right|^2} = 1.$$

Plugging into $(*)$, this concludes the proof. ☐

3.3 The symmetric group and its representations

3.3.1 Hook-length formula

We recall a few facts from the representation theory of the symmetric group, that we will naturally index by integer partitions λ. In a diagram associated to a partition, the hook of a box is the number of boxes which are above or on the right of our box (including our box). We call $\text{équ}(\lambda)$ the product of the hooks of the partition λ. For example, consider the partition $\lambda = (7, 3, 2, 1, 1)$ of the integer 14 filled with its hooks:

1						
2						
4	1					
6	3	1				
11	8	6	4	3	2	1

.

In this case, we have:

$$\begin{aligned}\text{équ}(7,3,2,1,1) &= 11 \times 8 \times 6 \times (4 \times 3 \times 2 \times 1) \times (6 \times 3 \times 1 \times 4 \times 1 \times 2 \times 1)\\ &= 11 \times 8 \times 6 \times 4! \times \text{équ}(3,2,1,1).\end{aligned}$$

We now recall the hook length formula, a proof of which can be found in Chapter 3 of [Mé17].

Proposition 3.3.1 (Hook-length formula)**.** If λ is a partition of some integer n, then $d_\lambda = \frac{n!}{\text{équ}(\lambda)}$. In particular, $d_{(n-j,\lambda_2,\lambda_3,\ldots)} \leq \binom{n}{j} d_{(\lambda_2,\lambda_3,\ldots)}$.

If $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ is an integer partition, we will denote by λ^* the truncated partition $(\lambda_2, \lambda_3, \lambda_4, \ldots)$, where the largest row has been removed. For example if $\lambda = (n-7, 3, 2, 1, 1)$, $\lambda^* = (3, 2, 1, 1)$ and in this case we have when $n \to \infty$,

$$\begin{aligned}d_\lambda &= \frac{n!}{(n-7+4)(n-8+2)(n-9+1)(n-10)!}\frac{1}{\text{équ}(\lambda^*)}\\ &= \frac{n!}{(n-7)!\,\text{équ}(\lambda^*)}(1 - \frac{7}{n} + O(\frac{1}{n^2})).\end{aligned}$$

This can be easily generalized and gives the following asymptotic formula.

Proposition 3.3.2 (Asymptotic hook-length formula)**.** Let $j \geq 1$ and λ_2, λ_3, ... be fixed integers such that $\lambda_2 + \lambda_3 + \cdots = j$. Then when $n \to \infty$,

$$d_{(n-j,\lambda_2,\lambda_3,\ldots)} = \binom{n}{j} d_{(\lambda_2,\lambda_3,\ldots)}(1 - \frac{j}{n} + O(\frac{1}{n^2})).$$

Proof. Let $n \in \mathbb{N}^*$ and $\lambda = \lambda(n) = (n-j, \lambda_2, \lambda_3, \ldots)$. Then when $n \to \infty$, denoting by $\lambda^{*'}$ the conjugated partition of the partition $\lambda^* = (\lambda_2, \lambda_3, \ldots)$,

$$\begin{aligned}
d_{(n-j,\lambda_2,\lambda_3,\ldots)} &= \frac{n!}{(n-j+\lambda_1^{*\prime})(n-j-1+\lambda_2^{*\prime})\cdots(n-2j+1+\lambda_j^{*\prime})}\frac{1}{\text{équ}(\lambda_2,\lambda_3,\ldots)} \\
&= \frac{n!}{(n-j)!\,\text{équ}(\lambda_2,\lambda_3,\ldots)}\frac{n-j}{n-j+\lambda_1^{*\prime}}\frac{n-j-1}{n-j-1+\lambda_2^{*\prime}}\cdots\frac{n-2j+1}{n-2j+1+\lambda_j^{*\prime}} \\
&= \frac{n!}{(n-j)!\,\text{équ}(\lambda_2,\lambda_3,\ldots)}(1-\frac{\lambda_1^{*\prime}}{n}+O(\frac{1}{n^2}))\cdots(1-\frac{\lambda_j^{*\prime}}{n}+O(\frac{1}{n^2})) \\
&= \binom{n}{j} d_{(\lambda_2,\lambda_3,\ldots)}(1-\frac{j}{n}+O(\frac{1}{n^2})).
\end{aligned}$$

□

3.3.2 Character ratios

Let τ be a transposition. We define as in [DS81] the character ratio $r(\lambda) = \frac{\mathrm{ch}^\lambda(\tau)}{d_\lambda}$. We can give different explicit formulas for this object, among which is the following symmetric one, which follows from Lemma 7.14 in [Mé17].

If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a partition of the integer n, then we have

$$r(\lambda) = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n} \binom{\lambda_i}{2} - \binom{\lambda_i'}{2},$$

where λ' denotes the conjugate partition of λ.

The modified character ratio, as defined in Section 3.2, writes as $s_\lambda = \frac{1}{n} + \frac{n-1}{n} r(\lambda)$ and takes into account that we pick the identity with probability $1/n$. The following upper bounds are given in [Dia88].

Proposition 3.3.3. If λ is a partition of the integer n, then

$$s_\lambda \leq \frac{\lambda_1}{n} \quad \text{and} \quad |r(\lambda)| \leq \frac{\lambda_1}{n}.$$

Moreover, if $\lambda_1 \geq \frac{n}{2}$, then

$$s_\lambda \leq 1 - \frac{2(\lambda_1+1)(n-\lambda_1)}{n^2}.$$

We will also need an asymptotic expansion of s_λ, easily obtainable from the explicit formula for $r(\lambda)$: If $j \in \mathbb{N}^*$ and $\lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_r$ are non-negative integers such that $\lambda_2 + \cdots + \lambda_r = j$, then when $n \to \infty$,

$$r(n-j, \lambda_2, \ldots, \lambda_r) = \frac{1}{\binom{n}{2}}(\binom{n-j}{2} + O(1)) = 1 - \frac{2j}{n} + O(\frac{1}{n^2}),$$

and so

$$s_\lambda = 1 - \frac{2j}{n} + O(\frac{1}{n^2}). \tag{3.3.1}$$

Remark 3.3.1. In the general case, to guess the cutoff time, we want to find a $t = t(n)$ for which $d_\alpha |s_\alpha|^t = \theta(1)$ as $n \to \infty$, for the representations α which have the most mass. In the case of the symmetric group, as $d_\lambda \approx n^j$, we want to find t such that $|s_\lambda|^t \approx n^{-j}$. For instance, for random transpositions, it is very natural to expect a cutoff at $\frac{1}{2} n \log(n)$ from the formula of s_λ, as $\left(1 - \frac{2j}{n}\right)^{\frac{1}{2} n \log(n)} \approx n^{-j}$.

3.3.3 Mass transfer in the Young graph

It will be convenient to use the formalism of the Young graph for some calculations. We will need them in Lemma 3.4.3, to study the behavior of the main term.

Let us first recall the definition of a Young diagram, quoting [Mé17]. Given an integer partition λ, we call Young diagram of λ the array of boxes with λ_1 boxes on the first row, λ_2 boxes on the second row, etc.

We will write $\lambda \vdash m$ for some $m \geq 1$ to indicate that λ is a partition of the integer m. We will also write $\lambda \nearrow \Lambda$ if $\lambda \vdash m$ and $\Lambda \vdash m+1$ to say that the diagram of Λ can be obtained from the diagram of λ by adding a box.

The Young graph is the infinite oriented graph of all integer partitions with an edge from λ to Λ if and only if $\lambda \nearrow \Lambda$. We are going to study, in the Young graph, a measure transfer from a row to the next one, which can be extended by recurrence to several lines.

Let us fix an integer $j \geq 1$. We recall the transition formula for the dimensions of diagrams, which we can find in [Ker03] or [Mé17]. If we fix $\lambda \vdash j$, then we have the following transfer, which may be of independent interest:

$$\sum_{\Lambda : \lambda \nearrow \Lambda} d_\Lambda = (j+1) d_\lambda.$$

Let j be an integer and $(\gamma_\lambda)_{\lambda \vdash j}$ a sequence of real numbers. We extend this line to the next line, $j+1$, as follows, following the edges of the graph. If $\Lambda \vdash j+1$, we set $\gamma_\Lambda = \sum_{\lambda \nearrow \Lambda} \gamma_\lambda$. Then we have the following transfer.

Proposition 3.3.4.

$$\sum_{\Lambda \vdash j+1} \gamma_\Lambda d_\Lambda = (j+1) \sum_{\lambda \vdash j} \gamma_\lambda d_\lambda.$$

Proof.

$$\begin{aligned}
\sum_{\Lambda \vdash j+1} \gamma_\Lambda d_\Lambda &= \sum_{\Lambda \vdash j+1} \Big(\sum_{\lambda : \lambda \nearrow \Lambda} \gamma_\lambda \Big) d_\Lambda \\
&= \sum_{\lambda \vdash j} \sum_{\Lambda \vdash j+1} \mathbf{1}_{\lambda \nearrow \Lambda} \gamma_\lambda d_\Lambda \\
&= \sum_{\lambda \vdash j} \gamma_\lambda \sum_{\Lambda : \lambda \nearrow \Lambda} d_\Lambda \\
&= (j+1) \sum_{\lambda \vdash j} \gamma_\lambda d_\lambda.
\end{aligned}$$

□

3.3.4 Permutations usually do not have only little cycles

We set, for $n \in \mathbb{N}^*$ and $1 \leq j \leq n$,

$$\mathfrak{S}_{n,j} = \{\sigma \in \mathfrak{S}_n : \text{all the cycles of } \sigma \text{ are of length} \leq j\}.$$

Let us show that when j is fixed, $\mathfrak{S}_{n,j}$ is asymptotically much smaller than $\mathfrak{S}_n$.

We will use the next proposition in the proof of Lemma 3.4.2.

Proposition 3.3.5. Let $j \geq 2$ be a fixed integer. Then for n large enough,

$$\log(\frac{|\mathfrak{S}_{n,j}|}{|\mathfrak{S}_n|}) \leq \frac{-n\log(n)}{\mathfrak{T}(j)},$$

where $\mathfrak{T}(j) = 1 + 2 + \cdots + j$.

Proof. We can see that in $\mathfrak{S}_{n,j}$, there are at most $(n+1)^j$ conjugacy classes, because such a conjugacy class is determined by the number of fixed points, 2-cycles, ..., j-cycles of a representative, each one necessarily between 0 and n. Let us give an upper bound on the cardinality of such a class. Let $n \geq j$ be a large integer, $\mu = (\mu_1, \ldots, \mu_r)$ a partition of the integer n such that $\mu_1 \leq j$ and $\mu_r \geq 1$, and $\mathcal{C}_\mu$ the associated conjugacy class. Then if k_q denotes the number of μ_i equal to q, we have for n big enough:

$$\begin{aligned}
|\mathcal{C}_\mu| &= \frac{n!}{2^{k_2}3^{k_3}\cdots j^{k_j}k_2!k_3!\cdots k_j!(n-2k_2-3k_3-\cdots-jk_j)!} \\
&\leq \frac{n!}{k_2!k_3!\cdots k_j!(n-2k_2-3k_3-\cdots-jk_j)!} \\
&= \binom{n}{(2k_2,\ldots,jk_j,(n-2k_2-\cdots-jk_j))}\frac{(2k_2)!}{k_2!}\cdots\frac{(jk_j)!}{k_j!} \\
&\leq \binom{n}{(\frac{n}{j},\frac{n}{j},\ldots,\frac{n}{j})}\frac{(2k_2)!}{k_2!}\cdots\frac{(jk_j)!}{k_j!} \\
&\leq j^n\frac{(2k_2)!}{k_2!}\cdots\frac{(jk_j)!}{k_j!}.
\end{aligned}$$

Moreover this latest product will be greater if the k_i increase, so we can assume without loss of generality that $2k_2 + \cdots + jk_j \geq n-1$. One of the k_i is therefore necessarily of cardinal greater than $\frac{n-1}{2+3+\cdots+j} = \frac{n-1}{T(j)-1}$. Furthermore, as $(2k_2)!\cdots(jk_j)! \leq n!$, we obtain:

$$|\mathcal{C}_\mu| \leq j^n\frac{n!}{(\frac{n-1}{\mathfrak{T}(j)-1})!}.$$

Thus for n large enough,

$$\frac{|\mathfrak{S}_{n,j}|}{|\mathfrak{S}_n|} \leq (n+1)^j j^n \frac{1}{(\frac{n-1}{\mathfrak{T}(j)-1})!}, \tag{3.3.2}$$

i.e.

$$\log(\frac{|\mathfrak{S}_{n,j}|}{|\mathfrak{S}_n|}) \leq j\log(n+1) + n\log(j) - \log((\frac{n-1}{\mathfrak{T}(j)-1})!) \sim -\frac{n\log(n)}{\mathfrak{T}(j)-1}.$$

As $\mathfrak{T}(j) - 1 < \mathfrak{T}(j)$, this leads to the desired asymptotic upper bound. $\square$

Remark 3.3.2. This upper bound proves in particular that the ratio $\frac{|\mathfrak{S}_{n,j}|}{|\mathfrak{S}_n|}$ tends to 0, even multiplied by any power function, or polynomial. It is this fact that we will use. The case $j = 1$ that we did not process is trivial because in this case $|\mathfrak{S}_{n,1}| = 1$.

Besides, if we had proceeded more carefully, we could have shown that $k_j \sim \frac{n}{j}$ maximizes the heavy terms of the cardinality of the conjugacy class, and therefore that $\log(|\mathfrak{S}_{n,j}|) \sim (1-\frac{1}{j})n\log(n)$.

3.3.5 Upper bound on the number of q-cycles

For every permutation $\sigma \in \mathfrak{S}_n$ and $q \in \mathbb{N}^*$, let $N_q(\sigma) = N_q^{(n)}(\sigma)$ denote the number of q-cycles in the cycle decomposition of σ. We recall the well-know law for the number of fixed points of a random permutation[3]

$$\mathbb{P}\left(\sigma \in \mathfrak{S}_n : N_1(\sigma) = m\right) = \frac{1}{m!} \sum_{i=0}^{n-m} \frac{(-1)^i}{i!}, \quad 0 \leq m \leq n. \tag{3.3.3}$$

In particular, we deduce that for all $0 \leq m \leq n$, $\mathbb{P}(\sigma \in \mathfrak{S}_n : N_1(\sigma) = m) \leq \frac{1}{m!}$. Now we generalize this upper bound to the number of q-cycles.

Proposition 3.3.6. Let $q, m \in \mathbb{N}^*$, then

$$\mathbb{P}\left(\sigma \in \mathfrak{S}_n : N_q(\sigma) = m\right) \leq \frac{1}{q^m m!}.$$

Proof. As in the previous paragraph, if μ_i is a partition of the integer n, we denote by k_q the number of μ_i equal to q.

$$\begin{aligned}
&\mathbb{P}\left(\sigma \in \mathfrak{S}_n : N_q(\sigma) = m\right) \\
&\quad = \frac{1}{n!} \sum_{\substack{\mu \vdash n \\ k_q = m}} |C_\mu| \\
&\quad = \sum_{r=1}^{\infty} \sum_{\substack{\mu = (\mu_1, \ldots, \mu_r \geq 1) \vdash n \\ k_q = m}} \frac{1}{2^{k_2} 3^{k_3} \cdots r^{k_r} k_2! k_3! \cdots k_r! (n - 2k_2 - 3k_3 - \cdots - rk_r)!} \\
&\quad = \frac{1}{q^m m!} \\
&\qquad \times \sum_{r=1}^{\infty} \sum_{\substack{\mu = (\mu_1, \ldots, \mu_r \geq 1) \vdash n - qm \\ k_q = 0}} \frac{1}{2^{k_2} 3^{k_3} \cdots r^{k_r} k_2! k_3! \cdots k_r! (n - qm - 2k_2 - 3k_3 - \cdots - rk_r)!} \\
&\quad \leq \frac{1}{q^m m!} \\
&\qquad \times \sum_{r=1}^{\infty} \sum_{\mu = (\mu_1, \ldots, \mu_r \geq 1) \vdash n - qm} \frac{1}{2^{k_2} 3^{k_3} \cdots r^{k_r} k_2! k_3! \cdots k_r! (n - qm - 2k_2 - 3k_3 - \cdots - rk_r)!} \\
&\quad = \frac{1}{q^m m!} \mathbb{P}\left(\sigma \in \mathfrak{S}_{n-qm} : N_q(\sigma) = 0\right) \\
&\quad \leq \frac{1}{q^m m!}.
\end{aligned}$$

□

3.4 Proof of Theorem 3.1.1

For this whole section, we fix $c \in \mathbb{R}$. We recall that $k = k(n, c) = \lfloor \frac{1}{2} n \log(n) + cn \rfloor$.

[3] For $m = 0$, we apply the inclusion-exclusion principle to $\bigcup_{i=1}^n F_i$, where $F_i = \{\sigma \in \mathfrak{S}_n : \sigma(i) = i\}$, and then generalize for any m.

3.4.1 Bounding the error

The upper bound is similar to the upper bound of the sum appearing in [Dia88] after applying Diaconis–Shahshahani upper bound lemma. However, as we want a more precise result, there will be some additional technical difficulties as c may be negative.

We can observe that the representations of the symmetric group which contribute the most in the sum

$$\mathrm{d}_1\left(P_n^{*k}, U_n\right) = \frac{1}{|\mathfrak{S}_n|}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda\in\widehat{\mathfrak{S}_n}^*} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\right|$$

correspond to partitions with a large first row. We will therefore naturally split according to λ_1. We set for all $M\in\mathbb{N}^*$, and integer n large enough,

$$S_M(n) = \left\{\lambda\in\widehat{\mathfrak{S}_n}^* : \lambda_1 \geq n-M\right\}.$$

From Lemma 3.2.1, we get that for all $M\geq 1$,

$$\left|\mathrm{d}_1\left(P_n^{*k}, U_n\right) - \frac{1}{|\mathfrak{S}_n|}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda\in S_M(n)} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\right|\right| \leq \sum_{\lambda\in\widehat{\mathfrak{S}_n};\lambda_1<n-M} d_\lambda \left|s_\lambda\right|^k.$$

It remains to prove that the right-hand side of this inequality tends to 0 uniformly in n when $M\to\infty$, and to estimate the second term in the left-hand side. Our first task is to bound the error in the approximation.

Lemma 3.4.1 (Upper bound on the remainder)**.** *For all $\epsilon>0$ there exist $M=M(c,\epsilon)\geq 1$ and $n_0=n_0(M)\in\mathbb{N}$ such that if $n\geq n_0$, then*

$$\sum_{\lambda_1\leq n-M} d_\lambda \left|s_\lambda\right|^k \leq \epsilon.$$

Proof. We recall that $s_\lambda = \frac{1}{n}+\frac{n-1}{n}r(\lambda)$. Observe that if λ is a partition of n such that $r(\lambda)\geq 0$, then $r(\lambda')=-r(\lambda)$ and so $s_\lambda = |s_\lambda| \geq |s_{\lambda'}|$. Let us first bound $\sum_{\lambda_1\leq n-1} d_\lambda \left|s_\lambda\right|^k$ splitting the sum into pieces. Note that $\lambda_1=n$ corresponds to $r(\lambda)=1$, that is, to $\lambda=(n)$, the trivial representation, which disappeared when we used the Fourier transform. Likewise, $r(\lambda)=-1$ corresponds to $\lambda=(1^n)$.

$$\begin{aligned}\sum_{r(\lambda)<1} d_\lambda\left|s_\lambda\right|^k &= d_{(1^n)}\left|s_{(1^n)}\right|^k + \sum_{-1<r(\lambda)\leq-\frac{2}{n}} d_\lambda\left|s_\lambda\right|^k + \sum_{-\frac{2}{n}<r(\lambda)<\frac{2}{n}} d_\lambda\left|s_\lambda\right|^k + \sum_{\frac{2}{n}\leq r(\lambda)<1} d_\lambda\left|s_\lambda\right|^k \\ &= S_1+S_2+S_3+S_4.\end{aligned}$$

Let us bound these different pieces separately, using Proposition 3.3.3. The first one is the easiest.

$$\begin{aligned}
S_1 &= (1-\frac{2}{n})^{\lfloor\frac{1}{2}n\log(n)+cn\rfloor} = o(1),\\
S_3 &\leq \sum_{-\frac{2}{n}<r(\lambda)<\frac{2}{n}} d_\lambda(\frac{3}{n})^k \leq (\sum_{\lambda\in\widehat{\mathfrak{S}_n}^*} d_\lambda^2)(\frac{3}{n})^k \leq n!(\frac{3}{n})^{\frac{1}{2}n\log(n)+cn} = o(1),\\
S_2 &= \sum_{\frac{2}{n}\leq r(\lambda)<1} d_\lambda(|s_\lambda|-\frac{2}{n})^k \leq \sum_{\frac{2}{n}\leq r(\lambda)<1} d_\lambda\left|s_\lambda\right|^k(1-\frac{2}{n})^k \leq e^{-\frac{2}{n}(\frac{1}{2}n\log(n)+cn)}S_4 = \frac{e^{-2c}}{n}S_4,
\end{aligned}$$

where we used in the upper bound for S_2 that $|s_\lambda| \leq 1$. If we succeed in proving that S_4 is bounded (in n), then we will be able to conclude that $\sum_{r(\lambda)<1} d_\lambda |s_\lambda|^k$ is bounded (in n). We will bound a sum a little larger than S_4, namely $\sum_{0\leq r(\lambda)<1} d_\lambda |s_\lambda|^k$. Let us begin by a crude bound which will prove useful in the sequel. If $1 \leq j \leq n$, we have

$$\sum_{\lambda_1=n-j} d_\lambda \leq \sum_{\lambda^*\vdash j} \binom{n}{j} d_{\lambda^*} \leq \binom{n}{j} \sqrt{(\sum_{\lambda^*\vdash j} 1^2)(\sum_{\lambda^*\vdash j} d^2_{\lambda^*})} \leq \frac{n^j}{j!}\sqrt{2^j j!} \leq \frac{n^j 2^{j/2}}{\sqrt{j!}}, \qquad (**)$$

where the two first inequalities come from Proposition 3.3.1 and Cauchy–Schwarz, and the before last inequality comes from the fact that each partitions of the integer j can be seen as one of the 2^j subsets of the set with j elements. Therefore we have, using Proposition 3.3.3 (note that $r(\lambda) \geq 0$ implies that $s(\lambda) > 0$)

$$\begin{aligned}
S_4 &\leq \sum_{0\leq r(\lambda)<1} d_\lambda s_\lambda^k \\
&= \sum_{j=1}^{n-1} \sum_{\substack{\lambda_1=n-j \\ 0\leq r(\lambda)<1}} d_\lambda s_\lambda^k \\
&\leq \sum_{j=1}^{\lfloor n/1000\rfloor} (\sum_{\lambda_1=n-j} d_\lambda)(1-\frac{2j(n-j+1)}{n^2})^k + \sum_{j=\lfloor n/1000\rfloor+1}^{n-1} (\sum_{\lambda_1=n-j} d_\lambda)(1-\frac{j}{n})^k \\
&= A_1 + A_2.
\end{aligned}$$

Let us bound A_1. We have, using $(**)$ and $1 + x \leq \exp x$,

$$\begin{aligned}
A_1 &\leq \sum_{j=1}^{\lfloor n/1000\rfloor} \frac{n^j 2^{j/2}}{\sqrt{j!}}(1-\frac{2j(n-j+1)}{n^2})^k \\
&\leq \sum_{j=1}^{\lfloor n/1000\rfloor} \frac{2^{j/2}}{\sqrt{j!}} e^{j\log(n)} e^{-\frac{2j(n-j+1)}{n^2}(\frac{1}{2}n\log(n)+cn)} \\
&= \sum_{j=1}^{\lfloor n/1000\rfloor} \frac{2^{j/2}}{\sqrt{j!}} e^{j\log(n)} e^{-j(1-\frac{j-1}{n})(\log(n)+2c)} \\
&= \sum_{j=1}^{\lfloor n/1000\rfloor} \frac{2^{j/2}}{\sqrt{j!}} e^{-2jc} e^{j(j-1)\frac{\log(n)+2c}{n}}.
\end{aligned}$$

Let $a_j(n)$ be the summand in the right-hand side, and note that

$$\frac{a_{j+1}(n)}{a_j(n)} = \frac{e^{\frac{\log(2)}{2}-2c}}{\sqrt{j+1}} e^{2j\frac{\log(n)+2c}{n}}.$$

As a function of j when n is fixed, this is decreasing until $j = \frac{n}{4(\log(n)+2c)}$ and then increasing. If the first and the last ratios are (strictly) less than 1, then we will have a subgeometric sum, which will hence be bounded. The last ratio, at $\frac{n}{1000}$, is equal to

$$\sqrt{1000}e^{\frac{\log(2)}{2}-2c+\frac{4c}{1000}} n^{\frac{2}{1000}-\frac{1}{2}} \xrightarrow[n\to\infty]{} 0.$$

For the first ratio, we need to be a little more careful. At $j = 1$, we can have a ratio much larger than 1, all the more when c is little (i.e., negative and far from 0). So we will need to split once more and consider the sum starting at a suitably chosen M, depending on c but not on n. Thus, though the convergence is fast in the case of a positive c, already treated by Diaconis and Shahshahani, if c is very negative, we will have to consider a very large amount of terms, and the convergence will be much slower. Let M be such that

$$\frac{e^{\frac{\log(2)}{2}-2c}}{\sqrt{M+1}} \leq \frac{1}{4},$$

and n large enough such that

$$e^{2M\frac{\log(n)+2c}{n}} \leq 2,$$

and that the ratio $\frac{a_{j+1}(n)}{a_j(n)}$ at $j = n/1000$ be less than $1/2$. Then as all the ratios from $j = M$ are less than $1/2$, we have:

$$\sum_{j=1}^{n/1000} a_j(n) \leq \sum_{j=1}^{M} a_j(n) + a_M(n)\sum_{i=1}^{\infty}\frac{1}{2^i} \xrightarrow[n\to\infty]{} \sum_{j=1}^{M}\frac{2^{j/2}}{\sqrt{j!}}e^{-2jc} + \frac{2^{M/2}}{\sqrt{M!}}e^{-2Mc}.$$

Thus, as $c \in \mathbb{R}$ is fixed, A_1 is bounded uniformly in n. Let us now treat A_2, which will be slightly easier.

We observe that for all $j \geq 0$, $j^j \leq j!3^j$, hence by $(**)$,

$$\sum_{\lambda_1=n-j} d_\lambda \leq \frac{n^j 6^{j/2}}{j^{j/2}}.$$

Let j be an integer between $n/1000$ and $n-1$. Then

$$\begin{aligned}
\frac{n^j 6^{j/2}}{j^{j/2}}(1-\frac{j}{n})^k &= \frac{n^j 6^{j/2}}{j^{j/2}}e^{k\log(1-\frac{j}{n})} \\
&\leq \frac{n^j 6^{j/2}}{j^{j/2}}e^{-k(\frac{j}{n}+\frac{j^2}{2n^2})} \\
&\leq \frac{n^j 6^{j/2}}{j^{j/2}}e^{-(\frac{1}{2}\log(n)+c)(j+\frac{n}{2\cdot 10^6})} \\
&= 6^{j/2}e^{\frac{j}{2}\log(\frac{n}{j})}e^{-c(j+\frac{n}{2\cdot 10^6})}e^{-\frac{1}{4\cdot 10^6}n\log(n)} \\
&\leq 6^{j/2}e^{\frac{j}{2}\log(1000)}e^{|c|(j+\frac{n}{2\cdot 10^6})}e^{-\frac{1}{4\cdot 10^6}n\log(n)} \\
&\leq e^{n\frac{\log(6)}{2}}e^{n\log(1000)}e^{|c|(n+\frac{n}{2\cdot 10^6})}e^{-\frac{1}{4\cdot 10^6}n\log(n)} \\
&= e^{Kn-K'n\log(n)},
\end{aligned}$$

where K is a real constant and K' is a positive constant. Thus,

$$A_2 \leq ne^{Kn-K'n\log(n)} \xrightarrow[n\to\infty]{} 0.$$

Now we are able to conclude, using the bounds in the proof for A_1. Let $\epsilon > 0$, and let

$M = M(c,\epsilon) \geq 1$ such that $\frac{e^{\frac{\log(2)}{2}-2c}}{\sqrt{M+1}} \leq \frac{1}{4}$ and $2\frac{2^{M/2}}{\sqrt{M!}}e^{-2Mc} < \epsilon$. Then for n large enough,

$$\begin{aligned}
\sum_{\lambda_1 \leq n-M} d_\lambda \left|s_\lambda\right|^k &\leq S_1 + S_2 + S_3 + \sum_{j=M}^{n/1000} a_j(n) + A_2 \\
&\leq \sum_{j=M}^{n/1000} a_j(n) + o(1) \\
&\leq a_M(n) \sum_{i=0}^{\infty} \frac{1}{2^i} + o(1) \\
&\leq 2\frac{2^{M/2}}{\sqrt{M!}} e^{-2Mc} + o(1) \\
&< \epsilon + o(1) \quad \text{as } n \to \infty.
\end{aligned}$$

□

3.4.2 Polynomial convergence lemma

We now start to estimate the main term.

Lemma 3.4.2. *Let $\ell \in \mathbb{N}^*$. Then when $n \to \infty$,*

$$\frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) \right| = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{j=1}^{\ell} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right| + o(1),$$

where we recall that

$$T_j(z) = \sum_{i=0}^{j} \binom{z}{j-i} \frac{(-1)^i}{i!}.$$

Let us first show how the polynomials T_j, a key element of the proof, arise naturally.

Lemma 3.4.3. *Let $j \in \mathbb{N}^*$ be a fixed integer, and $\sigma \in \mathfrak{S}_n$ a permutation with at least one cycle of length greater*[4] *than j (i.e., $\sigma \in \mathfrak{S}_n \backslash \mathfrak{S}_{n,j}$). Then*

$$\frac{1}{j!} \sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} d_{\lambda^*} \mathrm{ch}^\lambda(\sigma) = T_j\left(\mathrm{Fix}(\sigma)\right),$$

where we recall that $\lambda^* = (\lambda_2, \lambda_3, \lambda_4, \ldots)$ is the truncated partition of $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$.

Proof of Lemma 3.4.3. This proof is combinatorial and strongly relies on the Murnagham–Nakayama rule (Formula 3.10 page 108 in [Mé17]). We first consider $\sigma \in \mathfrak{S}_n \backslash \mathfrak{S}_{n,j}$ as an indeterminate in $\mathrm{ch}^\lambda(\sigma)$ and recall that, for any permutation σ and $q \in \mathbb{N}^*$, $N_q(\sigma)$ is the number of q-cycles in the cycle decomposition of σ. For example, if $\lambda = (n-4, 1, 1, 1, 1)$ and σ has a cycle of length greater than 4, we have, using the Murnagham–Nakayama formula and writing N_i for $N_i(\sigma)$,

$$\mathrm{ch}^\lambda(\sigma) = \binom{N_1}{4} + N_3 N_1 + \binom{N_2}{2} - N_4 - (\binom{N_1}{3} - N_2 N_1 + N_3) + (\binom{N_1}{2} - N_2) - N_1 + 1.$$

[4]It still works for $\sigma \in \mathfrak{S}_n \backslash \mathfrak{S}_{n,j-1}$.

We can observe that $\mathrm{ch}^\lambda(\sigma)$ is a polynomial in $N_1(\sigma) = \mathrm{Fix}(\sigma), N_2(\sigma), \ldots, N_j(\sigma)$. The key observation is that we will be able to compute everything when we take the sum at $\lambda_1 = j$ constant, and that our polynomial, which seemingly has j indeterminates, will in reality be a polynomial in only one variable, $N_1(\sigma)$, the number of fixed points of σ. This comes from the orthogonality of some characters and the mass transfer (Proposition 3.3.4), which will make all the other terms cancel. Let us give a little more details.

For the polynomial algebra $\mathbb{C}[z_1, z_2, \ldots]$, we will not use the canonical basis generated by the z_i^j, but rather the one generated by the $\binom{z_i}{j}$, better suited here.

Let $\sigma \in \mathfrak{S}_n \backslash \mathfrak{S}_{n,j}$. If λ is a partition of n such that $\lambda_1 = n-j$, then the coefficient of $\binom{N_1(\sigma)}{j}$ in $\mathrm{ch}^\lambda(\sigma)$ is naturally the number of ways we can fill the Young diagram of λ^* with all the numbers from 1 to j with line and column growth, that is, the number of standard tableaux of λ^*, which is $d_{\lambda^*} = \mathrm{ch}^{\lambda^*}(Id)$.

More generally, if $j_1, \ldots, j_r \in \mathbb{N}$ are such that $j_1 + 2j_2 + \cdots + rj_r = j$, then the coefficient of

$$\binom{N_1(\sigma)}{j_1}\binom{N_2(\sigma)}{j_2}\cdots\binom{N_r(\sigma)}{j_r}$$

in $\mathrm{ch}^\lambda(\sigma)$ is

$$\mathrm{ch}^{\lambda^*}\left(r^{j_r}, \ldots, 2^{j_2}, 1^{j_1}\right).$$

Thus, by orthogonality of the characters, the coefficient of $\binom{N_1(\sigma)}{j_1}\binom{N_2(\sigma)}{j_2}\cdots\binom{N_r(\sigma)}{j_r}$ in the sum

$$\sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} d_{\lambda^*}\mathrm{ch}^\lambda(\sigma)$$

is

$$\sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} d_{\lambda^*}\mathrm{ch}^{\lambda^*}\left(r^{j_r}, \ldots, 2^{j_2}, 1^{j_1}\right) = \sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} \mathrm{ch}^{\lambda^*}(Id)\mathrm{ch}^{\lambda^*}\left(r^{j_r}, \ldots, 2^{j_2}, 1^{j_1}\right) = 0.$$

By mass transfer (Proposition 3.3.4), we can also observe that for $1 \le j' \le j_1$, if σ has at least j' fixed points (if it has less, the coefficient is zero), the coefficient of

$$\binom{N_1(\sigma)}{j_1 - j'}\binom{N_2(\sigma)}{j_2}\cdots\binom{N_r(\sigma)}{j_r}$$

in the sum

$$\sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} d_{\lambda^*}\mathrm{ch}^\lambda(\sigma)$$

is $(-1)^{j'}$ times $j(j-1)\cdots(j-j'+1)$ times the coefficient of

$$\binom{N_2(\sigma)}{j_2}\cdots\binom{N_r(\sigma)}{j_r}$$

in the sum

$$\sum_{\lambda \in \widehat{\mathfrak{S}_{n-j'}} : \lambda_1 = n-j+j'} d_{\lambda^*}\mathrm{ch}^\lambda(\sigma'),$$

where σ' has j' less fixed points than σ, but as many i-cycles for each $i \geq 2$, coefficient which is zero except when $j_2 = \cdots = j_r = 0$, where it is equal to 1. To summarize, we have shown that

$$\begin{aligned}\frac{1}{j!} \sum_{\lambda \in \widehat{\mathfrak{S}_n} : \lambda_1 = n-j} d_{\lambda^*} \mathrm{ch}^\lambda(\sigma) &= \binom{N_1(\sigma)}{j} - \binom{N_1(\sigma)}{j-1} + \frac{1}{2}\binom{N_1(\sigma)}{j-2} + \cdots + \frac{(-1)^j}{j!} \\ &= T_j\left(\mathrm{Fix}(\sigma)\right).\end{aligned}$$

□

Proof of Lemma 3.4.2. Using the fact that $||a| - |b|| \leq |a-b|$ and the triangle inequality,

$$\begin{aligned}&\left| \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) \right| - \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{j=1}^{\ell} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right| \right| \\ &\leq \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \left| \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) - \sum_{j=1}^{\ell} e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right| \\ &\leq \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sum_{j=1}^{\ell} \left| \Big(\sum_{\lambda_1 = n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) \Big) - e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right|.\end{aligned}$$

Let us now split the sum on $\mathfrak{S}_n$ into two parts, along $\mathfrak{S}_{n,\ell}$ and $\mathfrak{S}_n \backslash \mathfrak{S}_{n,\ell}$, and let us bound each of these two sums separately. We recall that $\mathfrak{S}_{n,\ell} = \{\sigma \in \mathfrak{S}_n :$ all the cycles of σ are of length $\leq \ell\}$. We begin by the sum on $\mathfrak{S}_{n,\ell}$. By triangle inequality,

$$\begin{aligned}S_{\mathfrak{S}_{n,\ell}} &:= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n,\ell}} \sum_{j=1}^{\ell} \left| \sum_{\lambda_1 = n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma) - e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right| \\ &\leq \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n,\ell}} \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} \left(d_\lambda s_\lambda^k \left| \mathrm{ch}^\lambda(\sigma) \right| + \left| e^{-2jc} T_j\left(\mathrm{Fix}(\sigma)\right) \right| \right).\end{aligned}$$

Moreover, $0 \leq s_\lambda \leq 1$, $\mathrm{ch}^\lambda(\sigma) \leq d_\lambda$, and $T_j(\mathrm{Fix}(\sigma)) \leq \sum_{i=0}^{j} \mathrm{Fix}(\sigma)^{j-i} \leq (\ell+1)n^\ell$, so

$$S_{\mathfrak{S}_{n,\ell}} \leq \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n,\ell}} \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} \left(d_\lambda^2 + e^{-2jc}(\ell+1)n^\ell \right).$$

By Proposition 3.3.1, $d_\lambda \leq \binom{n}{j} d_{\lambda^*}$, so

$$S_{\mathfrak{S}_{n,\ell}} \leq \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n,\ell}} \sum_{j=1}^{\ell} \sum_{\lambda_1 = n-j} \Big(\Big(\binom{n}{j} d_{\lambda^*}\Big)^2 + e^{-2jc}(\ell+1)n^\ell \Big).$$

Finally, using $\binom{n}{j} d_{\lambda^*} \leq \frac{n^j}{j!} d_{\lambda^*} \leq n^j \leq n^\ell$ and Equation (3.3.2) in the proof of Proposition 3.3.5,

$$S_{\mathfrak{S}_{n,\ell}} \leq K(\ell, c) n^{2\ell} \frac{|\mathfrak{S}_{n,\ell}|}{|\mathfrak{S}_n|} = o(1),$$

where $K(\ell, c)$ is a constant depending only on l and c.

Let us treat the second sum, which we rewrite using Lemma 3.4.3.

$$\begin{aligned}
&\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}}\sum_{j=1}^{\ell}\left|\Big(\sum_{\lambda_1=n-j} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\Big)-e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|\\
&=\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}}\sum_{j=1}^{\ell}\left|\sum_{\lambda_1=n-j}\Big(d_\lambda s_\lambda^k-e^{-2jc}\frac{d_{\lambda^*}}{j!}\Big)\mathrm{ch}^\lambda(\sigma)\right|\\
&\leq\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}\left|d_\lambda s_\lambda^k-e^{-2jc}\frac{d_{\lambda^*}}{j!}\right|\left|\mathrm{ch}^\lambda(\sigma)\right|.
\end{aligned}$$

Let $1\leq j\leq\ell$, $1\leq r\leq j$, and $\lambda_2\geq\lambda_3\geq\cdots\geq\lambda_r\geq 1$ such that $\lambda_2+\cdots+\lambda_r=j$. Let $\lambda=(n-j,\ldots,\lambda_r)$ so that $\lambda^*=(\lambda_2,\ldots,\lambda_r)$.

From Equation (3.3.1), we have

$$s_\lambda^k=\Big(1-\frac{2j}{n}+O\big(\frac{1}{n^2}\big)\Big)^{\lfloor\frac{1}{2}n\log(n)+cn\rfloor}=n^{-j}e^{-2jc}\Big(1+O\big(\frac{\log(n)}{n}\big)\Big). \tag{3.4.1}$$

Using Equation (3.4.1), Proposition 3.3.2 and $n^{-j}\binom{n}{j}=\frac{1}{j!}(1+O(\frac{1}{n}))$, we deduce that

$$d_\lambda s_\lambda^k=e^{-2jc}\frac{d_{\lambda^*}}{j!}\Big(1+O\big(\frac{\log(n)}{n}\big)\Big),$$

which we can rewrite as

$$d_\lambda s_\lambda^k-e^{-2jc}\frac{d_{\lambda^*}}{j!}=O\big(\frac{\log(n)}{n}\big). \tag{3.4.2}$$

As ℓ is fixed, there are only a finite number of such λ, so the $O(\frac{\log(n)}{n})$ can be used for all these λ simultaneously.

We split the right-hand side according to whether $\max(N_1(\sigma),\ldots,N_\ell(\sigma))$ is larger or smaller than $n^{\frac{1}{2\ell}}$. On the one hand,

$$\begin{aligned}
&\frac{1}{n!}\sum_{\substack{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}\\ \max(N_1(\sigma),\ldots,N_\ell(\sigma))\leq n^{1/(2\ell)}}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}\left|d_\lambda s_\lambda^k-e^{-2jc}\frac{d_{\lambda^*}}{j!}\right|\left|\mathrm{ch}^\lambda(\sigma)\right|\\
&=O\big(\frac{\log(n)}{n}\big)\frac{1}{n!}\sum_{\substack{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}\\ \max(N_1(\sigma),\ldots,N_\ell(\sigma))\leq n^{1/(2\ell)}}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}K(\ell,c)\max\left(N_1(\sigma),\ldots,N_\ell(\sigma)\right)^\ell\\
&=O\big(\frac{\log(n)}{n}\big)\frac{1}{n!}\sum_{\substack{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}\\ \max(N_1(\sigma),\ldots,N_\ell(\sigma))\leq n^{1/(2\ell)}}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}O\left(n^{\frac{1}{2}}\right)\\
&=O\big(\frac{\log(n)}{n^{1/2}}\big).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\frac{1}{n!}\sum_{\substack{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}\\ \max(N_1(\sigma),\ldots,N_\ell(\sigma))>n^{\frac{1}{2\ell}}}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}\left|d_\lambda s_\lambda^k-e^{-2jc}\frac{d_{\lambda^*}}{j!}\right|\left|\mathrm{ch}^\lambda(\sigma)\right|\\
&=\frac{1}{n!}\sum_{\substack{\sigma\in\mathfrak{S}_n\setminus\mathfrak{S}_{n,\ell}\\ \max(N_1(\sigma),\ldots,N_\ell(\sigma))>n^{\frac{1}{2\ell}}}}\sum_{j=1}^{\ell}\sum_{\lambda_1=n-j}O(\frac{\log(n)}{n})K(\ell,c)\max\left(N_1(\sigma),\ldots,N_\ell(\sigma)\right)^\ell\\
&\leq\mathbb{P}\left(\sigma\in\mathfrak{S}_n:\max\left(N_1(\sigma),\ldots,N_\ell(\sigma)\right)>n^{\frac{1}{2\ell}}\right)O(\frac{\log(n)}{n})O\left(n^\ell\right)\\
&\leq\sum_{i=1}^{\ell}\mathbb{P}\left(\sigma\in\mathfrak{S}_n:N_i(\sigma)>n^{\frac{1}{2\ell}}\right)O(\frac{\log(n)}{n})O\left(n^\ell\right)\\
&=O(\frac{1}{(n^{\frac{1}{2\ell}})!})O(\frac{\log(n)}{n})O\left(n^\ell\right)\quad\text{from Proposition 3.3.6}\\
&=o(1).
\end{aligned}$$

□

3.4.3 Neglecting polynomials of high degree

Lemma 3.4.4. *Let $\epsilon>0$. There exists $M_0=M_0(\epsilon,c)$ such that for all $M\geq M_0$ and $n\in\mathbb{N}^*$,*

$$\left|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{M}e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|-\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty}e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|\right|\leq\epsilon.$$

Proof. Let $M,n\in\mathbb{N}^*$. Then we have, using again $||a|-|b||\leq|a-b|$, and splitting the symmetric group according to the number of fixed points of permutations,

$$\begin{aligned}
&\left|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{M}e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|-\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty}e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|\right|\\
&\leq\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=M+1}^{\infty}e^{-2jc}T_j\left(\mathrm{Fix}(\sigma)\right)\right|\\
&=\sum_{r=0}^{\infty}\mathbb{P}\left(\sigma\in\mathfrak{S}_n:N_1(\sigma)=r\right)\left|\sum_{j=M+1}^{\infty}e^{-2jc}T_j(r)\right|\\
&\leq\sum_{r=0}^{\infty}\frac{1}{r!}\sum_{j=M+1}^{\infty}e^{-2jc}\left|T_j(r)\right|\quad\text{from Proposition 3.3.6 again.}
\end{aligned}$$

Now we observe that if $r\geq j$,

$$|T_j(r)|=\left|\sum_{i=0}^{j}\binom{r}{j-i}\frac{(-1)^i}{i!}\right|\leq\sum_{i=0}^{j}\binom{r}{j-i}\left|\frac{(-1)^i}{i!}\right|\leq\sum_{i=0}^{j}\binom{r}{j-i}\leq 2^r,$$

and if $r \leq j$,

$$\frac{1}{r!}\left|T_j(r)\right| = \frac{1}{r!}\left|\sum_{i=j-r}^{j}\binom{r}{j-i}\frac{(-1)^i}{i!}\right| \leq \frac{1}{r!(j-r)!}\sum_{i=j-r}^{j}\binom{r}{j-i} \leq \frac{1}{((\frac{j}{2})!)^2}2^r.$$

We therefore conclude that

$$\begin{aligned}
&\sum_{r=0}^{\infty}\frac{1}{r!}\sum_{j=M+1}^{\infty} e^{-2jc}\left|T_j(r)\right| \\
&\quad= \sum_{j=M+1}^{\infty} e^{-2jc}\sum_{r=0}^{j}\frac{1}{r!}\left|T_j(r)\right| + \sum_{j=M+1}^{\infty}\sum_{r=j+1}^{\infty}\frac{1}{r!}e^{-2jc}\left|T_j(r)\right| \\
&\quad\leq \sum_{j=M+1}^{\infty}\frac{e^{-2jc}}{((\frac{j}{2})!)^2}\sum_{r=0}^{j}2^r + \sum_{j=M+1}^{\infty}\sum_{r=j+1}^{\infty}\frac{1}{r!}e^{2r|c|}2^r \\
&\quad\leq \sum_{j=M+1}^{\infty}\frac{e^{-2jc}}{((\frac{j}{2})!)^2}2^{j+1} + \sum_{j=M+1}^{\infty}\sum_{r=j+1}^{\infty}\frac{1}{r!}e^{2r|c|}2^r \\
&\quad= o(1)
\end{aligned}$$

when $M \to \infty$. □

Before proving the last approximation, let us rewrite the infinite sum inside the absolute values. Let us define

$$f_c : x \mapsto e^{-e^{-2c}}\left(1+e^{-2c}\right)^x - 1.$$

Proposition 3.4.1. Let $N \in \mathbb{N}$. Then

$$\sum_{j=1}^{\infty} e^{-2jc}T_j(N) = f_c(N).$$

Proof. We just need to make a change of variables and swap the two sums.

$$\begin{aligned}
\sum_{j=1}^{\infty} e^{-2jc}T_j(N) &= \sum_{j=1}^{\infty}\sum_{i=0}^{j} e^{-2jc}\binom{N}{j-i}\frac{(-1)^i}{i!} \\
&= \sum_{j=1}^{\infty}\sum_{i=0}^{j} e^{-2jc}\binom{N}{i}\frac{(-1)^{j-i}}{(j-i)!} \\
&= \sum_{j=0}^{\infty}\sum_{i=0}^{j} e^{-2jc}\binom{N}{i}\frac{(-1)^{j-i}}{(j-i)!} \quad - 1 \\
&= \sum_{i=0}^{\infty}\binom{N}{i}e^{-2ic}\sum_{j=i}^{\infty} e^{-2(j-i)c}\frac{(-1)^{j-i}}{(j-i)!} \quad - 1 \\
&= \sum_{i=0}^{N}\binom{N}{i}e^{-2ic}e^{-e^{-2c}} \quad - 1 \\
&= e^{-e^{-2c}}\left(1+e^{-2c}\right)^N - 1.
\end{aligned}$$

□

3.4.4 Conclusion of the proof

Before proving Theorem 3.1.1, let us show where the Poisson law comes from.

Lemma 3.4.5. *When $n \to \infty$, we have:*

$$\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|f_c\left(N_1^{(n)}(\sigma)\right)\right| \xrightarrow[n\to\infty]{} \mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right|,$$

where $\mathrm{Poiss}(1)$ *denotes the Poisson law of parameter* 1.

Proof. As factorials grow much faster than exponentials, and hence than f_c, we have as $n \to \infty$,

$$\begin{aligned}
&\left|\mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right| - \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|f_c\left(N_1^{(n)}(\sigma)\right)\right|\right| \\
&= \left|\sum_{r=0}^{\infty}\frac{e^{-1}}{r!}\left|f_c(r)\right| - \sum_{r=0}^{n}\frac{1}{r!}\left(\sum_{i=0}^{n-r}\frac{(-1)^i}{i!}\right)\left|f_c(r)\right|\right| \\
&= \left|\sum_{r=0}^{n}\frac{1}{r!}\left(\sum_{i=n-r+1}^{\infty}\frac{(-1)^i}{i!}\right)\left|f_c(r)\right| + \sum_{r=n+1}^{\infty}\frac{e^{-1}}{r!}\left|f_c(r)\right|\right| \\
&= o(1).
\end{aligned}$$

□

We are now ready to combine all our estimates.

Proof of Theorem 3.1.1. Let $\epsilon > 0$ and M, n_0 such that for $n \geq n_0$, all the approximations be true up to ϵ. Let $n \geq n_0$.

From Lemma 3.2.1 and Lemma 3.4.1,

$$\left|\mathrm{d}_1\left(P_n^{*k}, U_n\right) - \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda_1\geq n-M} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\right|\right| \leq \epsilon.$$

From Lemma 3.4.2,

$$\left|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{\lambda_1\geq n-M} d_\lambda s_\lambda^k \mathrm{ch}^\lambda(\sigma)\right| - \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{M} e^{-2jc}P_j\left(N_1(\sigma)\right)\right|\right| \leq \epsilon.$$

From Lemma 3.4.4,

$$\left|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{M} e^{-2jc}P_j\left(N_1(\sigma)\right)\right| - \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty} e^{-2jc}P_j\left(N_1(\sigma)\right)\right|\right| \leq \epsilon.$$

From Lemma 3.4.5,

$$\left|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\left|\sum_{j=1}^{\infty} e^{-2jc}P_j\left(N_1(\sigma)\right)\right| - \mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right|\right| \leq \epsilon.$$

Consequently, by triangle inequalities,

$$\left|\mathrm{d}_1\left(P_n^{*k}, U_n\right) - \mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right|\right| \leq 4\epsilon.$$

Thus, we proved that for all $c \in \mathbb{R}$,

$$\mathrm{d}_1\left(P_n^{*k}, U_n\right) \xrightarrow[n\to\infty]{} \mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right|.$$

To conclude, let us rewrite this expectation into the natural form of the wording.

$$\begin{aligned}
&\mathbb{E}\left|f_c\left(\mathrm{Poiss}(1)\right)\right| \\
&= \sum_{r=0}^{\infty} \frac{e^{-1}}{r!}\left|e^{-e^{-2c}}\left(1+e^{-2c}\right)^r - 1\right| \\
&= \sum_{r=0}^{\infty} \left|\frac{\left(e^{1+e^{-2c}}\right)^{-1}}{r!}\left(1+e^{-2c}\right)^r - \frac{e^{-1}}{r!}1^r\right| \\
&= \mathrm{d}_1\left(\mathrm{Poiss}\left(1+e^{-2c}\right), \mathrm{Poiss}(1)\right).
\end{aligned}$$

This concludes the proof of Theorem 3.1.1. □

Acknowledgements

I am very thankful to my former professor and advisor, Justin Salez, who introduced me to mixing times and then took great care of me during my master thesis. I would also like to thank Nathanaël Berestycki, for his hospitality when he invited me to the University of Vienna, and for numerous helpful suggestions.

3.5 A detailed proof of the polynomial character identity

In this section we present a detailed proof of Lemma 3.4.3. We thank Pierre-Loïc Méliot for expaining to us how to translate abstract formulas on symmetric functions into concrete combinatorial formulas via an exponential specialization in order to obtain Lemma 3.5.1. We also thank Evita Nestoridi and Sam Olesker–Taylor for helpful comments on this proof.

3.5.1 Preliminaries

In this section, we briefly present some definitions and statements from [Mé17, Section 2], related to symmetric functions. We denote by $\mathfrak{Y}_n$ the set of partitions of the integer $n \in \mathbb{N}$, and by $\mathfrak{Y} := \cup_{n\in\mathbb{N}}\mathfrak{Y}_n$ the set of all integer partitions. In the following, $X = \{x_1, x_2, ...\}$ is an infinite set of (independent) commuting variables (that is $x_i x_j = x_j x_i$ for all i, j).

Definition 3.5.1. ([Mé17, Definition 2.7]) We call *symmetric function* a power series $P \in \mathbb{C}\left[\left[x_1, x_2, ...\right]\right]$, bounded in degree, such that for any bijection $\sigma : \mathbb{N}^* \to \mathbb{N}^*$,

$$P(x_1, x_2, ...) = P(x_{\sigma(1)}, x_{\sigma(2)}, ...). \tag{3.5.1}$$

There are several bases of the algebra of symmetric functions $\mathbb{C}\left[\left[x_1, x_2, ...\right]\right]$. We will use the power sums $p_\lambda(X)$ and the Schur functions $s_\lambda(X)$. We refer to [Mé17, Section 2.2] for definitions. Let also, for $\mu \vdash n \in \mathbb{N}^*$, C_μ be the conjugacy class of $\mathfrak{S}_n$ of cycle type μ, and $z_\mu = n!/|C_\mu|$. We will need two identities which relate those families of symmetric functions. We also refer to [Mé17, Section 3.1] for the definitions of ribbon and ribbon height.

Theorem 3.5.1. *(Frobenius–Schur isomorphism, [Mé17, Theorem 2.32]) For every $n \in \mathbb{N}^*$ and $\lambda \in \mathfrak{Y}(n)$,*

$$s_\lambda(X) = \sum_{\mu \in \mathfrak{Y}(n)} \mathrm{ch}^\lambda(\mu) \frac{p_\mu(X)}{z_\mu}. \tag{3.5.2}$$

Theorem 3.5.2. *(Pieri rule, [Mé17, Theorem 3.9]) For every $k \geq 1$ and partition $\lambda \in \mathfrak{Y}$ (allowed to be empty), we have*

$$p_k(X) s_\lambda(X) = \sum_{\Lambda} (-1)^{\mathrm{ht}(\Lambda \backslash \lambda)} s_\Lambda(X), \tag{3.5.3}$$

where the sum is over partitions Λ such that $\Lambda \backslash \lambda$ is a ribbon of size k.

3.5.2 Mass transfer and mass compensations

Lemma 3.5.1. *Let n be an integer, and $\lambda \vdash n$. Then*

$$\sum_{\Lambda \,:\, \lambda \nearrow \Lambda} \dim \Lambda = (n+1) \dim \lambda, \tag{3.5.4}$$

and if $k \geq 2$,

$$\sum_{\Lambda \,:\, \Lambda \backslash \lambda \textit{ ribbon of size } k} (-1)^{\mathrm{ht}(\Lambda \backslash \lambda)} \dim \Lambda = 0. \tag{3.5.5}$$

Proof. By [Sta99, Proposition 7.8.4 b)] we can take the exponential specialization (with $t = 1$), that is instead of X we can use an alphabet E such that $p_1(E) = 1$ and $p_i(E) = 0$ for $i \geq 2$. Theorem 3.5.1 then becomes

$$s_\lambda(E) = \mathrm{ch}^\lambda(1^n) \frac{p_{1^n}}{z_{1^n}} = \frac{\dim \lambda}{n!}. \tag{3.5.6}$$

Injecting this in Theorem 3.5.2, we get for $k \geq 1$, and λ of size n,

$$\delta_{k,1} \frac{\dim \lambda}{n!} = \sum_{\Lambda} (-1)^{\mathrm{ht}(\Lambda \backslash \lambda)} \frac{\dim \Lambda}{(n+k)!}. \tag{3.5.7}$$

The result follows, noting that if $k = 1$ we have $(-1)^{\mathrm{ht}(\Lambda \backslash \lambda)} = 1$. □

Example 3.5.2. Consider the diagram $\lambda = [1^4]$. There are three possibilities to add a ribbon of size 2 to λ: we can add it at the top, leading to the diagram $\Lambda_1 = [1^6]$, and on the right, vertically or horizontally, leading to $\Lambda_2 = [2^2, 1^2]$ and $\Lambda_3 = [3, 1^3]$. The dimensions of $\Lambda_1, \Lambda_2, \Lambda_3$ are 1, 9, 10, and the height-sign of the added ribbons (the coefficient $(-1)^{\mathrm{ht}(\Lambda \backslash \lambda)}$) are -1,-1,+1, respectively.The sum in (3.5.5) is therefore $(-1) + (-9) + (+10) = 0$.

We well need an iterated version of Lemma 3.5.1.

3.5.3 Polynomial coefficients in characters

Let $j \geq 0$, $\sigma \in \mathfrak{S}_n$, and $\lambda \vdash n$ such that $\lambda_1 = n - j$. In particular, the truncated partition λ^* of λ is of size j. Let $\mu \vdash n$ be the cycle type of σ. Since characters depend only on the conjugacy class (i.e. of the cycle type) of σ, we may write $\mathrm{ch}^\lambda(\mu)$ for $\mathrm{ch}^\lambda(\sigma)$. We refer to [Mé17, Section 3.1] for the definitions of ribbon tableau and height of a ribbon tableau. We call $\mathrm{RT}(\lambda, \mu)$ the set of ribbon tableaux of shape λ and weight μ.

The Murnaghan–Nakayama formula for $\mathrm{ch}^\lambda(\mu)$ writes as

$$\mathrm{ch}^\lambda(\mu) = \sum_{m\in \mathrm{RT}(\lambda,\mu)} (-1)^{\mathrm{ht}(m)}. \tag{3.5.8}$$

Assume now that σ has a cycle of length at least $j+1$, i.e. that $\mu_1 \geq j+1$. A key observation is that in this case, calling m_i the i-th ribbon in m (which is of size μ_i), the diagram $\lambda\backslash m_1$ is disconnected: if $\mathrm{ht}(m_1) = \ell$, then $\lambda\backslash m_1$ consists of two diagrams, of shape $\lambda^*\backslash[1^\ell]$ (which is a skew diagram if $\ell \geq 1$ and is potentially disconnected itself), and a line $[\lambda_1 - \mu_1 + \ell]$. We can therefore rewrite, setting $\xi_\ell := [\mu_1 - \ell, 1^\ell]$,

$$\mathrm{ch}^\lambda(\mu) = \sum_{\ell=0}^{\lambda_1'-1} (-1)^\ell \sum_{m'\in \mathrm{RT}(\lambda\backslash\xi_\ell,\mu^*)} (-1)^{\mathrm{ht}(m')}. \tag{3.5.9}$$

Let $0 \leq \ell \leq \lambda_1'$. Call $\lambda_{up}^{(\ell)}$ the part of $\lambda\backslash\xi_\ell$ above the first line, and $\lambda_{down}^{(\ell)}$ the part of $\lambda\backslash\xi_\ell$ on the first line. Since $\lambda_{down}^{(\ell)}$ is simply a horizontal line, all its possible ribbon tableaux have height 0. The height of a ribbon tableau of $\lambda\backslash\xi_\ell$ is therefore determined solely by the ribbon configuration in $\lambda_{up}^{(\ell)}$.

Recall that we can write any partition $\nu = (\nu_1\nu_2, ...)$ also under the form $\nu = [1^{a_1(\nu)}, 2^{a_2(\nu)}, ...]$, where $a_i(\nu) := \#\{k \geq 1 : \nu_k = i\}$.

If $\mu_{up} \vdash j - \ell$, and $m_{up} \in \mathrm{RT}(\lambda_{up}^{(\ell)}, \mu_{up})$, the number of ribbon tableaux $m' \in \mathrm{RT}(\lambda\backslash\xi_\ell, \mu^*)$ which have the same shape (as an *ordered* union of ribbons) as m_{up} in $\lambda_{up}^{(\ell)}$ is exactly

$$\prod_{i\geq 1} \binom{a_i(\mu^*)}{a_i(\mu_{up})} = \prod_{i\geq 1} \binom{a_i(\mu)}{a_i(\mu_{up})}. \tag{3.5.10}$$

The equality above holds since $|\mu_{up}| \leq j$ and $\mu_1 \geq j+1$, so $a_{\mu_1}(\mu_{up}) = 0$, that is $\binom{a_{\mu_1}(\mu)}{a_{\mu_1}(\mu_{up})} = \binom{1}{0} = 1$. Note also that we can extend the first sum over ℓ in (3.5.9) up to j since if $\ell > \lambda_1' - 1$, then the sum over m_{up} is empty. We can therefore rewrite,

$$\mathrm{ch}^\lambda(\mu) = \sum_{\ell=0}^{j} (-1)^\ell \sum_{\mu_{up}\vdash j-\ell} \prod_{i\geq 1} \binom{a_i(\mu)}{a_i(\mu_{up})} \sum_{m_{up}\in \mathrm{RT}(\lambda_{up}^{(\ell)},\mu_{up})} (-1)^{\mathrm{ht}(m_{up})}, \tag{3.5.11}$$

3.5.4 Building diagrams with ribbons

Let $j \geq 0$ and $\sigma \in \mathfrak{S}_n$ with a cycle of length at least $j+1$. Denote again the cycle type of σ by $\mu \vdash n$. We set

$$S_j(\sigma) := \sum_{\lambda\in\widehat{\mathfrak{S}_n}\,:\,\lambda_1 = n-j} d_{\lambda^*}\mathrm{ch}^\lambda(\sigma). \tag{3.5.12}$$

In (3.5.10), we rewrote the character $\mathrm{ch}^\lambda(\mu)$ as a big sum. We now plug this formula into $S_j(\sigma)$, and exchange the sums. This gives

$$\begin{aligned} S_j(\sigma) &= \sum_{\lambda\,:\,\lambda_1=n-j} d_{\lambda^*} \sum_{\ell=0}^{j} (-1)^\ell \sum_{\mu_{up}\vdash j-\ell} \prod_{i\geq 1} \binom{a_i(\mu)}{a_i(\mu_{up})} \sum_{m_{up}\in \mathrm{RT}(\lambda_{up}^{(\ell)},\mu_{up})} (-1)^{\mathrm{ht}(m_{up})} \\ &= \sum_{\ell=0}^{j} (-1)^\ell \sum_{\nu\vdash j} d_\nu \sum_{\mu_{up}\vdash j-\ell} \prod_{i\geq 1} \binom{a_i(\mu)}{a_i(\mu_{up})} \sum_{m_{up}\in \mathrm{RT}(\nu\backslash[1^\ell],\mu_{up})} (-1)^{\mathrm{ht}(m_{up})}. \end{aligned} \tag{3.5.13}$$

We now want to look at the inner part of the sum iteratively, starting at $[1^\ell]$ and completing it with ribbons until it has size j. We call ribbon tableau completion of a diagram $\tilde{\lambda}$, with weight $\tilde{\mu}$, a sequence of ribbons $m = (m_1, m_2, ...)$ such that $|m_i| = \tilde{\mu}_i$, and for each i, $\tilde{\lambda} \cup m_1 \cup ... \cup m_i$ is a Young diagram. With a small abuse of notations we will write $\mathrm{dgr}(m) = \cup m_i$ for the diagram it corresponds to. Note that $m \in \mathrm{RT}(\mathrm{dgr}(m), \tilde{\mu})$. We call $\mathrm{RTC}(\tilde{\lambda}, \tilde{\mu})$ the set of such m's. With this definition we can write

$$S_j(\sigma) = \sum_{\ell=0}^{j} (-1)^\ell \sum_{\mu_{up} \vdash j-\ell} \prod_{i \geq 1} \binom{a_i(\mu)}{a_i(\mu_{up})} \sum_{m_{up} \in \mathrm{RTC}([1^\ell], \mu_{up})} (-1)^{\mathrm{ht}(m_{up})} d_{[1^\ell] \cup \mathrm{dgr}(m)}. \tag{3.5.14}$$

We are almost ready to conclude. Fix $1 \leq \ell \leq j$ and let $\mu_{up} \vdash j - \ell$ such that $\mu_{up} \neq 1^{j-\ell}$. Let $i_0 := \max\{i : \mu_{up,i} \geq 2\}$. We therefore have by the first part of Lemma 3.5.1 iterated $a_1(\mu_{up})$ times,

$$\sum_{m_{up}} (-1)^{\mathrm{ht}(m_{up})} d_{[1^\ell] \cup \mathrm{dgr}(m)} = \sum_{m'_{up}} (-1)^{\mathrm{ht}(m'_{up})} d_{[1^\ell] \cup \mathrm{dgr}(m'_{up})} j^{\downarrow a_1(\mu_{up})}, \tag{3.5.15}$$

where the first sum is over $m_{up} \in \mathrm{RTC}([1^\ell], \mu_{up})$, and the second over $m'_{up} \in \mathrm{RTC}([1^\ell], (\mu_{up,1}, ..., \mu_{up,i_0}))$. Splitting now the sum also according to m_{i_0}, we have

$$\sum_{m_{up}} (-1)^{\mathrm{ht}(m_{up})} d_{\lambda([1^\ell], m)} = \sum_{m'_{up}} (-1)^{\mathrm{ht}(m'_{up})} \left(\sum_{m_{i_0}} (-1)^{\mathrm{ht}(m_{i_0})} \right) j^{\downarrow a_1(\mu_{up})} = 0, \tag{3.5.16}$$

where the three sums are over $m_{up} \in \mathrm{RTC}([1^\ell], \mu_{up})$, $m'_{up} \in \mathrm{RTC}([1^\ell], (\mu_{up,1}, ..., \mu_{up,i_0-1}))$, and $m_{i_0} \in \mathrm{RTC}([1^\ell] \cup \mathrm{dgr}(m'_{up}), \mu_{up,i_0})$. In the last equality above, we used the second part of Lemma 3.5.1, since the term between parentheses is a sum as in (3.5.1), where m_{i_0} is the k-ribbon, $\lambda = [1^\ell] \cup \mathrm{dgr}(m'_{up})$ and the Λ correspond to the $[1^\ell] \cup \mathrm{dgr}(m'_{up}) \cup m_{i_0}$.

The same argument as in (3.5.15) shows that if $\mu_{up} = 1^{j-\ell}$ we have

$$\sum_{m_{up} \in \mathrm{RTC}([1^\ell], \mu_{up})} (-1)^{\mathrm{ht}(m_{up})} d_{\lambda([1^\ell], m)} = d_{[1^\ell]} j^{\downarrow j-\ell} = \frac{j!}{\ell!}. \tag{3.5.17}$$

Plugging (3.5.16) and (3.5.17) into (3.5.14) and since the only non-zero term is $\mu_{up} = 1^{j-\ell}$, we get

$$S_j(\sigma) = \sum_{\ell=0}^{j} (-1)^\ell \binom{a_1(\mu)}{j-\ell} \frac{j!}{\ell!}. \tag{3.5.18}$$

This is the desired result, since $a_1(\mu) = \mathrm{Fix}(\sigma)$.

Chapter 4

Cutoff profiles for quantum Lévy processes and quantum random transpositions

We consider a natural analogue of Brownian motion on free orthogonal quantum groups and prove that it exhibits a cutoff at time $N\ln(N)$. Then, we study the induced classical process on the real line and compute its atoms and density. This enables us to find the cutoff profile, which involves free Poisson distributions and the semi-circle law. We prove similar results for quantum permutations and quantum random transpositions.

4.1 Introduction

Let $(X_N)_{N\in\mathbf{N}}$ be a sequence of irreducible aperiodic finite state Markov chains, $\mu_N(t)$ the distribution of X_N after t steps, and $\mu_N(\infty)$ the stationary measure of X_N. Let also

$$\mathrm{d}_N(t) = \mathrm{d}_{\mathrm{TV}}(\mu_N(t), \mu_N(\infty))$$

be the distance of the process to equilibrium at time t, where the *total variation distance* $\mathrm{d}_{\mathrm{TV}}(\mu, \nu)$ between two probability measures μ and ν on a finite set E is defined by the formula

$$\mathrm{d}_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2}\sum_{x\in E}|\mu(x) - \nu(x)|.$$

Let $(t_N)_{N\in\mathbf{N}}$ be a sequence of times. We say that $(X_N)_{N\in\mathbf{N}}$ exhibits a *cutoff* in total variation distance at time $(t_N)_{N\in\mathbf{N}}$ if for all $\epsilon > 0$,

$$\mathrm{d}_N((1-\epsilon)t_N) \xrightarrow[N\to\infty]{} 1 \quad \text{and} \quad \mathrm{d}_N((1+\epsilon)t_N) \xrightarrow[N\to\infty]{} 0.$$

This means that the convergence to equilibrium occurs through a sharp phase transition, falling rapidly from 1 to 0 around time t_N[1].

To get a better understanding of this phenomenon, one may try to zoom in on the window where the "fall" occurs. The cutoff phenomenon tells us that the width of this window is negligible

[1]We do here (and sometimes in the sequel) a common abuse of notations, not writing the sequence indices. We also do not always write integer parts for random walk times.

with respect to the sequence $(t_N)_{N\in\mathbf{N}}$, and the next step is therefore to find the next significant "higher order term". Here is a way to formalize this. If there exists a sequence $(w_N)_{N\in\mathbf{N}}$ and a continuous function f decreasing from 1 to 0 such that for all $c \in \mathbf{R}$,

$$\mathrm{d}_N(t_N + cw_N) \xrightarrow[N\to\infty]{} f(c),$$

then we say that f is the *cutoff profile* or *limit profile* of $(X_N)_{N\in\mathbf{N}}$.

Computing the cutoff profile is a difficult task in general, but it could already be done for some important families of Markov chains and commonly involves important probability distributions shaping the profile. For instance, for the lazy random walk on the hypercube (which is equivalent to the Ehrenfest Urn) we have by [Voi96, Sal20]

$$\mathrm{d}_N\left(\frac{1}{2}N\ln(N) + cN\right) \xrightarrow[N\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(\mathcal{N}\left(e^{-c},1\right), \mathcal{N}(0,1)\right),$$

involving Gaussian distributions. Similar profiles were found for the dovetail shuffle [BD92], simple exclusion process on the circle [Lac16a], Ehrenfest Urn with multiple urns [NOT21], or Gibbs Sampler [NOT21]. For random transpositions, we have by [Tey20]

$$\mathrm{d}_N\left(\frac{1}{2}(N\ln(N) + cN)\right) \xrightarrow[N\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}\left(1+e^{-c}\right), \mathrm{Poiss}\left(1\right)\right),$$

involving Poisson distributions. The same profile appears also for k-cycles [NOT21].

This last result on random transpositions, by the second-named author, is one of the motivations of the present article, where we endeavour to compute the cutoff profile for some specific processes. One important difference however is that we will not work with finite classical groups, but with *infinite compact quantum groups*.

Compact quantum groups were introduced by S.L. Woronowicz in [Wor98] as a generalization of classical compact groups. In particular, many results from the representation theory of compact groups carry on to this setting, providing tools similar to those used in the study of random transpositions. A recent work of the first-named author [Fre19] showed that indeed, there are natural quantum Markov chains on compact quantum groups exhibiting a cutoff phenomenon in a way paralleling the classical case. However, the cutoff profile was not studied there.

In the present paper, we will push further the study of the cutoff phenomenon for stochastic processes on compact quantum groups in two ways. First, we will consider continuous processes instead of discrete ones and second, we will study and describe the cutoff profiles.

The most natural continuous process on a simple compact Lie group is certainly Brownian motion. Recall that this is the process whose diffusion kernel is the heat kernel corresponding to the canonical Riemannian structure on the group. Unfortunately, for quantum analogues of compact Lie groups there is to our knowledge no canonical Riemannian-like structure available to provide an analogue of the heat kernel. However, a result of M. Liao in [Lia04] shows that if $(g_t)_{t\in\mathbf{R}_+}$ is a Lévy process on a simple compact Lie group which is invariant under the adjoint action, then its infinitesimal generator is the sum of the Laplace-Beltrami operator (which is the infinitesimal generator of Brownian motion) and a "jump part" given by a so-called Lévy measure. It turns out that a similar decomposition also holds for some compact quantum groups. Indeed, F. Cipriani, U. Franz and A. Kula proved in [CFK14] that on the quantum orthogonal group O_N^+, there exists a distinguished process $(\psi_t)_{t\in\mathbf{R}_+}$ such that for any Lévy process which is invariant under the adjoint action, the corresponding infinitesimal generator splits as the sum of the infinitesimal generator of $(\psi_t)_{t\in\mathbf{R}_+}$ and a "jump part" characterized by a Lévy measure. As a consequence, $(\psi_t)_{t\in\mathbf{R}_+}$ can be seen as an analogue of Brownian motion.

Our main result is the computation in Section 4.3 of the cutoff profile for this Brownian motion on the quantum orthogonal group O_N^+, a compact quantum group which can be thought of as analogue of the group $SO(N)$, for which the cutoff phenomenon was proven by P.-L. Méliot in [Mé14]. More precisely, we prove in Theorem 4.3.9 that for any $c \in \mathbf{R}$ and suitable extensions $\widetilde{\mathrm{d}}_N$ of the distances d_N to the quantum setting,

$$\widetilde{\mathrm{d}}_N\left(N\ln(N) + cN\right) \xrightarrow[N\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}, \nu_{\mathrm{SC}}\right).$$

where ν_{SC} denotes the semi-circle distribution and Poiss^+ denotes the free Poisson distribution. It is known that the correspondence between $SO(N)$ (or rather $O(N)$) and O_N^+ has to do, at the probabilistic level, with the Bercovici-Pata bijection [BP99] (for instance the law of fundamental characters change under the liberation procedure according to that bijection, see [BS09, Sec 5]). From that point of view, the appearance of the semi-circle distribution in the cutoff profile is quite satisfying. On the contrary, the appearance of the free Poisson distribution is surprising because it is not a priori a "deformation" of the semi-circle distribution. The picture becomes clearer when written in terms of *free Meixner distributions* (see Section 4.3.2 for the definition) :

$$\widetilde{\mathrm{d}}_N\left(N\ln(N) + cN\right) \xrightarrow[N\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Meix}^+\left(-e^{-c}, 0\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,0)\right).$$

Let us briefly comment on the proof. On the one hand, the quantum group O_N^+ is easier to study than $SO(N)$, because its representation theory is simpler (the underlying combinatorics is essentially that of the representation theory of $SU(2)$). This enables to reduce the problem to the study of a classical process on the interval $[-N, N]$. But this is compensated by an analytic issue which is absent from the classical case : the measure associated with the quantum process has atoms as soon as $c < 0$, hence is not absolutely continuous with respect to the limiting distribution. These issues are the translation of a failure of absolute continuity of the process with respect to the Haar measure, which is a purely quantum phenomenon (see Proposition 4.3.6). As a consequence, our strategy is first to compute the cutoff profile for $c > 0$ in Proposition 4.3.10, where we have absolute continuity and can therefore reduce the problem to the convergence of the densities, and then to guess from it the form of the cutoff profile for $c < 0$. With this in hand and a method inspired from P. Biane in [Bia08], we are then able to compute the measure of the process also for $c < 0$ and prove the convergence to the cutoff profile in Proposition 4.3.14.

In the end of Section 4.3, we investigate other types of convergence and prove that the convergence to the cutoff profile for $c > 0$ also occurs in L^p-norm for all $1 \leqslant p \leqslant \infty$. Let us mention that for $c < 0$, the aforementioned analytic issues enter the picture again, making the very definition of the L^p-norm problematic, but we nevertheless have convergence of the absolutely continuous part. We furthermore investigate analogues of Brownian motion on some homogeneous spaces of O_N^+ called *free real spheres*, the computations essentially boiling down to the previous ones for O_N^+.

The article concludes in Section 4.4 with a second family of examples called the quantum permutation groups and denoted by S_N^+. Despite bearing strong analogies with the classical permutation group S_N justifying its name, S_N^+ is an "infinite" compact quantum group. In particular, it has a well-defined Brownian motion, given by a Lévy-Khintchine decomposition similar to that of O_N^+. After computing its cutoff profile, we turn to a problem which was left open in [Fre19] : the *quantum random transposition walk*. Here, the absolute continuity issue of the orthogonal case becomes critical : the measure of the corresponding classical process always has an atom so that we cannot resort to densities for $c > 0$. We therefore have to resort to another idea, which is to compare the process with the so-called *pure quantum transposition random walk* and prove that they asymptotically coincide. This is a specifically quantum phenomenon connected to the fact that the pure quantum transposition walk has no periodicity issue because there is no quantum alternating group. More precisely, we show in Theorem 4.4.4 that for $c > 0$,

$$\widetilde{\mathrm{d}}_N\left(\frac{1}{2}(N\ln(N) + cN)\right) \xrightarrow[N\to\infty]{} \mathrm{d}_{\mathrm{TV}}\left(D_{\sqrt{1+e^{-c}}}\left(\mathrm{Meix}^+\left(\frac{1-e^{-c}}{\sqrt{1+e^{-c}}}, \frac{-e^{-c}}{1+e^{-c}}\right)\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(1,0)\right).$$

As $\mathrm{Meix}^+(1,0) = \mathrm{Poiss}^+(1,1) * \delta_{-1}$ is the standard free Poisson distribution, this provides a quantum analogue of the result of [Tey20].

Acknowledgments

The authors are indebted to Uwe Franz for pointing out to them the article [Bia08], the ideas of which helped to improve significantly the results of an earlier version of the present work, and to P.-L. Méliot for discussions on topics linked to the subject of the present paper. A.F. and S.W. were partially funded by the ANR grant "Noncommutative analysis on groups and quantum groups" (ANR-19-CE40-0002) and the PHC Polonium "Quantum structures and processes", A.F. was also partially funded by the ANR grant "Operator algebras and dynamics on groups" (ANR-19-CE40-0008), the PHC Procope "Quantum groups and quantum probability" and the PHC Van Gogh "Quantum groups, harmonic analysis and quantum probability". S.W. was also partially supported by a public grant as part of the Fondation Mathématique Jacques Hadamard.

4.2 Preliminaries

Compact quantum groups will be one of our main objects of studies in this work, and the one the probabilist reader may be least acquainted with. We will therefore devote this preliminary section to some definitions and fundamental results concerning them. In order to keep things simple, we will only introduce free orthogonal quantum groups for the moment, as well as some results concerning Lévy processes on them. Details on quantum permutation groups will be given when needed later on.

4.2.1 Free orthogonal quantum groups

Free orthogonal quantum groups are examples of compact quantum groups in the sense of S.L. Woronowicz [Wor98] which were first introduced by Sh. Wang in [Wan95]. The original definition uses C*-algebras, as may be expected for objects of noncommutative topological nature. We will nevertheless use a different definition which we believe may be easier to understand for the non-expert reader, by focusing first on the purely algebraic aspects. We refer to the books [NT13] and [Tim08] for a comprehensive treatment of the theory and proofs of the main results.

4.2.1.1 Definition and representation theory

We recall that a $*$-algebra is an algebra A endowed with an involution $x \mapsto x^*$, i.e. an antimultiplicative linear map such that $(x^*)^* = x$ and $(\lambda x)^* = \bar{\lambda}x^*$ for all $x \in A$ and $\lambda \in \mathbf{C}$. Also, a $*$-ideal B of A is a $*$-subalgebra of A such that $\{ba, ab\} \subset B$ for all $a \in A$ and $b \in B$.

Definition 4.2.1. We define $\mathcal{O}(O_N^+)$ to be the universal $*$-algebra generated by N^2 *self-adjoint* elements u_{ij} (i.e. $u_{ij}^* = u_{ij}$) such that for all $1 \leqslant i, j \leqslant N$,

$$\sum_{k=1}^{N} u_{ik}u_{jk} = \delta_{ij} = \sum_{k=1}^{N} u_{ki}u_{kj}.$$

In other words,
$$\mathcal{O}(O_N^+) = \mathbf{C}\langle u_{ij} : 1 \leqslant i,j \leqslant N\rangle / I,$$
where $\mathbf{C}\langle u_{ij} : 1 \leqslant i,j \leqslant N\rangle$ denotes the $*$-algebra of noncommutative polynomials in variables u_{ij}, u_{ij}^* with $1 \leqslant i,j \leqslant N$, and I denotes the $*$-ideal generated by the elements

$$\left\{ u_{ij}^* - u_{ij}, \sum_{k=1}^{N} u_{ik}u_{jk} - \delta_{ij}, \sum_{k=1}^{N} u_{ki}u_{kj} - \delta_{ij}, 1 \leqslant i,j \leqslant N \right\}.$$

Let O_N be the usual orthogonal group, let $c_{ij} : O_N \to \mathbf{C}$ be the function sending a matrix to its (i,j)-th coefficient and let $\mathcal{O}(O_N)$ be the algebra of *regular functions* on O_N, i.e. the $*$-algebra generated by the functions c_{ij}, where the involution corresponds to the complex conjugation : $c_{ij}^* = \overline{c_{ij}}$. Then, quotienting $\mathcal{O}(O_N^+)$ by its commutator ideal yields a surjection

$$\pi : \mathcal{O}(O_N^+) \to \mathcal{O}(O_N)$$

so that O_N^+ can be seen as a "noncommutative version" of O_N. The group structure can be encoded in this setting thanks to the following remark : for any two orthogonal matrices g and h,

$$c_{ij}(gh) = \sum_{k=1}^{N} c_{ik}(g)c_{kj}(h) = \sum_{k=1}^{N} (c_{ik} \otimes c_{kj})(g,h),$$

where we identify $\mathcal{O}(O_N \times O_N)$ with $\mathcal{O}(O_N)\otimes\mathcal{O}(O_N)$. The "group law" of $\mathcal{O}(O_N^+)$ will therefore be given by the unique $*$-homomorphism $\Delta : \mathcal{O}(O_N^+) \to \mathcal{O}(O_N^+) \otimes \mathcal{O}(O_N^+)$, called the *comultiplication*, such that
$$\Delta(u_{ij}) = \sum_{k=1}^{N} u_{ik} \otimes u_{kj}.$$
The existence of Δ follows from the universal property of $\mathcal{O}(O_N^+)$, since the elements $U_{ij} = \sum u_{ik} \otimes u_{kj}$ satisfy the defining relations of the algebra $\mathcal{O}(O_N^+)$.

Probability measures can be generalized to this setting by identifying them with their integration linear form. They then correspond to *states*, i.e. linear maps

$$\psi : \mathcal{O}(O_N^+) \to \mathbf{C}$$

such that $\psi(1) = 1$ and $\psi(x^*x) \geqslant 0$ for all x. There is a particular state which plays the rôle of the uniform measure on O_N^+ :

Theorem 4.2.2 (Woronowicz)**.** There is a unique state h on $\mathcal{O}(O_N^+)$ such that for all $x \in \mathcal{O}(O_N^+)$,

$$(\mathrm{id} \otimes h) \circ \Delta(x) = h(x) \otimes 1 = (h \otimes \mathrm{id}) \circ \Delta(x).$$

It is called the *Haar state* of O_N^+.

Since the founding works of P. Diaconis and his coauthors, it is known that representation theory is a powerful tool to study the asymptotic behaviour of random walks on groups (see for instance [Dia88, Chap 4]). For O_N^+, the representation theory was computed by T. Banica in [Ban96]. However, for our purpose we will only need to understand the subalgebra $\mathcal{O}(O_N^+)_{\mathrm{central}}$ generated by the characters of the irreducible representations (we refer the reader for instance to [NT13, Sec 1.3] for the definitions of these notions and details).

Theorem 4.2.3 (Banica)**.** Let us set $\chi_0 = 1$ and $\chi_1 = \sum_{i=1}^{N} u_{ii}$. Then, the irreducible representations of O_N^+ are labelled by the integers such that if χ_n denotes the character associated to the integer $n \in \mathbf{N}$, we have the recurrence relation :

$$\forall n \geqslant 1, \quad \chi_1\chi_n = \chi_{n+1} + \chi_{n-1}.$$

Note that this implies that $\chi_n^* = \chi_n$ for all $n \in \mathbf{N}$. This recurrence relation is reminiscent of Chebyshev polynomials, and one can indeed express χ_n in terms of χ_1 using them. More precisely, let $(P_n)_{n\in\mathbf{N}}$ be the sequence of polynomials defined by $P_0(X) = 1$, $P_1(X) = X$ and

$$XP_n(X) = P_{n+1}(X) + P_{n-1}(X).$$

In particular, $P_n(2) = n+1$, $P_n(-2) = (-1)^n(n+1)$, and for $\theta \in]0, \pi[$ and $n \in \mathbf{N}$,

$$P_n(2\cos(\theta)) = \frac{\sin((n+1)\theta)}{\sin(\theta)}.$$

Then, the map $\iota : \chi_n \mapsto P_n$ yields an isomorphism between $\mathcal{O}(O_N^+)_{\text{central}}$ and $\mathbf{C}[X]$. Moreover, by [Ban97, Prop 1], the restriction of the Haar state to this subalgebra coincides with integration with respect to the semicircle distribution ν_{sc}. More precisely, for any $x \in \mathcal{O}(O_N^+)_{\text{central}}$ we have

$$h(x) = \int_{[-2,2]} \iota(x)\mathrm{d}\nu_{\text{sc}},$$

where ν_{sc} denotes the measure on [-2,2] with density

$$\mathrm{d}\nu_{\text{sc}} = \frac{1}{2\pi}\sqrt{4-x^2}\mathrm{d}x.$$

The polynomials P_n are exactly the orthogonal polynomials for this measure. Moreover, denoting by d_n the dimension of the n-th irreducible representation, we have for $n \in \mathbf{N}$,

$$d_n = P_n(N).$$

4.2.1.2 Central Lévy processes

Let us now describe what the analogue of a Lévy process is on O_N^+. On a classical group, this is a càdlàg stochastic process $(X_t)_{t\in\mathbf{R}_+}$ with independent and stationary increments. In particular, if μ_t is the distribution of $X_tX_0^{-1}$ then we have

- $\mu_0 = \delta_{\text{Id}}$,
- $\mu_t * \mu_s = \mu_{t+s}$,
- $\lim\limits_{t\to 0} \mu_t = \mu_0$ weakly.

In other words, we have a right-continuous convolution semigroup of probability measures. Because this semigroup contains most of the probabilistic information about the process, and in particular concerning its asymptotic behaviour, we will focus on it and define a *quantum Lévy process* to be a right-continuous convolution semigroup of states. More precisely, given two states φ, φ' on $\mathcal{O}(O_N^+)$, their *convolution* is defined by $\varphi * \varphi' = (\varphi \otimes \varphi') \circ \Delta$. A family $(\psi_t)_{t\in\mathbf{R}_+}$ of states on $\mathcal{O}(O_N^+)$ is called a right-continuous convolution semigroup of states if it satisfies

- $\psi_0 = \varepsilon$, where $\varepsilon : \mathcal{O}(O_N^+) \to \mathbf{C}$ is the *co-unit*, i.e. the unique $*$-homomorphism determined by $\varepsilon(u_{ij}) = \delta_{ij}$,
- $\psi_t * \psi_s = (\psi_t \otimes \psi_s) \circ \Delta = \psi_{t+s}$,
- $\lim_{t\to 0} \psi_t(x) = \psi_0(x)$ for all $x \in \mathcal{O}(O_N^+)$.

Let us mention that the theory of Lévy processes on compact quantum groups can be developped in full generality (not just restricted to marginals), see for instance the survey [Fra06].

As mentioned in the introduction, we will be interested in the case where the Lévy process is invariant under the adjoint action. In other words we will focus on states which are *central*, in the sense that they are invariant by conjugation (see the beginning of [CFK14, Sec 6] for the definitions). By [CFK14, Prop 6.9], such states have a specific form. First, there is a *conditional expectation*

$$\mathbb{E} : \mathcal{O}(O_N^+) \to \mathcal{O}(O_N^+)_{\text{central}} ,$$

that is to say a linear map satisfying $\mathbb{E}(x^*x) \geqslant 0$ for all x, and $\mathbb{E}(x) = x$ for all $x \in \mathcal{O}(O_N^+)_{\text{central}}$. Then, a state ψ is central if and only if there exists a state $\widetilde{\psi}$ on $\mathcal{O}(O_N^+)_{\text{central}}$ such that

$$\psi = \widetilde{\psi} \circ \mathbb{E}.$$

Definition 4.2.4. A *central Lévy process* on O_N^+ is a continuous convolution semigroup of states $(\psi_t)_{t\in\mathbf{R}_+}$ on $\mathcal{O}(O_N^+)$ such that ψ_t is central for all t.

Central Lévy processes on O_N^+ were classified by F. Cipriani, U. Franz and A. Kula in [CFK14, Thm 10.2] through an analogue of the Lévy-Khinchine formula. Note that because of centrality, it is enough to know the image of χ_n for all $n \in \mathbf{N}$.

Theorem 4.2.5 (Cipriani-Franz-Kula)**.** Any central Lévy process $(\psi_t)_{t\in\mathbf{R}_+}$ on O_N^+ is of the form

$$\psi_t : \chi_n \mapsto P_n(N)e^{-t\lambda_n},$$

where

$$\lambda_n = b\frac{P_n'(N)}{P_n(N)} + \frac{1}{P_n(N)}\int_{-N}^{N} \frac{P_n(N) - P_n(x)}{N - x} \mathrm{d}\nu(x) \tag{4.2.1}$$

for some $b \geqslant 0$ and a finite measure ν on $[-N, N]$ such that $\nu(\{N\}) = 0$.

Comparing this formula with the one proved by M. Liao for classical compact Lie groups in [Lia04], we see that the process corresponding to $\lambda_n = P_n'(N)/P_n(N)$ plays a role analogous to the one associated to the Laplace-Beltrami operator. As a consequence, we will call this process the *Brownian motion* on O_N^+. The term is further justified by a recent result of M. Brannan, M. Junge and L. Gao. They show in [BGJ22, Prop 3.9] that the semigroup $(\psi_t)_{\in\mathbf{R}_+}$ can also be constructed from the usual Brownian motion on O_N (that is to say with infinitesimal generator given by the Laplace-Beltrami operator) through an averaging procedure.

4.2.2 The cutoff phenomenon

This work is mostly concerned with the diffusion of central Lévy processes and in particular the time needed for the process to spread all over the group. This can be rigorously defined by measuring the distance between ψ_t and the Haar state h. Classically, one interesting and widely used distance for this is the *total variation distance*

$$\mathrm{d_{TV}}(\mu, \nu) = \|\mu - \nu\|_{TV} = \sup_{A\subset G} |\mu(A) - \nu(A)|,$$

where the supremum is taken over all Borel subsets A of the classical group G. For quantum groups, the corresponding definition requires the introduction of a suitable version of the Borel σ-algebra.

To this end, one may consider the universal enveloping C*-algebra (see for instance [Bla06, Sec II.8.3]) $C(O_N^+)$ of $\mathcal{O}(O_N^+)$. By definition, any state on $\mathcal{O}(O_N^+)$ has a unique extension to a state on $C(O_N^+)$, hence yields an element of the *Fourier-Stieltjes algebra*, which is the topological dual $C(O_N^+)^*$ of $C(O_N^+)$. In view of the Riesz representation theorem, this dual space is thought of as a noncommutative analogue of the measure algebra equipped with the total variation norm. Let us denote by $\|\cdot\|_{FS}$ the norm on this dual space and call it the *Fourier-Stieltjes norm*. Moreover, the topological double dual $C(O_N^+)^{**}$ of $C(O_N^+)$ is known to be the *universal enveloping von Neumann algebra* of $C(O_N^+)$, which is regarded as a noncommutative and universal analogue of measure spaces. Using the theory of Haagerup's noncommutative L^p-spaces, we may easily adapt the argument in [Fre19, Lemma 2.6] to show that

$$\frac{1}{2}\|\varphi-\psi\|_{FS} = \sup_{p\in\mathcal{P}} |\varphi(p)-\psi(p)|,$$

where $\mathcal{P}$ denotes the set of orthogonal projections in $C(O_N^+)^{**}$ (thought of as indicator functions of Borel subsets). We omit the details of the proof since we do not need it in this paper. We will use this generalized total variation distance in our cutoff statements and write :

$$\|\cdot\| = \frac{1}{2}\|\cdot\|_{FS}.$$

In the classical setting, a particularly important case is that both μ and ν are absolutely continuous with respect to the Haar measure. We may consider the similar situation in the quantum setting. Let us define an inner product on $\mathcal{O}(O_N^+)$ by the formula $\langle x,y\rangle = h(xy^*)$. Then, taking the completion yields a Hilbert space $L^2(O_N^+)$, and $\mathcal{O}(O_N^+)$ embeds through left multiplication into $\mathcal{B}(L^2(O_N^+))$ (see [NT13, Cor 1.7.5] and the comments thereafter). The weak closure of the image is denoted by $L^\infty(O_N^+)$ and is a von Neumann algebra. If $\varphi : \mathcal{O}(O_N^+)\to\mathbf{C}$ is a linear map which extends to a normal bounded map on $L^\infty(O_N^+)$, then φ becomes an element of the *Fourier algebra*, which is the Banach space predual $L^\infty(O_N^+)_*$ of $L^\infty(O_N^+)$ and $\|\varphi\|_{FS} = \|\varphi\|_{L^\infty(O_N^+)_*}$ (see for instance [BR17, Prop 3.14]), which further implies by [Fre19, Lem 2.6]:

$$\|\varphi-\psi\| = \sup_{p\in\tilde{\mathcal{P}}} |\varphi(p)-\psi(p)|,$$

where $\tilde{\mathcal{P}}$ is the set of orthogonal projections in $L^\infty(O_N^+)$. Note that in order for this formula to make sense, the states φ and ψ must extend to the von Neumann algebra $L^\infty(O_N^+)$. This is not always the case due to absolute continuity issues (see for instance Proposition 4.3.6).

Let us mention an elementary but useful fact on the monotonicity of that norm which is well-known in the classical case.

Lemma 4.2.6. *For N fixed, the map $t\mapsto\|\psi_t-h\|$ is decreasing.*

Proof. First note that for any two bounded linear forms φ,ψ on $C(O_N^+)$,

$$\|\varphi*\psi\|_{FS} = \|(\varphi\otimes\psi)\circ\Delta\|_{FS} \leqslant \|\varphi\otimes\psi\|_{FS} \leqslant \|\varphi\|_{FS}\|\psi\|_{FS}.$$

Thus, for any $t\geqslant s$,

$$\|\psi_t-h\|_{FS} = \|(\psi_s-h)*\psi_{t-s}\|_{FS} \leqslant \|\psi_s-h\|_{FS},$$

where we used the fact that any state ψ on a C*-algebra has norm one (see for instance [Tak02, Lem 9.9]). The result follows. □

The evolution of the distance from the process to the Haar state can exhibit various behaviours. One which is especially striking is the so-called *cutoff phenomenon*. Here is a precise definition of what we mean by this :

Definition 4.2.7. Let $(\mathbb{G}_N, (\psi_t^{(N)})_{t\in\mathbf{R}^+})_{N\in\mathbf{N}}$ be a family of compact quantum groups with a Lévy process $(\psi_t^{(N)})_{t\in\mathbf{R}^+}$ on each of them. We say that the processes exhibit a cutoff phenomenon at time $(t_N)_{N\in\mathbf{N}}$ if for any $\epsilon > 0$,

$$\lim_{N\to+\infty} \|\psi^{(N)}_{(1-\epsilon)t_N} - h_N\| = 1 \text{ and } \lim_{N\to+\infty} \|\psi^{(N)}_{(1+\epsilon)t_N} - h_N\| = 0.$$

One very useful tool to prove that such a phenomenon occurs is the following lemma originally due to P. Diaconis and M. Shahshahani in [DS81] for finite groups and to J.P. McCarthy in [McC19] for finite quantum groups. A proof for compact quantum groups can be found in [Fre19, Lem 2.7], but we simply state it in our particular case.

Lemma 4.2.8. *Let ψ be a central state on O_N^+. If for some $t > 0$ the sum $\sum_{n=1}^{+\infty} d_n^2 e^{-2t\lambda_n}$ is finite, then*

$$\|\psi_t - h\|^2 \leqslant \frac{1}{4}\sum_{n=1}^{+\infty} d_n^2 e^{-2t\lambda_n}. \tag{4.2.2}$$

Remark 4.2.9. The right-hand side is nothing but the L^2-norm of the density of $\psi_t - h$ with respect to h, computed using the Plancherel formula. We recover in that way the fact that as soon as a state has an L^2-density with respect to the Haar state, it has an extension to $L^\infty(O_N^+)$.

Let us end these preliminaries with some computational results concerning the Chebyshev polynomials introduced above. First, there is an explicit formula to compute these numbers for $t > 2$: setting

$$q(t) = \frac{t - \sqrt{t^2-4}}{2},$$

it is easily checked by induction that

$$P_n(t) = \frac{q(t)^{-(n+1)} - q(t)^{n+1}}{q^{-1} - q}.$$

This enables us to give a precise expansion of the polynomials P_n.

Proposition 4.2.10. For any $n \geqslant 1$, we have as $t \to \infty$

$$P_n(t) = t^n\left(1 - \frac{1}{t^2} + O\left(\frac{1}{t^4}\right)\right)^{n-1}.$$

Proof. Let $n \geqslant 1$. First, as $t \to \infty$, we have

$$q(t) = \frac{t - \sqrt{t^2-4}}{2} = \frac{1}{t} + \frac{1}{t^3} + O\left(\frac{1}{t^5}\right) \text{ and } q(t)^{-1} = t - \frac{1}{t} + O\left(\frac{1}{t^3}\right).$$

We deduce, writing q for $q(t)$, that for $t \geqslant 2$ we have

$$\ln(P_n(t)) = \ln\left(\frac{q^{-(n+1)}}{q^{-1}}\frac{1-q^{2(n+1)}}{1-q^2}\right) = n\ln(q^{-1}) + \ln\left(1 - q^{2(n+1)}\right) - \ln\left(1-q^2\right).$$

Note that $-\ln(1-q^2) = q^2 + O\left(q^4\right) = \frac{1}{t^2} + O\left(\frac{1}{t^4}\right)$ and $2(n+1) \geqslant 4$ so $\ln\left(1-q^{2(n+1)}\right) = O\left(\frac{1}{t^4}\right)$, hence

$$\ln(P_n(t)) = n\left(\ln(t) - \frac{1}{t^2} + O\left(\frac{1}{t^4}\right)\right) + O\left(\frac{1}{t^4}\right) + \frac{1}{t^2} + O\left(\frac{1}{t^4}\right).$$

Hence, putting the three $O\left(\frac{1}{t^4}\right)$ together (it might change the implicit constant, but we can still take it to be independent of n), we finally have

$$\ln(P_n(t)) = n\ln(t) + (n-1)\left(-\frac{1}{t^2} + O\left(\frac{1}{t^4}\right)\right),$$

i.e.

$$P_n(t) = t^n \exp\left(-\frac{1}{t^2} + O\left(\frac{1}{t^4}\right)\right)^{n-1} = t^n\left(1 - \frac{1}{t^2} + O\left(\frac{1}{t^4}\right)\right)^{n-1}.$$

□

4.3 The quantum orthogonal brownian motion

In this section we study the cutoff phenomenon for the analogue of Brownian motion on O_N^+. As explained above, this means that we will take $\nu = 0$ in Equation (4.2.1). Once that choice is made, changing the value of b is equivalent to rescaling the time, so that there is no loss in generality in fixing $b = 1$ for all N, leading to

$$\lambda_n = \frac{P_n'(N)}{P_n(N)}.$$

4.3.1 The cutoff phenomenon

We first want to prove that the process $(\psi_t)_{t\in\mathbf{R}_+}$ exhibits a cutoff phenomenon. It turns out that this result can be recovered as a consequence of the existence of a cutoff profile, proven in Theorem 4.3.9, as will be explained in Remark 4.3.17 below. Nevertheless, the proof below is independent of the computations of the cutoff profile and is interesting in its own right. One reason for this is that, in particular as far as the lower bound is concerned, the argument below fills a gap in the proof of previous similar theorems in [Fre19] and [Fre18] (see in particular Remark 4.3.2). Another reason is that to compute the profile, one must know the cutoff time and guess what the next order is, which we do when proving the cutoff phenomenon. That being said, let us give some additional details concerning the Fourier-Stieltjes norm needed for the proof.

As explained in Section 4.2.2, proving a cutoff phenomenon requires precise estimates for the Fourier-Stieltjes norm. As soon as the right-hand side of Equation (4.2.2) is finite, we can use it to bound the total variation distance, which is then well-defined and coincides with the half of the Fourier-Stieltjes norm. This is the strategy which was already used in [Fre19] and the computations will be similar. To obtain the lower bound, however, we need to deal directly with the Fourier-Stieltjes norm and this will require an alternate description which we now detail.

By [Bra12, Lem 4.2], the closure of $\mathcal{O}(O_N^+)_{\mathrm{central}}$ in $C(O_N^+)$ is a commutative C*-algebra isomorphic to $C([-N,N])$. Moreover, if $\varphi = \widetilde{\varphi}\circ\mathbb{E}$ is a bounded central linear form, then

$$\|\widetilde{\varphi}\|_{FS} = \left\|\varphi_{|\overline{\mathcal{O}(O_N^+)_{\mathrm{central}}}}\right\|_{FS} \leqslant \|\varphi\|_{FS} = \|\widetilde{\varphi}\circ\mathbb{E}\|_{FS} \leqslant \|\widetilde{\varphi}\|_{FS}$$

so that the problem reduces to the computation of the norm of a bounded linear form on a commutative C*-algebra. By the Riesz Representation Theorem, there exists a measure μ on $[-N, N]$ such that

$$\widetilde{\varphi}(x) = \int_{-N}^{N} x\mathrm{d}\mu$$

and moreover, the Fourier-Stieltjes norm of $\widetilde{\varphi}$ coincides with twice the total variation of μ. Using that observation, we can now establish that Brownian motion on O_N^+ exhibits a cutoff phenomenon.

Theorem 4.3.1. Brownian motion on O_N^+ exhibits a cutoff phenomenon at time $t_N = N\ln(N)$.

Proof. We start with the upper bound and we will obtain an estimate which is finer than what is actually needed. It was proven in [FHL+17, Lem 1.7] that (we will give in Lemma 4.3.5 on a finer estimate, but that one is sufficient for our present purpose)

$$\frac{n}{N} \leqslant \lambda_n \leqslant \frac{n}{N-2}.$$

Using this and the estimates of [Fre19, Lem 3.3], and writing $q = q(N)$ for simplicity, the sum in the right-hand side of Lemma 4.2.8 applied to ψ_t can be bounded as soon as $q^{-1}e^{-t/N} < 1$ by

$$\sum_{n=1}^{+\infty} \frac{q^{-2n}}{(1-q^2)^2} e^{-2tn/N} = \frac{1}{(1-q^2)^2} \frac{q^{-2}e^{-2t/N}}{1-q^{-2}e^{-2t/N}} = \frac{1}{(1-q^2)^2} \frac{1}{q^2 e^{2t/N} - 1}.$$

For $c > 0$ and $t = N\ln(N) + cN$, we get, using $1/2 > q(N) > 1/N$ (see for instance [Fre19, Lem 3.8]),

$$\frac{1}{(1-q^2)^2} \frac{1}{q^2 N^2 e^{2c} - 1} \leqslant \frac{4}{3} \frac{e^{-2c}}{1 - e^{-2c}}.$$

Taking $c = \epsilon \ln(N)$ with $\epsilon > 0$ then yields the upper bound part of the cutoff phenomenon.

For the lower bound, we will show that the character χ_1 is a good witness of the distance between ψ_t and h. To do that, let us first estimate its mean and variance. We have, for $c < 0$ and $t = N\ln(N) + cN$,

$$\psi_t(\chi_1) = Ne^{-t/N} = e^{-c}$$

and the variance $\mathrm{var}_{\psi_t}(\chi_1) = \psi_t(\chi_1^2) - \psi_t(\chi_1)^2$ is given, using the fact that $\chi_1^2 = 1 + \chi_2$, by

$$\begin{aligned}
\mathrm{var}_{\psi_t}(\chi_1) &= 1 + (N^2 - 1)e^{-2tN/(N^2-1)} - N^2 e^{-2t/N} \\
&\leqslant 1 + N^2 e^{-2t/N} - N^2 e^{-2t/N} \\
&\leqslant 1.
\end{aligned}$$

Let us now view χ_1 as a continuous function on $[-N, N]$ and consider the Borel subset

$$B = \{s \in [-N, N] \mid |\chi_1(s)| \leqslant e^{-c}/2\}.$$

The indicator function $p = \mathbf{1}_B$ can be seen as a projection in the von Neumann algebra $L^\infty([-N, N])$ of essentially bounded functions. If we denote by ν_h (respectively μ_t) the unique Borel probability measure on $[-N, N]$ such that for any $x \in \mathcal{O}(O_N^+)_{\mathrm{central}}$,

$$h(x) = \int_{-N}^{N} x\mathrm{d}\nu_h \qquad \left(\text{respectively} \quad \psi_t(x) = \int_{-N}^{N} x\mathrm{d}\mu_t\right),$$

then these formulæ provide norm-preserving extensions of the states h and ψ_t to $L^\infty([-N,N])$. Moreover, because $\psi_t(\chi_1) = e^{-c}$, we have $B \subset \{s \in [-N,N] \mid |\chi_1(s) - \psi_t(\chi_1)| \geqslant e^{-c}/2\}$ so that by Chebyshev's inequality

$$\mu_t(B) \leqslant \left(e^{-c}/2\right)^{-2} \mathrm{var}_{\mu_t}(\chi_1) \leqslant 4e^{2c}.$$

Using again Chebyshev's inequality for ν_h with $h(\chi_1) = 0$ and $\mathrm{var}_h(\chi_1) = 1$, we eventually get

$$\begin{aligned}
|\mu_t(B) - \nu_h(B)| &\geqslant \nu_h(B) - \mu_t(B) \\
&= 1 - \nu_h([-N,N] \setminus B) - \mu_t(B) \\
&\geqslant 1 - 4e^{2c} - 4e^{2c} \\
&= 1 - 8e^{2c}.
\end{aligned}$$

To conclude, recall that because μ_t and ν_h are probability measures, the total variation norm of their difference coincides with twice their total variation distance, so that

$$\|\psi_t - h\| = \frac{1}{2}|\mu_t - \nu_h|([-N,N]) = \|\mu_t - \nu_h\|_{TV} \geqslant |\mu_t(B) - \nu_h(B)| \geqslant 1 - 8e^{2c}.$$

For $c = -\epsilon\ln(N)$ ($\epsilon > 0$ fixed), the right-hand side becomes $1 - 8N^{-2\epsilon}$ which tends to 1 as $N \to +\infty$, hence the proof is complete. □

Remark 4.3.2. In the papers [Fre19] and [Fre18], the lower bounds were only proven for c such that the corresponding state is absolutely continuous with respect to the Haar state. As a consequence, this does not yield a cutoff phenomenon in the sense of Definition 4.2.7 unless one makes sure that the states are asymptotically always absolutely continuous. As we will show in Proposition 4.3.6, and this was already observed in the aforementioned papers, this is never true. Hence, the term "cutoff" was slightly abusive there. However, using the same argument as in the above proof together with the estimates of [Fre19] and [Fre18], one can easily show that the random walks studied there indeed exhibit a *bona fide* cutoff phenomenon for the Fourier-Stieltjes norm.

Remark 4.3.3. In [Mé14], P.-L. Méliot proved that Brownian motion on $SO(N)$ exhibits a cutoff phenomenon at time $2\ln(N)$. The factor 2 comes from the fact that he chooses one half of the Laplace-Beltrami operator as an infinitesimal generator. As for the additional factor N in our result, it could be removed through setting $b_N = N$. A scaling-free statement on our case would therefore be that the cutoff time t_N satisfies $t_N b_N = N\ln(N)$ while in the case of P.-L. Méliot we have $t_N b_N = \ln(N)$.

Remark 4.3.4. Thanks to the work of F. Cipriani, U. Franz and A. Kula [CFK14], it is possible to construct a non-commutative Riemannian structure (a *spectral triple*) on O_N^+ out of a Lévy process. However, it was already noted in [CFK14, Sec 10] that in our case, and independently from the choice of b, the dimension of the resulting object is infinite, while one would expect a canonical Riemannian-like structure to be able to recover N through a notion of dimension.

We mentioned earlier that the use of the Fourier-Stieltjes norm was necessary because ψ_t cannot be extended to $L^\infty(O_N^+)$ in general. This can be thought of as an absolute continuity issue in the following sense. Let us denote by $L^1(O_N^+)$ the completion of $L^\infty(O_N^+)$ with respect to the norm $\|x\|_1 = h(|x|)$, where $|x|$ is obtained by functional calculus. A state ψ_t is then said to be *absolutely continuous* (with respect to the Haar state) if there exists $\rho_t \in L^1(O_N^+)$ such that $\psi_t(x) = h(\rho_t x)$ for all $x \in \mathcal{O}(O_N^+)$. It follows from the general theory (see [Tak02, Thm V.2.18]) that a state on $\mathcal{O}(O_N^+)$ is absolutely continuous with respect to the Haar state if and only if it extends to a normal linear map on $L^\infty(O_N^+)$.

We now want to give a precise result about absolute continuity, and this requires a finer estimate on λ_n than that of [FHL+17, Lem 1.7]. In fact, we will prove that λ_n is very close to being affine (and even linear).

Lemma 4.3.5. *Set $a_N = 1/\sqrt{N^2-4}$ and $b_N = (N-\sqrt{N^2-4})/(N^2-4)$. Then, for any $n \in \mathbf{N}$ and $N \geqslant 4$,*

$$\lambda_n - (a_N n + b_N) = c_N(n) := \frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{2}{N}\right)^{2j}\left(\frac{n+1}{2^{2j}}\sum_{\substack{\ell=j+1\\ \ell\equiv j \bmod n+1}}^{2j}\binom{2j}{\ell}\right),$$

and moreover,

$$0 \leqslant c_N(n) \leqslant \left(\frac{2}{N}\right)^n.$$

Proof. Let $n \in \mathbf{N}$ and $N \geq 4$. Recall that the roots of P_n are $x_k = 2\cos(k\pi/(n+1))$ for $1 \leqslant k \leqslant n$, so that

$$\lambda_n = \frac{P_n'(N)}{P_n(N)} = \sum_{k=1}^{n}\frac{1}{N-x_k} = \sum_{k=1}^{n}\frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{x_k}{N}\right)^j = \frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{2}{N}\right)^j\sum_{k=1}^{n}\cos\left(\frac{k\pi}{n+1}\right)^j.$$

We are therefore led to compute some sums of powers of trigonometric functions.

Let $j \in \mathbf{N}$ and set for convenience in the next computations $\theta = \frac{\pi}{n+1}$. Then we have,

$$\sum_{k=0}^{n}\cos(k\theta)^j = \sum_{k=0}^{n}\left(\frac{e^{ik\theta}+e^{-ik\theta}}{2}\right)^j = \sum_{k=0}^{n}\frac{1}{2^j}\sum_{\ell=0}^{j}\binom{j}{\ell}e^{ik\theta(2\ell-j)} = \frac{1}{2^j}\sum_{\ell=0}^{j}\binom{j}{\ell}\sum_{k=0}^{n}e^{ik\theta(2\ell-j)}.$$

Now observe that

$$\sum_{k=0}^{n}e^{ik\theta(2\ell-j)} = \begin{cases} 0 & \text{if } 2(n+1)\nmid 2\ell-j \\ n+1 & \text{if } 2(n+1)\mid 2\ell-j. \end{cases}$$

From this we see immediately that if j is odd, then the sum vanishes. If j is even, it is possible that $2(n+1) \mid 2\ell - j$ (i.e. that $n+1 \mid \ell - j/2$), and we have to consider all such ℓ's. We can hence rewrite the sum as

$$\sum_{k=0}^{n}\cos(k\theta)^j = \frac{n+1}{2^j}\sum_{\substack{\ell=0\\ \ell\equiv j/2 \bmod n+1}}^{j}\binom{j}{\ell}.$$

We deduce, splitting the previous sum according to whether $\ell = j/2$ or not, that

$$\lambda_n + \frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{2}{N}\right)^{2j} = \frac{n+1}{N}\sum_{j=0}^{+\infty}\left(\frac{1}{N}\right)^{2j}\binom{2j}{j} + \frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{2}{N}\right)^{2j}\left(\frac{n+1}{2^{2j}}\sum_{\substack{\ell=j+1\\ \ell\equiv j \bmod n+1}}^{2j}\binom{2j}{\ell}\right),$$

i.e., observing that $(1-4x)^{-1/2} = \sum_{j=0}^{\infty} \binom{2j}{j} x^j$ for $|x| < 1/4$,

$$\lambda_n + \frac{1}{N}\frac{1}{1-4/N^2} = \frac{n+1}{N}\frac{1}{\sqrt{1-4/N^2}} + c_N(n),$$

which rewrites exactly as

$$\lambda_n = a_N n + b_N + c_N(n).$$

Let us now bound $c_N(n)$. Because the sum $\sum_{\ell=j+1, \ell\equiv j \mod n+1}^{2j}$ is empty if $j < n$, we get

$$c_N(n) = \frac{1}{N}\sum_{j=0}^{+\infty}\left(\frac{2}{N}\right)^{2j}\left(\frac{n+1}{2^{2j}}\sum_{\substack{\ell=j+1\\ \ell\equiv j \mod n+1}}^{2j}\binom{2j}{\ell}\right) \leqslant \frac{1}{N}\sum_{j=n}^{+\infty}\left(\frac{2}{N}\right)^{2j}\left(\frac{n+1}{2^{2j}}2^{2j}\right) = \frac{n+1}{N-2}\left(\frac{2}{N}\right)^{2n}.$$

As $N \geq 4$, we finally obtain

$$c_N(n) \leqslant \frac{n+1}{N-2}\left(\frac{2}{N}\right)^{2n} \leqslant \frac{1}{N-2}\frac{n+1}{2^n}\left(\frac{2}{N}\right)^{n} \leqslant \left(\frac{2}{N}\right)^{n}.$$

□

We can now give a precise criterion for absolute continuity.

Proposition 4.3.6. Let $N \geqslant 4$. Then, there exists a positive time $t_{abscont}(N)$ such that

- If $t < t_{abscont}(N)$, then ψ_t is not absolutely continuous with respect to the Haar state,
- If $t > t_{abscont}(N)$, then ψ_t is absolutely continuous with respect to the Haar state. Moreover,

$$t_{abscont}(N) = \frac{-\ln(q(N))}{a_N} = N\ln(N) - \frac{2\ln(N)}{N} + \underset{N\to+\infty}{O}\left(\frac{1}{N}\right).$$

Proof. Let $N \geqslant 4$, and set

$$\rho_t = \sum_{n=0}^{+\infty} d_n e^{-t\lambda_n}\chi_n.$$

If this series converges in $L^1(O_N^+)$, then ψ_t is absolutely continuous with respect to the Haar state with density ρ_t. Moreover, if it converges in $L^2(O_N^+)$, then it converges also in $L^1(O_N^+)$ (as the L^1-norm is dominated by the L^2-norm), and

$$\|\rho_t\|_2^2 = \sum_{n=0}^{+\infty} d_n^2 e^{-2t\lambda_n}.$$

Recall that $d_n \leqslant q(N)^{-n}/(1-q(N)^2)$. Using Lemma 4.3.5,

$$\begin{aligned}
\ln\left(d_n^2 e^{-2t\lambda_n}\right) &\leqslant 2n\ln(1/q(N)) - \ln(1-q(N)^2) - 2t\lambda_n \\
&\leqslant 2n\left(\ln(1/q(N)) - ta_N\right) + \ln(1/q(N)) - \ln(1-q(N)^2) - (b_N + c_N)t \\
&\leqslant 2n\left(\ln(1/q(N)) - ta_N\right) + g(t,N)
\end{aligned}$$

where $g(t, N)$ is a quantity independent of n. Consequently, $\|\rho_t\|_2$ is finite (so that ψ_t is absolutely continuous with respect to the Haar state) as soon as

$$\ln(1/q(N)) - ta_N < 0.$$

As for the second point, observe that

$$\frac{\psi_t(\chi_n)}{\|\chi_n\|_\infty} = \frac{d_n e^{-t\lambda_n}}{n+1}.$$

If the right-hand side is not uniformly bounded with respect to n, then ψ_t cannot extend to $L^\infty(O_N^+)$. Using the previous lemma and the fact (see for instance [Fre19, Lem 3.8]) that $d_n \geqslant Nq(N)^{-(n-1)}$, and proceeding as above, we see that $\frac{d_n e^{-t\lambda_n}}{n+1}$ will not be bounded as soon as

$$\ln(1/q(N)) - ta_N > 0.$$

It follows from our two inequalities that $t_{abscont} = -\ln(q(N))a_N^{-1}$. Using the Taylor expansion of $q(N)^{-1}$ computed in the proof of 4.2.10, we see that $\ln(1/q(N)) = \ln(N) + O\left(\frac{1}{N^2}\right)$ so that

$$\begin{aligned}
t_{abscont}(N) &= \left(\ln(N) + O\left(\frac{1}{N^2}\right)\right)\sqrt{N^2-4} \\
&= \left(N\ln(N) + O\left(\frac{1}{N}\right)\right)\left(1 - \frac{2}{N^2} + O\left(\frac{1}{N^4}\right)\right) \\
&= N\ln(N) - \frac{2\ln(N)}{N} + O\left(\frac{1}{N}\right).
\end{aligned}$$

This concludes the proof. $\square$

Remark 4.3.7. This is in sharp contrast with the classical case, where any non-degenerate (a condition analogous to requiring $b > 0$) Lévy process automatically has an L^2-density with respect to the Haar measure by [Lia04, Thm 1], and is thus absolutely continuous. Indeed, the existence of bounded approximate identities by convolutions of L_1-densities is equivalent to the co-amenability of the quantum group, which can not hold for O_N^+ (see for instance [Bra17]).

Remark 4.3.8. As we will see in Proposition 4.3.14, ψ_t is absolutely continuous with respect to the Haar state if and only if the measure of the corresponding classical process is absolutely continuous with respect to the semi-circle distribution. Moreover, the lack of absolute continuity is witnessed by the appearance of an atom in that measure.

In particular, $\psi_{N\ln(N)+cN}$ is absolutely continuous if and only if $c \geqslant -2\ln(N)/N^2 + O\left(1/N^2\right)$, which goes to 0 as N goes to infinity. Thus, the total variation distance is asymptotically only defined for $c \geqslant 0$.

4.3.2 Cutoff profile

We will now try to get a better understanding of the cutoff phenomenon by computing the corresponding *cutoff profile*, that is to say the limit of the distance between the process at time $t_c = N\ln(N) + cN$ and the Haar state as N goes to infinity, c being fixed. Our main result is an expression of this limit as the distance between two explicit probability measures. Before stating it, let us give some heuristics.

In the proof of Theorem 4.3.1, we saw that it was enough to consider the element χ_1 to obtain a lower bound of the correct order for the mixing time. In the case of a classical compact matrix group, χ_1 is nothing but the trace function, and this would mean that the trace of the matrices is the last thing to be mixed by Brownian motion. In the case of O_N^+, we know that the distribution of χ_1 under the Haar state is the semi-circle distribution ν_{SC}, so that we may expect the cutoff profile to be given by the distance between ν_{SC} and a "deformation" of it. The whole problem of course lies in the vague meaning of the word "deformation".

We will show in the first part of Theorem 4.3.9 that the profile indeed appears as the distance between ν_{SC} and a family of closely related laws called the *free Poisson distributions*. Let us recall that the free Poisson distribution with rate λ and jump size α is given, for $\lambda > 1$, by (see [NS06, Def 12.12] for details)

$$\mathrm{d}\,\mathrm{Poiss}^+(\lambda, \alpha)(t) = \frac{1}{2\pi\alpha t}\sqrt{4\lambda\alpha^2 - (t - \alpha(1+\lambda))^2}\mathrm{d}t.$$

Unfortunately, there is no value of the parameters for which the free Poisson distribution equals the semi-circle one.

One can nevertheless write things differently using a larger family of probability distribution called the *free Meixner distributions*. Let us denote by $\mathrm{Meix}^+(a, b)$ the standardised (i.e. with mean 0 and variance 1) free Meixner law with parameters a and b (see for instance [BB06, Sec 2.2] for details). Its absolutely continuous part with respect to the Lebesgue measure is given by

$$\mathrm{d}\,\mathrm{Meix}^+(a, b)(t) = \frac{\sqrt{4(1+b) - (t-a)^2}}{2\pi(bt^2 + at + 1)}\mathbf{1}_{[a-2\sqrt{1+b}, a+2\sqrt{1+b}]}\mathrm{d}t.$$

For $a = b = 0$, the formula reduces to the density of the semi-circular distribution, while for $b = 0$ it yields the density of a free Poisson distribution with mean 0 and variance 1.

We will now state our result using both the free Poisson and the free Meixner settings, after introducing some extra notations. If X is a random variable with law μ, then we denote by $D_r(\mu)$ the r-dilation of μ (that is to say the law of rX) and by $\mu * \delta_a$ its translation by a (that is to say the law of $X + a$). Moreover, we denote by d_{TV} the usual total variation distance for Borel measure on $\mathbf{R}$ and by $\widetilde{\mathrm{d}}_N$ the distance associated to the norm $\|\cdot\|$ on O_N^+.

Theorem 4.3.9. Let $c \in \mathbf{R}$, and recall $t_c = N\ln(N) + cN$. Then

$$\begin{aligned}\widetilde{\mathrm{d}}_N\left(\psi_{t_c}, h\right) \xrightarrow[N\to\infty]{} f(c) &:= \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c + e^{-c}}, \nu_{\mathrm{SC}}\right) \\ &= \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Meix}^+\left(-e^{-c}, 0\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,0)\right).\end{aligned}$$

As the proof of this result is long, we will split it in two, depending on the sign of c. For convenience and clarity, the two cases will be stated in the separate Propositions 4.3.10 and 4.3.14, each treated in a proper subsection.

4.3.2.1 The profile on the right

We start with the case $c > 0$, which turns out to be the simplest one. The reason for this is that it follows from Proposition 4.3.6 that ψ_{t_c} is absolutely continuous with respect to the Haar state in that case, so that the convergence in total variation distance boils down to L^1-convergence of the densities, which in turn follows from L^2-convergence. The main part of the work is therefore rather the identification of the limit and its expression in terms of Poisson or free Meixner laws.

Proposition 4.3.10. For $c > 0$,

$$\begin{aligned}\widetilde{\mathrm{d}}_N\left(\psi_{t_c}, h\right) \xrightarrow[N\to\infty]{} f(c) &= \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c + e^{-c}}, \nu_{\mathrm{SC}}\right) \\ &= \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Meix}^+\left(-e^{-c}, 0\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,0)\right).\end{aligned}$$

Proof. Recall that as $c > 0$, ψ_{t_c} has an L^1-density given by

$$\rho_{t_c} = \sum_{n=0}^{+\infty} d_n e^{-t_c\lambda_n}\chi_n.$$

Moreover, we know from Lemma 4.3.5 that

$$\lambda_n = n\left(\frac{1}{N} + \underset{N\to\infty}{O}\left(\frac{1}{N^3}\right)\right),$$

and an easy computation yields $d_n \underset{N\to\infty}{\sim} N^n$. In particular, for each n, $d_n e^{-t_c\lambda_n}$ converges to e^{-cn} as N goes to $+\infty$. Moreover,

$$\|d_n e^{-t_c\lambda_n}\chi_n\|_1 \leqslant \|d_n e^{-t_c\lambda_n}\chi_n\|_2 = d_n e^{-t_c\lambda_n}$$

and because $qN \geqslant 1$, for $N \geqslant 3$,

$$d_n e^{-t_c\lambda_n} \leqslant \frac{q^{-n}}{(1-q^2)} N^{-n} e^{-nc} \leqslant \frac{3}{2} e^{-nc}.$$

The latter being summable and independent of N, we can exchange the sum over n and the limit in N. This yields (recall that there is an isomorphism between $\mathcal{O}(O_N^+)_{\mathrm{central}}$ and $\mathbf{C}[X]$ sending χ_n to P_n and sending the measure associated to h to the semi-circle distribution)

$$\lim_{N\to+\infty} \|\psi_{t_c} - h\| = \frac{1}{2}\left\|\sum_{n=1}^{+\infty} e^{-cn}P_n\right\|_1,$$

where the L^1-norm is computed with respect to the standard semi-circular distribution. Using the generating series of the Chebyshev polynomials of the second kind (which is easily computed, multiplying by t and using the recursion relation), we get for every $t \in [-2, 2]$,

$$\begin{aligned}
\sum_{n=1}^{+\infty} e^{-cn}P_n(t) &= \frac{1}{1 - te^{-c} + e^{-2c}} - 1 \\
&= \frac{1}{1+\beta^2}\frac{1}{1-\gamma t} - 1 \\
&= F_c(t) - 1.
\end{aligned}$$

where $\beta = e^{-c}$ and $\gamma = \beta/(1+\beta^2) < 1/2$. Thus, the cutoff profile is equal to

$$\lim_{N\to+\infty} \|\psi_{t_c} - h\| = \frac{1}{2}\int_{-2}^{2} |F_c(t) - 1|\, \mathrm{d}\nu_{\mathrm{SC}}(t).$$

Performing the change of variables $u = 1 - \gamma t$,

$$\begin{aligned}
F_c(t)\mathrm{d}\nu_{\mathrm{SC}}(t) = F_c(t)\frac{\sqrt{4-t^2}}{2\pi}\mathbf{1}_{[-2,2]}(t)\mathrm{d}t &= \frac{1}{2\pi(1+\beta^2)u}\sqrt{4 - \left(\frac{1-u}{\gamma}\right)^2}\mathbf{1}_{[1-2\gamma,1+2\gamma]}(u)\frac{\mathrm{d}u}{\gamma} \\
&= \frac{1}{2\pi\gamma^2(1+\beta^2)u}\sqrt{4\gamma^2 - (1-u)^2}\mathbf{1}_{[1-2\gamma,1+2\gamma]}(u)\mathrm{d}u.
\end{aligned}$$

Setting $\alpha = \beta\gamma = \gamma^2(1+\beta^2)$ and $\lambda = \beta^{-2} > 1$, this density becomes

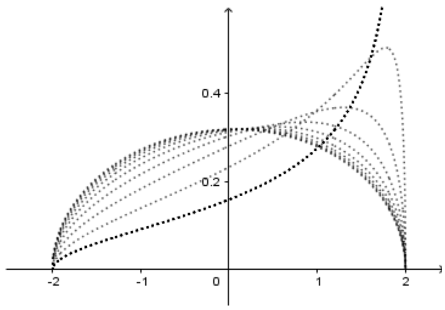
$$\frac{1}{2\pi\alpha u}\sqrt{4\lambda\alpha^2 - (u-\alpha(1+\lambda))^2}\mathbf{1}_{[\alpha(1-\sqrt{\lambda})^2,\alpha(1+\sqrt{\lambda})^2]}(u)\mathrm{d}u.$$

This is exactly the free Poisson distribution with rate $\lambda = e^{2c}$ and jump size $\alpha = e^{-2c}/(1+e^{-2c})$. Reversing the change of variables, we see that $F_c(t)\mathrm{d}\nu_{\mathrm{SC}}(t)$ is the density of the law

$$D_{-1/\gamma}\left(\mathrm{Poiss}^+\left(\beta^{-2}, -\beta\gamma\right) * \delta_{-1}\right) = \mathrm{Poiss}^+\left(\beta^{-2}, -\beta\right) * \delta_{1/\gamma} = \mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}},$$

hence the result. Using the facts that $\mathrm{Poiss}^+\left(a^{-2}, a\right) * \delta_{-a^{-1}} = \mathrm{Meix}^+(a,0)$ and that $\nu_{\mathrm{SC}} = \mathrm{Meix}^+(0,0)$, the second formula follows. $\square$

As explained heuristically at the beginning of this subsection, the fact that Brownian motion is not completely mixed is witnessed by the "trace" it can attain, and the cutoff profile gives a precise quantitative description of this phenomenon. In particular, it shows that the "trace" of Brownian motion is averagely shifted to the right and more concentrated around its mean. Here is a plot of the density of $\mathrm{Meix}^+(-e^{-c},0) * \delta_{e^{-c}}$ with respect to the Lebesgue measure for values of c between 0 and 5 :



For $c = 0$ we get a free Poisson law (this is the curve with a peak on the right) while for $c = 5$ the density is already indistinguishable to the naked eye from that of the semi-circle distribution.

Remark 4.3.11. The quantity $\mathrm{d_{TV}}\left(\mathrm{Meix}^+\left(-e^{-c},0\right) * \delta_{e^{-c}}, \mathrm{Meix}^+(0,0)\right)$ can be computed explicitly in terms of c through integration, yielding the formula :

$$f(c) = \frac{e^{2c}-1}{2\pi}\arcsin\left(\frac{e^{-3c}-3e^{-c}}{2}\right) + \frac{1-e^{2c}}{2\pi}\arcsin\left(\frac{e^{-c}}{2}\right) + \frac{e^{-2c}+2}{4\pi}\sqrt{4e^{2c}-1}.$$

Using the facts that $0 < \frac{e^{-c}}{2} < 1/2$ and that $3\arcsin(x) = \arcsin(3x-4x^3)$ for $0 \leqslant x \leqslant 1/2$, this can be rewritten as

$$f(c) = (1-e^{2c})\frac{2}{\pi}\arcsin\left(\frac{e^{-c}}{2}\right) + \frac{e^{-2c}+2}{4\pi}\sqrt{4e^{2c}-1}.$$

4.3.2.2 The profile on the left

By Proposition 4.3.6, the computations above can only make sense for $c \geqslant 0$. However, the free Poisson distribution makes sense even for $\lambda < 1$, with the only difference that some mass is carried by an atom :

$$\mathrm{d}\,\mathrm{Poiss}^{+}(\lambda, \alpha)(t) = (1-\lambda)\,\delta_0 + \frac{\lambda}{2\pi\alpha t}\sqrt{4\lambda\alpha^2 - (t-\alpha(1+\lambda))^2}\,\mathbf{1}_{[\alpha(1-\sqrt{\lambda})^2, \alpha(1+\sqrt{\lambda})^2]}(t)\mathrm{d}t$$

As a consequence, the formula

$$f(c) = \mathrm{d}_{\mathrm{TV}}\left(\mathrm{Poiss}^{+}\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}, \nu_{\mathrm{SC}}\right)$$

still makes sense for $c < 0$ and we will now show that this is indeed the profile. Let us start with a characterization of the limit distribution in terms of "Chebyshev moments".

Lemma 4.3.12. *For any $c \in \mathbf{R}$, the measure*

$$\mu_c = \mathrm{Poiss}^{+}\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}$$

is the unique probability measure on $\mathbf{R}$ such that for any $n \in \mathbf{N}$,

$$\int_{\mathbf{R}} P_n \mathrm{d}\mu_c = e^{-cn}.$$

Proof. Let us first recall that the free cumulants (see [NS06, Def 11.3] for the definition of free cumulants and [NS06, Prop 12.11] for the free Poisson case) of the free Poisson distribution $\mathrm{Poiss}^{+}(\lambda, \alpha)$ are given by

$$\kappa_n = \lambda\alpha^n.$$

As a consequence, the free cumulants of $\mathrm{Poiss}^{+}\left(e^{2c}, -e^{-c}\right)$ are Laurent polynomials in e^c. The free additive convolution with $\delta_{e^c+e^{-c}}$ only modifies the first cumulant κ_1 by adding $e^c + e^{-c}$ to it, hence the free cumulants of μ_c are Laurent polynomials in e^c. Because the moments are polynomial functions of the free cumulants (by virtue of the moment-cumulant formula, see [NS06, Prop 11.4]), we conclude that there exist Laurent polynomials L_n such that for any $c \in \mathbf{R}$,

$$\int_{\mathbf{R}} P_n(x)\mathrm{d}\mu_c(x) = L_n(e^c).$$

Let us now assume that $c > 0$. Then, we know from the proof of Theorem 4.3.9 in Subsection 4.3.2.1 that for $c > 0$, μ_c is absolutely continuous with respect to the semi-circle distribution ν_{SC} with density F_c. Using the fact that the polynomials P_n are orthonormal for the semi-circle distribution, we get

$$\int_{\mathbf{R}} P_n \mathrm{d}\mu_c = \int_{\mathbf{R}} P_n \left(\sum_{n=0}^{+\infty} e^{-cn} P_n\right) \mathrm{d}\nu_{\mathrm{SC}} = e^{-cn}.$$

As a consequence, $L_n(x) = x^{-n}$ for any $x > 1$, so that by uniqueness of the decomposition of a Laurent polynomial, $L_n(X) = X^{-n}$ and the formula for the integral of Chebyshev polynomials also holds for $c \leqslant 0$.

As for the uniqueness assertion, it simply follows from the fact that μ_c has compact support and is consequently determined by its moments. ∎

With this in hand, we can complete our proof of the cutoff profile. To do so, let us define, for $t > 0$ and $N \in \mathbf{N}$, a probability measure $m_t^{(N)}$ on $\mathbf{R}$ by requiring that $m_t^{(N)}(\mathbf{R} \setminus [-N, N]) = 0$ and for all $n \in \mathbf{N}$,

$$\int_{-N}^{N} P_n(x) \mathrm{d}m_t^{(N)} = e^{-t\lambda_n} P_n(N).$$

Such a measure exists because it corresponds to the restriction of the state ψ_t to the algebra $\mathcal{O}(O_N^+)_{\text{central}}$. Using the centrality of the process and the Riesz representation theorem (see the discussion at the beginning of Section 4.3.1), we know that

$$\|\psi_{t_c} - h\| = \left\| m_{t_c}^{(N)} - \nu_{\text{SC}} \right\|_{TV}.$$

To find the limit of the right-hand side, we will compute $m_{t_c}^{(N)}$ explicitly. Our approach is based on an idea of P. Biane in [Bia08, Sec 12.2] for the study of a similar process on the duals of free groups. One would naively want the measure to be absolutely continuous with respect to the Haar measure with density

$$\sum_{n=0}^{+\infty} e^{-t\lambda_n} P_n(N) P_n$$

However, we know that for $t = N\ln(N) + cN$ with $c < 0$, this series does not converge in $L^2([-2, 2])$ because $e^{-t\lambda_n} P_n(N)$ grows exponentially (this is the lack of absolute continuity with respect to the Haar state). The idea is to find a point $\widetilde{N}(t)$ such that $e^{-t\lambda_n} P_n(N) - P_n(\widetilde{N}(t))$ is small enough to be summable. Then summing the previous difference to produce an absolutely continuous measure and adding a Dirac mass at $\widetilde{N}(t)$ will produce an expression of the measure. To find that $\widetilde{N}(t)$, let us first recall that for $x > 2$, if $0 < q(x) < 1$ denotes the unique number such that $q(x) + q(x)^{-1} = x$, then

$$P_n(x) = \frac{q(x)^{-(n+1)} - q(x)^{n+1}}{q(x)^{-1} - q(x)}.$$

In particular, considering (with the notations of Lemma 4.3.5) that λ_n is roughly equal to $a_N n$, we have

$$\begin{aligned} e^{-t\lambda_n} P_n(N) &\approx e^{-nta_N} \frac{q(N)^{-(n+1)} - q(N)^{n+1}}{q(N)^{-1} - q(N)} \\ &\approx e^{-nta_N} q(N)^{-n} \\ &\approx P_n\left(e^{-a_N t} q(N)^{-1} + e^{a_N t} q(N)\right) \end{aligned}$$

so that setting

$$\widetilde{N}(t) = e^{-a_N t} q(N)^{-1} + e^{a_N t} q(N)$$

has a good chance of giving the correct order of magnitude. There is still a normalization coefficient required, and the next result is meant to make all this precise and rigorous. We will write q for $q(N)$ in order to lighten notations.

Lemma 4.3.13. *For any fixed $c < 0$ and $t_c = N\ln(N) + cN$, there exists N_c such that for all $N \geqslant N_c$,*

$$m_{t_c}^{(N)} = \alpha(t_c)\delta_{\widetilde{N}(t_c)} + \sum_{n=0}^{+\infty} \left(e^{-\lambda_n t_c} P_n(N) - P_n(\widetilde{N}(t_c))yyy\right) P_n(y) \mathrm{d}\mu_{SC}(y),$$

where
$$\alpha(t) = e^{a_N t - b_N t} \frac{e^{-a_N t} q^{-1} - e^{a_N t} q}{q^{-1} - q}$$
and
$$\sum_{n \geqslant 0} \left(e^{-t_c \lambda_n} P_n(N) - \alpha(t_c) P_n(\widetilde{N}(t_c)) \right) P_n \in L^2([-2,2], \nu_{SC}).$$

Proof. To lighten notations, let us simply write a, b, c for a_N, b_N, c_N. Since by definition $q\left(\widetilde{N}(t)\right) = (e^{at}q)^{\pm 1}$ (depending on which one is less than one, but this does not change the expression of $P_n(\widetilde{N}(t))$),

$$\begin{aligned}
\left(q^{-1} - q\right) \alpha(t) P_n\left(\widetilde{N}(t)\right) &= e^{at-bt}\left(e^{-at}q^{-1} - e^{at}q\right) \frac{q\left(\widetilde{N}(t)\right)^{-n-1} - q\left(\widetilde{N}(t)\right)^{n+1}}{q\left(\widetilde{N}(t)\right)^{-1} - q\left(\widetilde{N}(t)\right)} \\
&= e^{at-bt}\left(e^{-a(n+1)t}q^{-(n+1)} - e^{a(n+1)t}q^{n+1}\right) \\
&= e^{-(an+b)t}q^{-(n+1)} - e^{a(n+2)t-bt}q^{n+1},
\end{aligned}$$

from which we deduce that

$$\alpha(t) P_n\left(\widetilde{N}(t)\right) = e^{-(an+b)t} P_n(N) + \frac{q^{n+1}}{q^{-1} - q}\left(e^{-(an+b)t} - e^{(a(n+2)-b)t}\right),$$

which eventually leads to the equality

$$\alpha(t) P_n\left(\widetilde{N}(t)\right) - e^{-\lambda_n t} P_n(N) = \left(e^{c(n)t} - 1\right) e^{-\lambda_n t} P_n(N) + \frac{q^{n+1}}{q^{-1} - q}\left(e^{-(an+b)t} - e^{(a(n+2)-b)t}\right).$$

Using that $e^y - 1 \leqslant (e-1)y \leqslant 2y$ for $0 \leqslant y \leqslant 1$ and using Lemma 4.3.5, we see that for $N \geqslant 4$,

$$0 \leqslant e^{c(n)t} - 1 \leqslant 2\left(\frac{2}{N}\right)^n.$$

Moreover, $a \geqslant 1/N$ and $b, c(n) \geqslant 0$ so that $e^{-\lambda_n t} \leqslant e^{-ant} \leqslant e^{-nt/N}$ (remember that $c < 0$). Taking $t = t_c = N \ln(N) + cN$ (and of course N large enough so that $t_c > 0$), and using also $|P_n(N)| \leqslant N^n$ we can bound the first term :

$$\left| \left(e^{c(n)t_c} - 1\right) e^{-\lambda_n t_c} P_n(N) \right| \leqslant 2\left(\frac{2}{N}\right)^n e^{-nt_c/N} N^n = 2\left(\frac{2e^{-c}}{N}\right)^n.$$

Taking N greater than some N_c depending on c, this converges exponentially fast to 0.

Let us now bound the second term. First observe that $q^{-1} - q = \sqrt{N^2 - 4}$ and $e^{-(an+b)t} \leqslant 1 \leqslant e^{((n+2)-b)t}$, hence we have

$$\left| \frac{q^{n+1}}{q^{-1} - q}\left(e^{-(an+b)t} - e^{(a(n+2)-b)t}\right) \right| \leqslant \frac{q^{n+1}}{\sqrt{N^2 - 4}} e^{(a(n+2)-b)t} \leqslant \frac{q^{n+1}}{\sqrt{N^2 - 4}} e^{a(n+2)t}.$$

At $t = t_c$, we then get as $N \to \infty$, using $a = \frac{1}{N} + O\left(\frac{1}{N^3}\right)$, that

$$\begin{aligned}\frac{q^{n+1}}{\sqrt{N^2-4}} e^{a(n+2)t_c} &= \left(\frac{1}{N} + O\left(\frac{1}{N^3}\right)\right)^{n+2} \left(e^{a(N\ln(N)+cN)}\right)^{n+2} \\ &= \left(e^{-\ln(N)+O(1/N^2)}\right)^{n+2} \left(e^{\ln(N)+c+O(\ln(N)/N^2)}\right)^{n+2} \\ &= \left(e^{c+o(1)}\right)^{n+2}.\end{aligned}$$

Hence, as $c < 0$ and $N \geqslant N_c$ is fixed, this term converges exponentially fast to 0 as $n \to \infty$. Combining those two bounds, we conclude that for N fixed large enough,

$$\left|\alpha(t)P_n\left(\widetilde{N}(t)\right) - e^{-\lambda_n t}P_n(N)\right|$$

converges exponentially fast to 0 as $n \to \infty$. Because $(P_n)_{n\in\mathbf{N}}$ is a Hilbert basis for $L^2([-2,2],\nu_{\mathrm{SC}})$, this proves that the sum stated in the lemma belongs to $L^2([-2,2],\nu_{\mathrm{SC}})$.

Now the first formula in the statement makes sense for N large enough and defines a probability measure on $[-N,N]$. Moreover, the integral of any Chebyshev polynomial with respect to that measure coincides by construction with its integral with respect to $m_{t_c}^{(N)}$. Because Chebyshev polynomials form a basis of $\mathbf{C}[X]$, it follows that the two measures have the same moments and hence coincide since they have compact support. $\square$

We are now ready to establish the cutoff profile for negative c.

Proposition 4.3.14. For any fixed $c < 0$ and $t_c = N\ln(N) + cN$, we have

$$\widetilde{\mathrm{d}}_N(\psi_{t_c}, h) = \left\|m_{t_c}^{(N)} - \nu_{\mathrm{SC}}\right\|_{TV} \underset{N\to+\infty}{\longrightarrow} \left\|\mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}, \nu_{\mathrm{SC}}\right\|_{TV}.$$

Proof. To compute the limit of the total variation distance, let us first notice that if $\widetilde{m}_{t_c}^{(N)} = \mathbf{1}_{[-2,2]} m_{t_c}^{(N)}$, then

$$\begin{aligned}\left\|m_{t_c}^{(N)} - \nu_{\mathrm{SC}}\right\|_{TV} &= \left\|\mathbf{1}_{[-2,2]}\left(m_{t_c}^{(N)} - \nu_{\mathrm{SC}}\right)\right\|_{TV} + \left\|\mathbf{1}_{\mathbf{R}\setminus[-2,2]}\left(m_{t_c}^{(N)} - \nu_{\mathrm{SC}}\right)\right\|_{TV} \\ &= \left\|\widetilde{m}_{t_c}^{(N)} - \nu_{\mathrm{SC}}\right\|_{TV} + \alpha(t_c).\end{aligned}$$

We will deal with each part separately :

- It is straightforward to see that as N goes to infinity, $\alpha(t_c) \to 1 - e^{2c}$ and that $\widetilde{N}(t_c) \to e^c + e^{-c}$, which are exactly the parameters of the atom of the free Poisson distribution in the statement.
- Observe moreover that $m_{t_c}^{(N)}$ converges in moments to a measure for which the integral of P_n equals e^{cn}. By Lemma 4.3.12, this must be the free Poisson distribution in the statement. Therefore, if we can prove that $\widetilde{m}_{t_c}^{(N)}$ converges in total variation distance, then its limit must be the absolutely continuous part of $\mathrm{Poiss}^+\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}$. The density of $\widetilde{m}_{t_c}^{(N)}$ with respect to ν_{SC} is

$$\sum_{n=0}^{+\infty}\left(e^{-\lambda_n t_c}P_n(N) - P_n(\widetilde{N}(t))\right)P_n$$

and each term of the sum converges as N goes to infinity. Therefore, if we can bound these terms in L^2-norm by a summable sequence not depending on N, then we can conclude. But this was already done in Lemma 4.3.13, since it follows from it that such a bound exists at least for N large enough depending on c.

□

Remark 4.3.15. Note that we do not prove that $m_{t_c}^{(N)}$ converges in total variation to $\mathrm{Poiss}^{+}\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}$. This is actually false since their atoms are not at the same place. If we change slightly the time by setting

$$\widehat{t_c} = \frac{-\ln(q(N)) + c}{a_N} = t_{abscont}(N) + \frac{c}{a_N},$$

then $\tilde{N}(\widehat{t_c}) = e^c + e^{-c}$ and $m_{\widehat{t_c}}^{(N)}$ converges in total variation distance to $\mathrm{Poiss}^{+}\left(e^{2c}, -e^{-c}\right) * \delta_{e^c+e^{-c}}$. However, this would be somewhat unnatural, and we prefer to state only results which are true for all times $t_c + o(N)$ than for a very specific sequence of times.

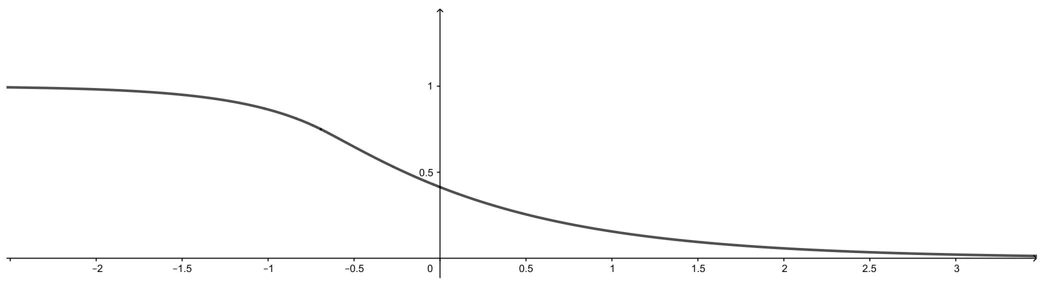
Remark 4.3.16. Once again, the integral giving the total variation distance can be computed explicitly in terms of c. The computation differs however depending on whether c is greater or smaller than $-\ln(2)$ (because of the absolute value in the integral). For $-\ln(2) < c < 0$, the result is the same as the first expression given in Remark 4.3.11 for $c > 0$, except for an extra $(1 - e^{2c})/2$. However, this time $1/2 \leqslant e^{-c}/2 \leqslant 1$ and for $x \in [1/2, 1]$, $\pi - 3\arcsin(x) = \arcsin(3x - 4x^3)$ so that we end up with the same result, namely

$$f(c) = (1 - e^{2c})\frac{2}{\pi}\arcsin\left(\frac{e^{-c}}{2}\right) + \frac{e^{-2c} + 2}{4\pi e^{-c}}\sqrt{4 - e^{-2c}}$$

while for $c \leqslant -\ln(2)$,

$$f(c) = 1 - e^{2c}.$$

The combination of Propositions 4.3.10 and 4.3.14 yields the complete proof of Theorem 4.3.9, therefore yielding the cutoff profile for Brownian motion on O_N^+. As an illustration, here is a plot of the profile :



Remark 4.3.17. Theorem 4.3.9 is not just a refinement of the cutoff phenomenon established in Theorem 4.3.1, but a stronger result. In other words, the existence of a cutoff profile implies the cutoff phenomenon, as was already mentioned in the beginning of the proof of Theorem 4.3.1. This follows from the fact that the profile converges to 1 at $-\infty$ and 0 at $+\infty$, together with the monotonicity of Proposition 4.2.6.

4.3.3 Further results

Let us complete this section with some additional remarks and results concerning various generalizations of the original problem.

4.3.3.1 Other norms

We have worked so far with the Fourier-Stieltjes norm, because it is the only natural norm available which makes sense for all $t \in \mathbf{R}_+$. However, the upper bound was computed using the total variation distance, and one may wonder whether the cutoff upper bound also holds with respect to other distances. It turns out that the answer is yes.

Corollary 4.3.18. *Brownian motion on O_N^+ satisfies, for all $1 \leqslant p \leqslant \infty$ and $c > 0$, with $t_c = N\ln(N) + cN$,*

$$\lim_{N\to+\infty} \|\psi_{t_c} - h\|_{L^p} = \left\|\mathrm{Meix}^+\left(-e^{-c}, 0\right) * \delta_{e^{-c}} - \mathrm{Meix}^+(0,0)\right\|_{L^p}.$$

Proof. Recall that the density of the process at time t is, if the series makes sense,

$$\rho_t = \sum_{n=0}^{+\infty} d_n e^{-t\lambda_n}\chi_n.$$

Using the fact that $\|\chi_n\|_\infty = P_n(2) = n+1$, we see that the density converges in L^∞-norm at t_c as soon as $c > 0$ since (for $N \geqslant 3$) and

$$\left\|d_n e^{-t_c\lambda_n}\chi_n\right\|_\infty \leqslant (n+1)d_n e^{-t_c\lambda_n} \leqslant \frac{3}{2}(n+1)e^{-nc}.$$

Moreover, using this bound the same strategy as for Proposition 4.3.10 yields the cutoff profile in the case $p = \infty$. As for finite p, it follows from the noncommutative Hölder inequality (see for instance [Tak03, Thm 2.13.iv]) that for any $1 \leqslant p \leqslant \infty$,

$$\|d_n e^{-t_c\lambda_n}\chi_n\|_p \leqslant \|d_n e^{-t_c\lambda_n}\chi_n\|_\infty$$

hence we can once again resort to the same argument. □

Remark 4.3.19. Note that for $c < 0$, according to the discussions in the previous subsection, we may decompose ψ_{t_c} into an absolutely continuous part $\widetilde{\psi}_{t_c}$ which admits a L^p-density for all $1 \leqslant p \leqslant \infty$ and such that $\psi_{t_c} - \widetilde{\psi}_{t_c}$ is singular with respect to h. It is easy to see that the above theorem still holds for $\widetilde{\psi}_{t_c}$ even for $c < 0$.

Let us compare this with the classical case. P.-L. Méliot proved in [Mé14, Thm 7], building on results of G. Chen and L. Saloff-Coste in [CSC08], that the cutoff phenomenon for Brownian motion on $SO(N)$ indeed occurs for all $1 \leqslant p \leqslant \infty$ and that the cutoff time is the same as for the L^1-norm for all $1 \leqslant p < \infty$. However, for $p = \infty$, the cutoff time is doubled and becomes $4\ln(N)$. It is therefore quite surprising that in the quantum case, the difference between the case of finite and infinite parameter p disappears. This must nevertheless be tempered with the fact that there is no lower bound in our case, because our definition of the L^p norm only makes sense for absolutely continuous states.

4.3.3.2 The free real sphere

In [Mé14], P.-L. Méliot did not only prove the cutoff phenomenon for compact simple Lie groups, but also for their homogeneous spaces. In the quantum setting, there is no structure theory of compact homogeneous spaces paralleling the classical one, but there are nevertheless some explicit examples. We will now consider the simplest of them, which is an analogue of the real sphere on which the classical orthogonal group acts. The same idea of "liberation" as for the definition of free orthogonal quantum groups suggests that the free analogue of the real sphere should be described by the universal $*$-algebra generated by N self-adjoint elements $(x_i)_{1\leqslant i\leqslant N}$ such that

$$\sum_{i=1}^{N} x_i^2 = 1.$$

Denoting by $\mathcal{O}(S_+^{N-1})$ this object, it is endowed with an action of O_N^+ through the map

$$\alpha : x_i \mapsto \sum_{i=1}^{N} u_{ij} \otimes x_j.$$

Note that the abelianization of $\mathcal{O}(S_+^{N-1})$ is exactly the algebra of polynomial functions on the $N-1$ dimensional sphere in $\mathbf{R}^N$ and that the formula defining α also defines the usual action of O_N on that sphere.

Intuitively, Brownian motion on such a space should be a Lévy process invariant under the action α, and the analogue of the uniform measure should be the unique probability measure invariant under α. Such an α-invariant state does indeed exist and can be constructed in the following way. Consider the subalgebra $\mathcal{O}(X_N) \subset \mathcal{O}(O_N^+)$ generated by the elements u_{i1} for $1 \leqslant i \leqslant N$. Then, there is a surjective $*$-homomorphism $\pi : \mathcal{O}(S_+^{N-1}) \to \mathcal{O}(X_N)$ sending x_i to u_{i1}. Moreover, one has

$$(\pi \otimes \mathrm{id}) \circ \alpha = (\Delta \otimes \mathrm{id}) \circ \pi$$

so that the state $\omega = h \circ \pi$ is invariant under the action α. As a consequence, we will only consider the "concrete" model $\mathcal{O}(X_N)$ instead of $\mathcal{O}(S_+^{N-1})$.

Brownian motion considered in [Mé14] on a homogeneous space is then the projection of Brownian motion coming from the group. In our case, this simply amounts to restricting ψ_t to $\mathcal{O}(X_N)$. Before giving the expression, let us first recall that by [DFW21, Lem 7.3], one may find a basis for the carrier Hilbert space of each irreducible representation u^n of O_N^+ such that

$$\mathcal{O}(X_N) = \mathrm{Span}\{u_{i1}^n \mid 1 \leqslant i \leqslant N, n \in \mathbf{N}\}.$$

Proposition 4.3.20. The Lévy process given by the restriction of ψ_t to $\mathcal{O}(X_N)$ exhibits a cutoff phenomenon at time $t_N = \frac{1}{2}N\ln(N)$. Morevoer, it has the same cutoff profile as Brownian motion on O_N^+.

Proof. For t large enough, the density of $\psi_t - h$ is

$$\sum_{n=1}^{+\infty} d_n e^{-tP_n'(N)/P_n(N)} u_{11}^n$$

whose L^2-norm squared is

$$\sum_{n=1}^{+\infty} d_n e^{-2tP_n'(N)/P_n(N)}.$$

The difference with the previous case is that the dimension d_n is not squared, due to the fact that coefficients of irreducible representations form an orthogonal but not orthonormal basis. This accounts for the factor $1/2$ in the cutoff time, exactly as in [Mé14]. The proof is now exactly the same as for Theorems 4.3.1 and 4.3.9. □

There is also another candidate for a Brownian motion on the real free sphere. Lévy processes on the later quantum space were classified by B. Das, U. Franz and X. Wang in [DFW21, Thm 7.5] using a formula similar to Equation (4.2.1), i.e. involving a positive constant b and a Lévy measure ν. Taking as before $b = 1$ and $\nu = 0$ yields a reasonable notion of a Brownian motion on X_N which *is not* the projection of the one on O_N^+. The convolution semigroup of states $(\varphi_t)_{t\in\mathbf{R}_+}$ we are interested in is then given by :

$$\varphi_t : u_{i1}^n \mapsto \delta_{i1}e^{-tR_n'(1)},$$

where the polynomials $(R_n)_{n\in\mathbf{N}}$ are the orthogonal polynomials associated to the spectral measure of u_{11} (see [BCZJ09] for details and explicit computations).

Proposition 4.3.21. The O_N^+-invariant Lévy process on X_N given by $(\varphi_t)_{t\in\mathbf{R}_+}$ exhibits a cutoff phenomenon at time $t_N = \frac{1}{2}\ln(N)$.

Proof. For t large enough, φ_t has an L^2-density with respect to $\omega = h\circ\pi$ given by

$$\sum_{n=0}^{+\infty} d_n e^{-tR_n'(1)}u_{11}^n$$

so that

$$\|\varphi_t - \omega\|^2 \leqslant \frac{1}{4}\|\rho_t - 1\|_1^2 \leqslant \frac{1}{4}\|\rho_t - 1\|_2^2 = \frac{1}{4}\sum_{n=1}^{+\infty} d_n^2 e^{-2tR_n'(1)}\|u_{11}^n\|_2^2 = \frac{1}{4}\sum_{n=1}^{+\infty} d_n e^{-2tR_n'(1)}.$$

Now, we know from [DFW21, Cor 7.14] that

$$n \leqslant R_n'(1) \leqslant \frac{N-1}{N-2}n.$$

and combining this with [Fre19, Lem 3.3] shows that $t = \frac{1}{2}(\ln(N) + c)$ is enough to ensure the existence of the L^2-density and that

$$\|\varphi_t - \omega\|^2 \leqslant \frac{1}{4}\frac{1}{(1-q^2)}\sum_{n=1}^{+\infty} q^{-n}e^{-2tn} = \frac{1}{4}\frac{1}{(1-q^2)}\frac{1}{qe^{2t}-1} \leqslant \frac{1}{2}\frac{e^{-c}}{1-e^{-c}}$$

and the upper bound follows.

As for the lower bound, it is obtained by using the element $x = \sqrt{N}u_{11}$ instead of χ_1 and the computations are similar to the proof of Theorem 4.3.1. $\square$

Remark 4.3.22. Note that there is an abuse of notations since the polynomials R_n also depend on the integer N. This is different from the case O_N^+ where for all N, the orthogonal polynomials were always the same Chebyshev polynomials P_n. That fact, combined with the cumbersome available descriptions of R_n (see for instance [DFW21, Sec 7.3]), make it difficult to compute the cutoff profile. However, because the spectral density of $\sqrt{N+2}u_{11}$ converges uniformly to the semi-circle distribution when N goes to infinity by [BCZJ09, Thm 5.3], and because all these densities are supported on the closed unit disk, the convergence also holds in moments hence $\sqrt{N+2}^n R_n(X/\sqrt{N+2})$ converges to $P_n(X)$ since the coefficients of the orthogonal polynomials are continuous functions of the moments. It is therefore reasonable to expect that the process has the same cutoff profile as the Brownian motion on O_N^+.

4.4 Quantum permutations

Our second family of examples will be quantum permutations. The quantum permutation groups S_N^+ were introduced by Sh. Wang in [Wan98]. The corresponding $*$-algebra $\mathcal{O}(S_N^+)$ is the quotient of $\mathcal{O}(O_N^+)$ by the relations $u_{ij}^2 = u_{ij}$. The coproduct factors through this and yields the compact quantum group structure. The connection to classical permutation may seem loose from that definition, but one easily shows that if $c_{ij} : S_N \to \mathbf{C}$ is the function sending a permutation σ to $\left(\delta_{\sigma(i)j}\right)_{ij}$, then there is a surjective $*$-homomorphism $\mathcal{O}(S_N^+) \to \mathcal{O}(S_N)$ sending u_{ij} to c_{ij}, and that $\mathcal{O}(S_N)$ is in fact the abelianization of $\mathcal{O}(S_N^+)$. Thus, S_N^+ is a quantum version of S_N somehow like O_N^+ is the quantum version of O_N. Beyond this fact which motivated the original definition, several connections between classical and quantum permutations have emerged which strongly support the idea that S_N^+ is the correct generalization of S_N. An example of particular interest from the probabilistic point of view is the free De Finetti theorem of C. Köstler and R. Speicher [KS09].

The representation theory of S_N^+ is close to that of O_N^+, with the difference that when multiplying two characters (which are still indexed by the integers with $\chi_0 = 1$ and $\chi_0 + \chi_1 = \sum_{i=1}^N u_{ii}$), one gets the formula

$$\chi_1\chi_n = \chi_{n+1} + \chi_n + \chi_{n-1}.$$

The corresponding orthogonal polynomials are then given by the restriction to $[0, 4]$ of $Q_n(t) = P_{2n}(\sqrt{t})$, yielding the free Poisson law $\mathrm{Poiss}^+(1, 1)$ as spectral measure of χ_1 under the Haar state. The associated dimensions of irreducible representations are $d_n = Q_n(N)$. We are now going to study two examples of processes on S_N^+, one continuous and one discrete.

4.4.1 Brownian motion

The natural candidate for Brownian motion on S_N^+ can be constructed exactly as in the case of O_N^+. Indeed, U. Franz, A. Kula and A. Skalski proved in [FKS16, Thm 10.10] a decomposition result for central Lévy processes on S_N^+ involving as before a positive constant b and a Lévy measure ν. Setting $b = 1$ and $\nu = 0$ leads to a central Lévy process. We will again denote by $(\lambda_n)_{n\in\mathbf{N}}$ the sequence determining the process, which is in this case given by

$$\lambda_n = \frac{Q_n'(N)}{Q_n(N)}.$$

The previous arguments carry on almost verbatim to yield the cutoff phenomenon and one can once again describe the cutoff profile as a distance between two free Meixner laws.

Theorem 4.4.1. The central Lévy process defined above exhibits a cutoff phenomenon at time $N\ln(N)$. Moreover, setting again $t_c = N\ln(N) + cN$, for every $c \in \mathbf{R}$, we have

$$\lim_{N\to+\infty} \|\psi_{t_c} - h\| = \left\| D_{\sqrt{1+e^{-c}}}\left(\mathrm{Meix}^+\left(\frac{1-e^{-c}}{\sqrt{1+e^{-c}}}, \frac{-e^{-c}}{1+e^{-c}}\right)\right) * \delta_{e^{-c}} - \mathrm{Meix}^+(0,1)\right\|.$$

Proof. As mentionned in Remark 4.3.17, the existence of the cutoff profile implies the existence of the cutoff phenomenon, hence we will focus on the former. The proof proceeds as for Theorem 4.3.9 and involves similar estimates. We will therefore focus on the features which differ.

Assuming first $c > 0$ and setting

$$G_c(t) = \sum_{n=0}^{+\infty} e^{-cn}Q_n(t),$$

the cutoff profile equals

$$\|G_{2c}(t) - 1\|_1,$$

where the L^1 norm is computed with respect to the spectral measure of χ_1 with respect to the Haar state, which is $\mathrm{Poiss}^+(1,1)$. Note that because P_{2n} is an even function and P_{2n+1} an odd one,

$$G_{2c}(t) = \sum_{n=0}^{+\infty} e^{-2cn} P_{2n}(\sqrt{t}) = \frac{F_c(\sqrt{t}) + F_c(-\sqrt{t})}{2}.$$

Setting $\beta = e^{-c}$ and $\gamma = \beta/(1+\beta^2)$ as in Subsection 4.3.2, this leads to the formula

$$G_{2c}(t) = \frac{1}{2(1+\beta^2)}\left(\frac{1}{1-\gamma\sqrt{t}} + \frac{1}{1+\gamma\sqrt{t}}\right) = \frac{1}{1+\beta^2}\frac{1}{1-\gamma^2 t}.$$

Let us also set

$$\eta = \frac{1-\sqrt{1-4\gamma^2}}{2\gamma^2} = 1+\beta^2.$$

Then, making the changes of variables $u = t-\eta$ and $v = u/\sqrt{\eta}$, and observing that $\gamma^2 = (\eta-1)\eta^{-2}$, the density of $G_{2c}(t)\mathrm{d}\,\mathrm{Poiss}^+(1,1)(t)$ becomes

$$\begin{aligned}
&\frac{1}{\eta}\frac{1}{1-\gamma^2 t}\frac{1}{2\pi t}\sqrt{4-(t-2)^2}\mathbf{1}_{[0,4]}(t)\mathrm{d}t \\
&= \frac{1}{2\pi\eta}\frac{1}{(1-\gamma^2(u+\eta))(u+\eta)}\sqrt{4-(u-(2-\eta))^2}\mathbf{1}_{[-\eta,4-\eta]}(u)\mathrm{d}u \\
&= \frac{1}{2\pi\eta}\frac{1}{(1-(\eta-1)\eta^{-2}(v\sqrt{\eta}+\eta))(v\sqrt{\eta}+\eta)}\sqrt{4-(v\sqrt{\eta}-(2-\eta))^2}\mathbf{1}_{[-\sqrt{\eta},\frac{4}{\sqrt{\eta}}-\eta]}(v)\sqrt{\eta}\mathrm{d}v \\
&= \frac{1}{2\pi}\frac{1}{1+v(2-\eta)/\sqrt{\eta}+v^2(1-\eta)/\eta}\sqrt{4/\eta-(v-(2-\eta)/\sqrt{\eta})^2}\mathbf{1}_{[-\sqrt{\eta},\frac{4}{\sqrt{\eta}}-\eta]}(v)\mathrm{d}v.
\end{aligned}$$

Setting $a = (2-\eta)/\sqrt{\eta}$, and $b = (1-\eta)/\eta$, this is exactly the density of the standardised free Meixner law with parameters a and b,

$$\frac{1}{2\pi}\frac{\sqrt{4(1+b)-(v-a)^2}}{1+av+bv^2}\mathbf{1}_{a-2\sqrt{1+b},a+2\sqrt{1+b}}\mathrm{d}v.$$

Thus, $G_{2c}(t)\mathrm{d}\,\mathrm{Poiss}^+(1,1)(t)$ is the density of the law

$$D_{\sqrt{\eta}}\left(\mathrm{Meix}^+\left(\frac{2-\eta}{\sqrt{\eta}},\frac{1-\eta}{\eta}\right)\right) * \delta_\eta.$$

Writing $\mathrm{Poiss}^+(1,1) = \mathrm{Meix}^+(0,1) * \delta_1$, applying $*\,\delta_{-1}$ on both sides and replacing $2c$ by c now yields the desired result.

Assume now that $c < 0$. Let us first mention that the free Meixner distribution in the statement then has an atom given by the following formula :

$$(1-e^c)\delta_{e^c\sqrt{1+e^{-c}}}.$$

Applying the dilation by a factor $\sqrt{1+e^{-c}}$ and the translation by e^{-c} changes the indices of the Dirac mass into e^c+1+e^{-c} and translating again by 1 to turn $\mathrm{Meix}^+(0,1)$ into $\mathrm{Poiss}^+(1,1)$, we see that the atom of the measures $m_{t_c}^{(N)}$ has to converge to $e^c+2+e^{-c} = (e^{c/2}+e^{-c/2})^2$.

The same argument as in Proposition 4.3.14 yields an explicit formula for the measure of the corresponding classical process at time t_c and the proof is done similarly. Note that by [BB06], the cumulants of $\mathrm{Meix}^+(a, b)$ are polynomials in a and b, hence in our case Laurent polynomials in e^{-c}, so that the analogue of Lemma 4.3.12 holds. □

Here is an interpretation of this result inspired by the analogy with the classical permutation groups. Indeed, the function giving the number of fixed points of a permutation is, in terms of the generators of $\mathcal{O}(S_N)$, $F = \sum c_{ii}$. Therefore, the elements $\chi_1 = \sum u_{ii}$ is the quantum version of the number of fixed points. In particular, its law with respect to the Haar state, which is $\mathrm{Poiss}^+(1, 1)$, can be considered as the "fixed points law for quantum permutations". As a consequence, the difference between Brownian motion and the uniform measure on S_N^+ is asymptotically due to the fact that Brownian motion has "too many fixed points".

4.4.2 Quantum random transpositions

We will conclude with a discrete example, namely the quantum random transposition walk on the quantum permutation group. The reason for this is that the second-named author recently computed the cutoff profile for the classical version of that walk, while nothing is known in the quantum case.

Recall that if μ_{tr} is the uniform measure on the set of transpositions, then the classical random transposition walk has increment distribution

$$\mu = \frac{N-1}{N}\mu_{\mathrm{tr}} + \frac{1}{N}\delta_e.$$

One of the first results in the theory of cutoff phenomenon was the proof by P. Diaconis and M. Shahshahani in [DS81] that the random transposition walk exhibits a cutoff phenomenon at $\frac{1}{2}N\ln(N)$ steps. The second named author proved in [Tey20] that the cutoff profile has the following form : for any $c \in \mathbf{R}$,

$$\mathrm{d_{TV}}\left(\mu^{*\frac{1}{2}(N\ln(N)+cN)}, h\right) \xrightarrow[N\to\infty]{} \mathrm{d_{TV}}\left(\mathrm{Poiss}\left(1+e^{-c}\right), \mathrm{Poiss}(1)\right).$$

Note that Poiss(1) is the asymptotic law of the number of fixed points of a uniformly distributed permutation, which is the same as the law of the trace of a permutation matrix under the Haar measure, i.e. the law of $\chi_1 + \chi_0 = \chi_1 + 1$.

The δ_e-part in the definition of μ appears in a natural way. If we had decided to work with the *pure* random transposition walk, working with the transition law μ_{tr} instead of μ, we would have had periodicity issues, as the signature would alternate from 1 to -1. Including an extra δ_e-part (sometimes refered to as the "laziness" of the walk) in the definition of μ is the way used by P. Diaconis and M. Shahshahani to rule out this problem. Note that in this case it is very natural to have a coefficient $1/N$ if one thinks of the random walk as a card shuffle : spread a deck of N cards on a table and then choose two cards uniformly at random and swap them if they are different. The probability that the same card has been selected twice is then exactly $1/N$.

On the quantum side, there is a natural analogue of μ_{tr} introduced in [Fre19] and denoted by φ_{tr}. This is a central state given on the characters by

$$\varphi_{\mathrm{tr}}(\chi_n) = Q_n(N-2),$$

where $Q_n(N) = P_{2n}(\sqrt{N})$ as in the previous subsection. We may then consider the quantum analogue of μ, the walk given by

$$\varphi = \frac{N-1}{N}\varphi_{\mathrm{tr}} + \frac{1}{N}\varepsilon,$$

where ε is the unique state on $\mathcal{O}(S_N^+)$ such that $\varepsilon(u_{ij}) = \delta_{ij}$ and hence φ is a well-defined state in the Fourier-Stieltjes algebra $C(O_N^+)^*$.

It was proven in [Fre19] that the random walk on S_N^+ corresponding to φ_{tr} exhibits a cutoff phenomenon (with the same caveat as in Remark 4.3.2), and that there is no periodicity issue, unlike for the classical walk. Consequently, in the quantum setting we can also naturally work with the pure random walk without extra laziness. As for the quantum lazy random walk associated with φ, the study of cutoff phenomenon becomes more delicate and was left as an open problem in [Fre19]. In this subsection we will show that the two random walks are asymptotically identical and hence have the same cutoff profile, which we will then compute.

For notational simplicity, we write for a state ψ on $\mathcal{O}(S_N^+)$

$$\psi(n) = \frac{1}{d_n}\psi(\chi_n).$$

It is well-known and easy to see from definition that $\psi^{*k}(n) = \psi(n)^k$ for all $k \in \mathbf{N}$. Lemma 4.2.8 still holds if we replace ψ_t by ψ^{*k} and $d_n^{-1}\psi_t(\chi_n) = e^{-t\lambda_n}$ by $\psi(n)^k$.

4.4.2.1 The pure walk

We start by revisiting the work of [Fre19] concerning the pure quantum transposition walk. There, the upper bound of cutoff phenomenon was shown to happen at time $N\ln(N)/2$, while the lower bound can be deduced from these computations and the methods used in the proof of Theorem 4.3.1.

It is therefore natural to wonder whether the same strategy as in the orthogonal case can also yield the cutoff profile. The answer turns out to be yes, but requires a fine estimate on $\varphi_{\mathrm{tr}}(n)$, in the same spirit as the affine approximation of λ_n in Lemma 4.3.5.

Lemma 4.4.2. *There exist $a_N, b_N < 0$ depending only on N and a function c_N of n such that for all $n \geqslant 0$,*

$$\varphi_{\mathrm{tr}}(n) = e^{a_N n + b_N + c_N(n)}.$$

Moreover, $0 \geqslant c_N(n) \geqslant -2q(\sqrt{N-2})^{4n+2}$.

Proof. First, set for this proof $q = q(\sqrt{N})$ and $p = q(\sqrt{N-2})$, and recall that

$$\varphi_{\mathrm{tr}}(n) = \frac{Q_n(N-2)}{Q_n(N)} = \frac{P_{2n}(\sqrt{N-2})}{P_{2n}(\sqrt{N})} = \frac{p^{-(2n+1)} - p^{2n+1}}{p^{-1} - p}\left(\frac{q^{-(2n+1)} - q^{2n+1}}{q^{-1} - q}\right)^{-1}.$$

Factoring, we obtain

$$\varphi_{\mathrm{tr}}(n) = \frac{p^{-(2n+1)} - p^{2n+1}}{q^{-(2n+1)} - q^{2n+1}}\frac{q^{-1} - q}{p^{-1} - p} = \left(\frac{p}{q}\right)^{-(2n+1)}\frac{q^{-1} - q}{p^{-1} - p}\frac{1 - p^{4n+2}}{1 - q^{4n+2}} = \left(\frac{q}{p}\right)^{2n}\frac{1 - q^2}{1 - p^2}\frac{1 - p^{4n+2}}{1 - q^{4n+2}},$$

so that setting $a_N = 2\ln(q/p)$, $b_N = \ln\left(\frac{1-q^2}{1-p^2}\right)$ and $c_N(n) = \ln\left(\frac{1-p^{4n+2}}{1-q^{4n+2}}\right)$ yields the first part of the statement. As $p \geqslant q$, we have that $b_N \leqslant 0$ and $c_N(n) \leqslant 0$.

Let us now prove the lower bound on $c_N(n)$. Using that for $x < 1$, $\ln(1-x) \leqslant -x$ and for $0 \leqslant x \leqslant 1/2$, $\ln(1-x) \geqslant -2\ln(2)x$, we have

$$c_N(n) = \ln\left(1 - p^{4n+2}\right) - \ln\left(1 - q^{4n+2}\right) \geqslant -2\ln(2)p^{4n+2} + q^{4n+2} \geq -2p^{4n+2}.$$

$\square$

We can now compute the measure of the classical process to prove the convergence of the complete profile, defined through the formula

$$\int_0^N Q_n(x)\mathrm{d}m_k^{(N)}(x) = \varphi(n)^k Q_n(N).$$

Because of Lemma 4.4.2, it is natural to look for an $\widetilde{N}(k)$ such that $Q_n(\widetilde{N}(k)) \approx \varphi_{\mathrm{tr}}^{*k}(n)Q_n(N)$, which leads heuristically to

$$\widetilde{N}(k) = \left(e^{-ka_N/2}q + e^{ka_N/2}q^{-1}\right)^2$$

Theorem 4.4.3. Set

$$\alpha(k) = e^{-a_N k/2+b_N k}\frac{e^{a_N k/2}q^{-1} - e^{-a_N k/2}q}{q^{-1}-q}.$$

Then, for any $c<0$, and N large enough, setting $k_c = \lceil (N\ln(N)+cN)/2\rceil$,

$$m_{k_c}^{(N)} = \alpha(k_c)\delta_{\widetilde{N}(k_c)} + \sum_{n=0}^{+\infty}\left[\varphi_{\mathrm{tr}}^{*k_c}(n)Q_n(N) - Q_n(\widetilde{N}(k_c))\right] Q_n \mathrm{d}\,\mathrm{Poiss}^+(1,1).$$

Moreover, for all $c\in\mathbf{R}$,

$$\left\|\varphi_{\mathrm{tr}}^{*k_c} - h\right\| \underset{N\to+\infty}{\longrightarrow} \left\| D_{\sqrt{1+e^{-c}}}\left(\mathrm{Meix}^+\left(\frac{1-e^{-c}}{\sqrt{1+e^{-c}}}, \frac{-e^{-c}}{1+e^{-c}}\right)\right) * \delta_{e^{-c}} - \mathrm{Meix}^+(0,1)\right\|_{TV}. \tag{4.4.1}$$

Proof. We start by proving (4.4.1) for $c>0$. This follows from the same argument as in Proposition 4.3.10 using the following computations : by Proposition 4.2.10, we have for every fixed $n\geq 1$, as $N\to\infty$,

$$Q_n(N) = P_{2n}(\sqrt{N}) = N^n\left(1-\frac{1}{N}+O\left(\frac{1}{N^2}\right)\right)^{2n-1}$$

and

$$\varphi_{\mathrm{tr}}(n) = \frac{Q_n(N-2)}{Q_n(N)} = \left(1-\frac{2}{N}\right)^n\left(\frac{1-\frac{1}{N-2}+O\left(\frac{1}{N^2}\right)}{1-\frac{1}{N}+O\left(\frac{1}{N^2}\right)}\right)^{2n-1} = \left(1-\frac{2}{N}\right)^n\left(1+O\left(\frac{1}{N^2}\right)\right)^{2n-1}$$

where we recall that $\varphi_{\mathrm{tr}}(n) = d_n^{-1}\varphi_{\mathrm{tr}}(\chi_n)$. For $c\in\mathbf{R}$ fixed and $k_c = \lceil (N\ln(N)+cN)/2\rceil$, we have

$$\begin{aligned}\left(1-\frac{2}{N}\right)^{nk_c} &= \exp\left(n\left\lceil\frac{1}{2}(N\ln(N)+cN)\right\rceil\left(-\frac{2}{N}+O\left(\frac{1}{N^2}\right)\right)\right)\\ &= \exp\left(n\left(-\ln(N)-c+O\left(\frac{\ln(N)}{N}\right)\right)\right)\\ &= \frac{e^{-n(c+o(1))}}{N^n}.\end{aligned}$$

It follows that

$$d_n\varphi_{\mathrm{tr}}(n)^{k_c} = N^n e^{no(1)}\frac{e^{-n(c+o(1))}}{N^n} = e^{-n(c+o(1))} \underset{N\to+\infty}{\longrightarrow} e^{-nc}.$$

Moreover, for N large enough we have $c + o(1) > c/2 > 0$, hence we have a uniform bound $d_n \varphi_{\mathrm{tr}}(n)^{k_c} \leqslant e^{-nc/2}$, which is summable with respect to n and enables us to conclude when $c > 0$.

We now assume, until the end of the proof, that $c < 0$. Let us set $q = q(\sqrt{N})$ and omit the N indices for $a_N, b_N, c_N(n)$ for convenience. Then,

$$\begin{aligned}(q^{-1} - q)\alpha(k)Q_n(\widetilde{N}(k)) &= e^{-ak/2+bk}\left(e^{a(2n+1)k/2}q^{-(2n+1)} - e^{-a(2n+1)k/2}q^{2n+1}\right)\\ &= e^{ank+bk}q^{-(2n+1)} - e^{-a(n+1)k+bk}q^{2n+1}\\ &= (q^{-1}-q)e^{ank+bk}Q_n(N) + e^{ank+bk}q^{2n+1} - e^{-a(n+1)k+bk}q^{2n+1}\\ &= (q^{-1}-q)\varphi_{\mathrm{tr}}(n)^k e^{-c(n)k}Q_n(N) + e^{bk}\left(e^{ank} - e^{-a(n+1)k}\right)q^{2n+1},\end{aligned}$$

so that

$$\left|\varphi_{\mathrm{tr}}^{*k}(n)Q_n(N) - \alpha(k)Q_n(\widetilde{N}(k))\right| \leqslant \left|1 - e^{-c(n)k}\right|\varphi_{\mathrm{tr}}(n)^kQ_n(N) + \frac{e^{bk}q^{-1}}{q^{-1}-q}\left|e^{ank} - e^{-a(n+1)k}\right|q^{2n+2}.$$

Let us now bound both terms for $k = k_c$:

- For the first term we use the fact that $\varphi_{\mathrm{tr}}(n)^kQ_n(N) = O(e^{-cn})$ together with
$$\left|1 - e^{-c(n)k}\right| \leqslant 2|c(n)|k \leqslant 4kq\left(\sqrt{N-2}\right)^{4n+2}.$$
- As for the second term, the fraction involving b in front of the absolute value does not depend on n, hence has no impact on the summability, while $a < 0$ for N large enough so that we can bound the rest by
$$e^{-a(n+1)k}q^{2n+2} = \left(e^{\ln(N)+c+o\left(\frac{\ln(N)}{N}\right)}q^2\right)^{n+1} = \left(e^{c+o\left(\frac{\ln(N)}{N}\right)}\left(1 + O\left(1/N\right)\right)\right)^{n+1}.$$
For N large enough, the term in parenthesis is bounded by $e^{c/2}$, hence the whole thing is summable with respect to n.

The proof is now concluded as in the proof of Proposition 4.3.14 to obtain the measure and the convergence of the profile for $c < 0$. Note that $\alpha(k)$ converges to $1 - e^c$, which is the mass of the discrete part of the limit, and that $\widetilde{N}(k)$ converges to $\left(e^{c/2} + e^{-c/2}\right)^2$.

Remark that the case $c = 0$, which was not treated above, follows from the monotonicity of the distance proven in Lemma 4.2.6 and the continuity of the profile at 0. □

4.4.2.2 The lazy walk

We now turn to the lazy random walk associated with the state

$$\varphi = \frac{N-1}{N}\varphi_{\mathrm{tr}} + \frac{1}{N}\varepsilon.$$

As mentioned previously, the study of the corresponding cutoff phenomenon is subtler. Indeed, as pointed out in [Fre19], the states φ^{*k} never admit L^2-densities and hence the previous method based on Lemma 4.2.8 for $c > 0$ does not work any more. Our idea in the present work will be to approximate the lazy random walk by the pure one by mimicking an alternative classical way of avoiding periodicity for Markov chains, which was used for example to study the cutoff for k-cycles by Berestycki, Schramm and Zeitouni [BSZ11]. It consists in working in continuous time and considering a clock which rings at a random time given by an exponential law of parameter one. Each time the clock rings, we make one step, and reset the clock. Note that the standard deviation of a sum of order $N\ln(N)$ independent variables of law $\mathrm{Exp}(1)$ is of the order $\sqrt{N\ln(N)}$, which is negligible when compared to N, the size of the cutoff window, so it doesn't change the cutoff profile at all. A similar comment can be made about adding extra laziness as long as the laziness coefficient is not too large.

As the next result will show, one can transfer this idea to the quantum setting and the result is formally equivalent to adding laziness. Moreover, it leads to a simple proof of the cutoff phenomenon.

Theorem 4.4.4. The random walk associated to $\varphi = \frac{N-1}{N}\varphi_{\mathrm{tr}} + \frac{1}{N}\varepsilon$ exhibits a cutoff phenomenon at $N\ln(N)/2$ steps in Fourier-Stieltjes norm. Moreover, the associated cutoff profile is given, for every $c \in \mathbf{R}$, by

$$\lim_{N\to+\infty} \left\| \varphi^{*\left\lceil \frac{1}{2}(N\ln(N)+cN)\right\rceil} - h \right\| = y.$$

Proof. As explained in Remark 4.3.17, it suffices to prove the cutoff profile. For each $j \in \mathbf{N}$, we denote by X_j a random variable following the binomial distribution of parameters j and $p = 1 - 1/N$. Then for any $j \in \mathbf{N}$, we have

$$\mathbb{E}\left(\varphi_{\mathrm{tr}}^{*X_j}\right) = \sum_{k=0}^{j} \binom{j}{k} \left(\frac{N-1}{N}\right)^k \left(\frac{1}{N}\right)^{n-k} \varphi_{\mathrm{tr}}^{*k} = \varphi^{*j},$$

where we make the convention that $\varphi_{\mathrm{tr}}^{*0} = \varepsilon$. Write $k_c = \left\lceil \frac{1}{2}(N\ln(N) + cN)\right\rceil$. We have

$$\left\|\mathbb{E}\left(\varphi_{\mathrm{tr}}^{*X_{k_c}}\right) - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS} = \left\|\mathbb{E}\left(\varphi_{\mathrm{tr}}^{*X_{k_c}} - \varphi_{\mathrm{tr}}^{*k_c}\right)\right\|_{FS} \leqslant \mathbb{E}\left\|\varphi_{\mathrm{tr}}^{*X_{k_c}} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS}.$$

Note that $\mathbb{E}\left(X_{k_c}\right) = pk_c = k_c + O(\ln(N))$, so that by the Markov inequality,

$$\mathbb{P}\left(|X_{k_c} - k_c| < \sqrt{N}\right) \xrightarrow[N\to\infty]{} 1.$$

Using this together with the dominated convergence theorem yields

$$\begin{aligned}\lim_{N\to+\infty} \left\|\mathbb{E}\left(\varphi_{\mathrm{tr}}^{*X_{k_c}}\right) - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS} &\leqslant \lim_{N\to+\infty} \mathbb{E}\left(\mathbf{1}_{|X_{k_c}-k_c|<\sqrt{N}} \left\|\varphi_{\mathrm{tr}}^{*X_{k_c}} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS}\right)\\ &= \mathbb{E}\left(\lim_{N\to+\infty} \mathbf{1}_{|X_{k_c}-k_c|<\sqrt{N}} \left\|\varphi_{\mathrm{tr}}^{*X_{k_c}} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS}\right).\end{aligned}$$

Let us first assume $c > 0$. The proof of Theorem 4.4.3 shows that for any sequence $k_N = k_c + o(N)$, we have

$$d_n\varphi_{\mathrm{tr}}(n)^{k_N} = e^{-n(c+o(1))}$$

as $N \to +\infty$, and moreover $d_n\varphi_{\mathrm{tr}}(n)^{k_N} \xrightarrow[N\to\infty]{} e^{-nc}$ and $d_n\varphi_{\mathrm{tr}}(n)^{k_N} \leqslant e^{-nc/2}$ for N large enough. In particular, applying the same strategy of exchange of sum and limit as in the proof of Proposition 4.3.14, we have

$$\left\|\varphi_{\mathrm{tr}}^{*k_N} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS} \xrightarrow[N\to\infty]{} 0.$$

Together with the previous estimates, this yields

$$\lim_{N\to+\infty} \left\|\varphi^{*k_c} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS} \leqslant \mathbb{E}\left(\lim_{N\to+\infty} \mathbf{1}_{|X_{k_c}-k_c|<\sqrt{N}} \left\|\varphi_{\mathrm{tr}}^{*X_{k_c}} - \varphi_{\mathrm{tr}}^{*k_c}\right\|_{FS}\right) = 0.$$

and the result follows from Theorem 4.4.3.

Now assume $c<0$ and consider the measures $m_k^{(N)}$ associated with the pure random walk which were defined just before Theorem 4.4.3. Note that for any $k \in \mathbf{N}$, $\mathbb{E}\left(m_{X_k}^{(N)}\right)$ is a bounded measure on $[0,N]$ such that

$$\left\|\varphi^{*k} - h\right\| = \left\|\mathbb{E}\left(m_{X_k}^{(N)}\right) - \mathrm{Poiss}^{+}(1,1)\right\|_{TV}.$$

Note also that

$$\left\|\mathbf{1}_{[0,4]}\left(\mathbb{E}\left(m_{X_{k_c}}^{(N)}\right) - m_{k_c}^{(N)}\right)\right\|_{TV} = \left\|\mathbb{E}\left(\mathbf{1}_{[0,4]}\left(m_{X_{k_c}}^{(N)} - m_{k_c}^{(N)}\right)\right)\right\|_{TV} \leqslant \mathbb{E}\left\|\mathbf{1}_{[0,4]}\left(m_{X_{k_c}}^{(N)} - m_{k_c}^{(N)}\right)\right\|_{TV}$$

The same argument as for the case $c>0$ yields that the right hand side tends to 0. So together with Theorem 4.4.3, we deduce that

$$\lim_{N\to+\infty}\left\|\mathbf{1}_{[0,4]}\left(\mathbb{E}\left(m_{X_{k_c}}^{(N)}\right) - \mathrm{Poiss}^{+}(1,1)\right)\right\|_{TV} = \lim_{N\to+\infty}\left\|\mathbf{1}_{[0,4]}\left(m_{k_c}^{(N)} - \mathrm{Poiss}^{+}(1,1)\right)\right\|_{TV}.$$

On the other hand, using again Theorem 4.4.3, we see that

$$\begin{aligned}\left\|\mathbf{1}_{\mathbf{R}\setminus[0,4]}\left(\mathbb{E}\left(m_{X_{k_c}}^{(N)}\right) - \mathrm{Poiss}^{+}(1,1)\right)\right\|_{TV} &= \left\|\mathbf{1}_{\mathbf{R}\setminus[0,4]}\mathbb{E}\left(m_{X_{k_c}}^{(N)}\right)\right\|_{TV} = \left\|\mathbb{E}\left(\alpha(X_{k_c})\delta_{\widetilde{N}(X_{k_c})}\right)\right\|_{TV} \\ &= \left\|\sum_{i\in\mathbf{N}}\mathbb{P}(X_{k_c}=i)\alpha(i)\delta_{\widetilde{N}(i)}\right\|_{TV} = \sum_{i\in\mathbf{N}}\mathbb{P}(X_{k_c}=i)\alpha(i) \\ &= \mathbb{E}\left(\alpha(X_{k_c})\right),\end{aligned}$$

which tends to the mass of the discrete part of the free Meixner law in the desired profile. The proof is now concluded as in the proof of Theorem 4.4.3. □

Bibliography

[AB92] David J. Aldous and Mark Brown. Inequalities for rare events in time-reversible Markov chains. I. In *Stochastic inequalities (Seattle, WA, 1991)*, volume 22 of *IMS Lecture Notes Monogr. Ser.*, pages 1–16. Inst. Math. Statist., Hayward, CA, 1992.

[AB22] Yoshihiro Abe and Marek Biskup. Exceptional points of two-dimensional random walks at multiples of the cover time. *Probab. Theory Related Fields*, 183(1-2):1–55, 2022.

[Abe21] Yoshihiro Abe. Second-order term of cover time for planar simple random walk. *J. Theoret. Probab.*, 34(3):1689–1747, 2021.

[AD86] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986.

[AF] David Aldous and James A. Fill. Local expansion of symmetrical graphs.

[Ald82] David J. Aldous. Some inequalities for reversible Markov chains. *J. London Math. Soc. (2)*, 25(3):564–576, 1982.

[Ald83a] David Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin, 1983.

[Ald83b] David J. Aldous. On the time taken by random walks on finite groups to visit every state. *Z. Wahrsch. Verw. Gebiete*, 62(3):361–374, 1983.

[Ald89] David Aldous. Hitting times for random walks on vertex-transitive graphs. *Math. Proc. Cambridge Philos. Soc.*, 106(1):179–191, 1989.

[Bab89] László Babai. The probability of generating the symmetric group. *J. Combin. Theory Ser. A*, 52(1):148–153, 1989.

[Ban96] T. Banica. Théorie des représentations du groupe quantique compact libre $O(n)$. *C. R. Acad. Sci. Paris Sér. I Math.*, 322(3):241–244, 1996.

[Ban97] T. Banica. Le groupe quantique compact libre $U(n)$. *Comm. Math. Phys.*, 190(1):143–172, 1997.

[BB06] M. Bożejko and W. Bryc. On a class of free Lévy laws related to a regression problem. *J. Funct. Anal.*, 236(1):59–77, 2006.

[BCMR21] Michael E. Bate, Stephen B. Connor, and Oliver Matheau-Raven. Cutoff for a one-sided transposition shuffle. *Ann. Appl. Probab.*, 31(4):1746–1773, 2021.

[BCZJ09] T. Banica, B. Collins, and P. Zinn-Justin. Spectral analysis of the free orthogonal matrix. *Int. Math. Res. Not.*, 2009(17):3286–3309, 2009.

[BD92] Dave Bayer and Persi Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, 2(2):294–313, 1992.

[Bel13] David Belius. Gumbel fluctuations for cover times in the discrete torus. *Probab. Theory Related Fields*, 157(3-4):635–689, 2013.

[Ber16] Nathanaël Berestycki. Mixing times of markov chains: Techniques and examples. 2016.

[Ber18] Megan Bernstein. A random walk on the symmetric group generated by random involutions. *Electron. J. Probab.*, 23:Paper No. 26, 28, 2018.

[BGJ22] M. Brannan, L. Gao, and M. Junge. Complete Logarithmic Sobolev Inequalities via Ricci Curvature Bounded Below II. *J. Topol. Anal.*, 2022.

[BH91] MJAM Brummelhuis and HJ Hilhorst. Covering of a finite lattice by a random walk. *Physica A: Statistical Mechanics and its Applications*, 176(3):387–408, 1991.

[BH92] MJAM Brummelhuis and HJ Hilhorst. How a random walk covers a finite lattice. *Physica A: Statistical Mechanics and its Applications*, 185(1-4):35–44, 1992.

[BH05] László Babai and Thomas P. Hayes. Near-independence of permutations and an almost sure polynomial bound on the diameter of the symmetric group. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1057–1066. ACM, New York, 2005.

[BHP17] Riddhipratim Basu, Jonathan Hermon, and Yuval Peres. Characterization of cutoff for reversible Markov chains. *Ann. Probab.*, 45(3):1448–1487, 2017.

[BHT22] Nathanaël Berestycki, Jonathan Hermon, and Lucas Teyssier. On the universality of fluctuations for the cover time. 2022.

[Bia08] Philippe Biane. Introduction to random walks on noncommutative spaces. In *Quantum potential theory*, pages 61–116. Springer, 2008.

[BK05] Itai Benjamini and Gady Kozma. A resistance bound via an isoperimetric inequality. *Combinatorica*, 25:645–650, 2005.

[BK17] David Belius and Nicola Kistler. The subleading order of two dimensional cover times. *Probab. Theory Related Fields*, 167(1-2):461–552, 2017.

[Bla06] B. Blackadar. *Operator algebras*, volume 122 of *Encyclopædia of Mathematical Sciences*. Springer, 2006.

[BLPS18] Nathanaël Berestycki, Eyal Lubetzky, Yuval Peres, and Allan Sly. Random walks on the random graph. *Ann. Probab.*, 46(1):456–490, 2018.

[Blu11] Elena Blumberg. *Permutations with Interval Restrictions.* PhD thesis, Stanford University, 2011.

[BN19] Megan Bernstein and Evita Nestoridi. Cutoff for random to random card shuffle. *Ann. Probab.*, 47(5):3303–3320, 2019.

[BN22] Alexey Bufetov and Peter Nejjar. Cutoff profile of ASEP on a segment. *Probab. Theory Related Fields*, 183(1-2):229–253, 2022.

[BP99] H. Bercovici and V. Pata. Stable laws and domains of attraction in free probability theory. With an appendix by P. Biane. *Ann. Math.*, pages 1023–1060, 1999.

[BR17] M. Brannan and Z-J. Ruan. L_p-representations of discrete quantum groups. *J. Reine Angew. Math.*, 732:165–210, 2017.

[Bra12] M. Brannan. Approximation properties for free orthogonal and free unitary quantum groups. *J. Reine Angew. Math.*, 672:223–251, 2012.

[Bra17] M. Brannan. Approximation properties for locally compact quantum groups. *Banach Center Publications*, 111:185–232, 2017.

[BRZ20] David Belius, Jay Rosen, and Ofer Zeitouni. Tightness for the cover time of the two dimensional sphere. *Probab. Theory Related Fields*, 176(3-4):1357–1437, 2020.

[BS09] T. Banica and R. Speicher. Liberation of orthogonal Lie groups. *Adv. Math.*, 222(4):1461–1501, 2009.

[BŞ19] Nathanaël Berestycki and Batı Şengül. Cutoff for conjugacy-invariant random walks on the permutation group. *Probab. Theory Related Fields*, 173(3-4):1197–1241, 2019.

[BSZ11] Nathanaël Berestycki, Oded Schramm, and Ofer Zeitouni. Mixing times for random k-cycles and coalescence-fragmentation chains. *Ann. Probab.*, 39(5):1815–1843, 2011.

[CE22] Yuansi Chen and Ronen Eldan. Localization schemes: a framework for proving mixing bounds for Markov chains. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science—FOCS 2022*, pages 110–122. IEEE Computer Soc., Los Alamitos, CA, [2022] ©2022.

[CFK14] F. Cipriani, U. Franz, and A. Kula. Symmetries of Lévy processes, their Markov semigroups and potential theory on compact quantum groups. *J. Funct. Anal.*, 266(5):2789–2844, 2014.

[Chu97] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997.

[CSC08] G. Chen and L. Saloff-Coste. The cutoff phenomenon for ergodic Markov processes. *Electronic J. Probab.*, 13:26–78, 2008.

[DDMP19] Amir Dembo, Jian Ding, Jason Miller, and Yuval Peres. Cut-off for lamplighter chains on tori: dimension interpolation and phase transition. *Probab. Theory Related Fields*, 173(1-2):605–650, 2019.

[DFW21] B. Das, U. Franz, and X. Wang. Invariant Markov semigroups on quantum homogeneous spaces. *J. Noncommut. Geom.*, 15(2):531–580, 2021.

[DHESC21] Persi Diaconis, Kelsey Houston-Edwards, and Laurent Saloff-Coste. Gambler's ruin estimates on finite inner uniform domains. *Ann. Appl. Probab.*, 31(2):865–895, 2021.

[Dia88] Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988.

[Din14] Jian Ding. Asymptotics of cover times via Gaussian free fields: bounded-degree graphs and general trees. *Ann. Probab.*, 42(2):464–496, 2014.

[DLP12] Jian Ding, James R. Lee, and Yuval Peres. Cover times, blanket times, and majorizing measures. *Ann. of Math. (2)*, 175(3):1409–1471, 2012.

[DM09] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.

[DM15] Persi Diaconis and Laurent Miclo. On quantitative convergence to quasi-stationarity. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):973–1016, 2015.

[DP22] Persi Diaconis and Soumik Pal. Shuffling cards by spatial motion. *Stochastic Process. Appl.*, 152:149–176, 2022.

[DPRZ04] Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math. (2)*, 160(2):433–464, 2004.

[DS81] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981.

[DSC93] Persi Diaconis and Laurent Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.*, 21(4):2131–2156, 1993.

[DSC94] P. Diaconis and L. Saloff-Coste. Moderate growth and random walk on finite groups. *Geom. Funct. Anal.*, 4(1):1–36, 1994.

[ER61] Paul Erdős and Alfréd Rényi. On a classical problem of probability theory. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 6(1-2):215–220, 1961.

[FHL+17] U. Franz, G. Hong, F. Lemeux, M. Ullrich, and H. Zhang. Hypercontractivity of heat semigroups on free quantum groups. *J. Operator Theory*, 77(1):61–76, 2017.

[FJR03] Avital Frumkin, Gordon James, and Yuval Roichman. On trees and characters. *J. Algebraic Combin.*, 17(3):323–334, 2003.

[FKS16] U. Franz, A. Kula, and A. Skalski. Lévy processes on quantum permutation groups. In *Noncommutative analysis, operator theory and applications*, pages 193–259. Birkhäuser, 2016.

[Fol11] Matthew Folz. Gaussian upper bounds for heat kernels of continuous time simple random walks. *Electron. J. Probab.*, 16:no. 62, 1693–1722, 2011.

[Fra06] U. Franz. Lévy processes on quantum groups and dual groups. In M. Schürmann and U. Franz, editors, *Quantum independent increment processes II*, volume 1866 of *Lecture Notes in Mathematics*, pages 161–257. Springer, 2006.

[Fre18] A. Freslon. Quantum reflections, random walks and cut-off. *Internat. J. Math.*, 29(14):1850101, 2018.

[Fre19] Amaury Freslon. Cut-off phenomenon for random walks on free orthogonal quantum groups. *Probab. Theory Related Fields*, 174(3-4):731–760, 2019.

[FTW22] Amaury Freslon, Lucas Teyssier, and Simeng Wang. Cutoff profiles for quantum Lévy processes and quantum random transpositions. *Probab. Theory Related Fields*, 183(3-4):1285–1327, 2022.

[GG] Shirshendu Ganguly and Reza Gheissari. Cutoff for the glauber dynamics of the lattice free field.

[Gro81] Mikhael Gromov. Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.*, (53):53–73, 1981.

[HLP16] Jonathan Hermon, Hubert Lacoin, and Yuval Peres. Total variation and separation cutoffs are not equivalent and neither one implies the other. *Electron. J. Probab.*, 21:Paper No. 44, 36, 2016.

[Hou16] Bob Hough. The random k cycle walk on the symmetric group. *Probab. Theory Related Fields*, 165(1-2):447–482, 2016.

[HP18] Jonathan Hermon and Yuval Peres. A characterization of L_2 mixing and hypercontractivity via hitting times and maximal inequalities. *Probab. Theory Related Fields*, 170(3-4):769–800, 2018.

[HS21] Jonathan Hermon and Justin Salez. The interchange process on high-dimensional products. *Ann. Appl. Probab.*, 31(1):84–98, 2021.

[HSZ15] Harald A. Helfgott, Ákos Seress, and Andrzej Zuk. Random generators of the symmetric group: diameter, mixing time and spectral gap. *J. Algebra*, 421:349–368, 2015.

[Jeg20] Antoine Jego. Planar Brownian motion and Gaussian multiplicative chaos. *Ann. Probab.*, 48(4):1597–1643, 2020.

[Ker03] S. V. Kerov. *Asymptotic representation theory of the symmetric group and its applications in analysis*, volume 219 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2003. Translated from the Russian manuscript by N. V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski.

[KS09] C. Köstler and R. Speicher. A noncommutative de Finetti theorem : invariance under quantum permutations is equivalent to freeness with amalgamation. *Comm. Math. Phys.*, 291(2):473–490, 2009.

[Lac16a] Hubert Lacoin. The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.*, 44(5):3399–3430, 2016.

[Lac16b] Hubert Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *The Annals of Probability*, 44(2):1426 – 1487, 2016.

[Lia04] M. Liao. Lévy processes and Fourier analysis on compact Lie groups. *Ann. Probab.*, pages 1553–1573, 2004.

[LMS08] Russell Lyons, Benjamin J. Morris, and Oded Schramm. Ends in uniform spanning forests. *Electron. J. Probab.*, 13:no. 58, 1702–1725, 2008.

[LOG18] Russell Lyons and Shayan Oveis Gharan. Sharp bounds on random walk eigenvalues via spectral embedding. *Int. Math. Res. Not. IMRN*, (24):7555–7605, 2018.

[LP02] Nathan Lulov and Igor Pak. Rapidly mixing random walks and bounds on characters of the symmetric group. *J. Algebraic Combin.*, 16(2):151–163, 2002.

[LP16a] Eyal Lubetzky and Yuval Peres. Cutoff on all ramanujan graphs. *Geometric and Functional Analysis*, 26(4):1190–1216, 2016.

[LP16b] Russell Lyons and Yuval Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016.

[LP17] David A. Levin and Yuval Peres. *Markov chains and mixing times.* American Mathematical Society, Providence, RI, 2017. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[LS08] Michael Larsen and Aner Shalev. Characters of symmetric groups: sharp bounds and applications. *Invent. Math.*, 174(3):645–687, 2008.

[LS13] Eyal Lubetzky and Allan Sly. Cutoff for the Ising model on the lattice. *Invent. Math.*, 191(3):719–755, 2013.

[LS17] Eyal Lubetzky and Allan Sly. Universality of cutoff for the Ising model. *Ann. Probab.*, 45(6A):3664–3696, 2017.

[Mat88a] Peter Matthews. Covering problems for Markov chains. *Ann. Probab.*, 16(3):1215–1228, 1988.

[Mat88b] Peter Matthews. A strong uniform time for random transpositions. *J. Theoret. Probab.*, 1(4):411–423, 1988.

[McC19] J.P. McCarthy. Diaconis-Shahshahani upper bound lemma for finite quantum groups. *J. Fourier Anal. App.*, 25:2463—-2491, 2019.

[Mor94] Moshe Morgenstern. Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *J. Combin. Theory Ser. B*, 62(1):44–62, 1994.

[MP05] B. Morris and Yuval Peres. Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields*, 133(2):245–266, 2005.

[MP12] Jason Miller and Yuval Peres. Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *Ann. Probab.*, 40(2):535–577, 2012.

[MS17] Jason Miller and Perla Sousi. Uniformity of the late points of random walk on $\mathbb{Z}_n^d, d \geq 3$. *Probability Theory and Related Fields*, 167(3):1001–1056, 2017.

[Mé14] Pierre-Loïc Méliot. The cut-off phenomenon for Brownian motions on compact symmetric spaces. *Potential Anal.*, 40(4):427–509, 2014.

[Mé17] Pierre-Loïc Méliot. *Representation Theory of Symmetric Groups (1st ed.).* Chapman and Hall/CRC., 2017.

[Nes21] Evita Nestoridi. The limit profile of star transpositions, https://arxiv.org/abs/2209.12859. 2021.

[NOT21] E. Nestoridi and S. Olesker-Taylor. Limit profiles for reversible markov chains. *Probab. Theory Related Fields*, 2021.

[NOT22] E. Nestoridi and S. Olesker-Taylor. Limit profile for projections of random walks on groups, https://arxiv.org/abs/2209.12859. 2022.

[NS06] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *Lecture Note Series*. London Mathematical Society, 2006.

[NT13] S. Neshveyev and L. Tuset. *Compact quantum groups and their representation categories*, volume 20 of *Cours Spécialisés*. Société Mathématique de France, 2013.

[OTS20] Sam Olesker-Taylor and Perla Sousi. Chen–Stein method for the uncovered set of random walk on $\mathbb{Z}_n^d, d \geq 3$. *Electronic Communications in Probability*, 25:1–11, 2020.

[PdP12] Alan Prata de Paula. *Stochastic processes on finite graphs*. PhD thesis, Instituto de Matemática Pura e Aplicada, 2012.

[PR04] Yuval Peres and David Revelle. Mixing times for random walks on finite lamplighter groups. *Electron. J. Probab.*, 9:no. 26, 825–845, 2004.

[PRS23] Alexis Prévost, Pierre-François Rodriguez, and Perla Sousi. Phase transition for the late points of random walk. *In preparation*, 2023.

[PS15] Yuval Peres and Perla Sousi. Mixing times are hitting times of large sets. *J. Theoret. Probab.*, 28(2):488–519, 2015.

[Roi96] Yuval Roichman. Upper bound on the characters of the symmetric groups. *Invent. Math.*, 125(3):451–485, 1996.

[Sal20] Justin Salez. Temps de mélange des chaînes de markov, notes de cours. 2020.

[Sal22] Justin Salez. Cutoff for non-negatively curved markov chains. *Journal of the European Mathematical Society, to appear*, 2022.

[Sch] Dominik Schmid. Mixing times for the tasep in the maximal current phase.

[Sch05] Oded Schramm. Compositions of random transpositions. *Israel J. Math.*, 147:221–243, 2005.

[SE] Dominik Schmid and Dor Elboim. Mixing times and cutoff for the tasep in the high and low density phase.

[Ser78] Jean-Pierre Serre. *Représentations linéaires des groupes finis*. Hermann, Paris, revised edition, 1978.

[SS] Dominik Schmid and Allan Sly. Mixing times for the tasep on the circle.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Szn10] Alain-Sol Sznitman. Vacant set of random interlacements and percolation. *Ann. of Math. (2)*, 171(3):2039–2087, 2010.

[Tak02] M. Takesaki. *Theory of operator algebras I*, volume 124 of *Encyclopædia of Mathematical Sciences*. Springer, 2002.

[Tak03] M. Takesaki. *Theory of operator algebras II*, volume 125 of *Encyclopædia of Mathematical Sciences*. Springer, 2003.

[Tey20] Lucas Teyssier. Limit profile for random transpositions. *Ann. Probab.*, 48(5):2323–2343, 2020.

[Tim08] T. Timmermann. *An invitation to quantum groups and duality. From Hopf algebras to multiplicative unitaries and beyond.* EMS Textbooks in Mathematics. European Mathematical Society, 2008.

[Tro84] V. I. Trofimov. Graphs with polynomial growth. *Mat. Sb. (N.S.)*, 123(165)(3):407–421, 1984.

[TT20] Romain Tessera and Matthew Tointon. Sharp relations between volume growth, isoperimetry and resistance in vertex-transitive graphs. 2020.

[TT21] Romain Tessera and Matthew C. H. Tointon. A finitary structure theorem for vertex-transitive graphs of polynomial growth. *Combinatorica*, 41(2):263–298, 2021.

[Voi96] Michael Voit. Asymptotic distributions for the Ehrenfest urn and related random walks. *J. Appl. Probab.*, 33(2):340–356, 1996.

[Wan95] Sh. Wang. Free products of compact quantum groups. *Comm. Math. Phys.*, 167(3):671–692, 1995.

[Wan98] Sh. Wang. Quantum symmetry groups of finite spaces. *Comm. Math. Phys.*, 195(1):195–211, 1998.

[Whi19] Graham White. A strong stationary time for random transpositions. 2019.

[Wil03] David Bruce Wilson. Mixing time of the Rudvalis shuffle. *Electron. Comm. Probab.*, 8:77–85, 2003.

[Wor98] S.L. Woronowicz. Compact quantum groups. *Symétries quantiques (Les Houches, 1995)*, pages 845–884, 1998.

[Zha18] Alex Zhai. Exponential concentration of cover times. *Electron. J. Probab.*, 23:Paper No. 32, 22, 2018.