THE GENERALIZED POLAND-SCHERAGA MODEL: A BIVARIATE RENEWAL APPROACH TO DNA DENATURATION

Maha Khatib

UNIVERSITÉ PARIS DIDEROT-PARIS 7
ÉCOLE DOCTORALE DE SCIENCES MATHÉMATIQUES
DE PARIS CENTRE

THÈSE DE DOCTORAT

Discipline : Mathématiques appliquées

Présentée par

Maha KHATIB

THE GENERALIZED POLAND-SCHERAGA MODEL: A BIVARIATE RENEWAL APPROACH TO DNA DENATURATION

Thèse dirigée par **Giambattista GIACOMIN** et **Arnaud LE NY**

Soutenue le 12 octobre 2016 devant le Jury composé de :

M. Giambattista Giacomin	Université Paris Diderot	Directeur
M. Arnaud Le Ny	Université Paris-est	Directeur
M. Matthias Birkner	Université Johannes Gutenberg de Mayence	Rapporteur
M. Philippe Carmona	Université de Nantes	Rapporteur
M. Fabio Toninelli	Université Claude Bernard Lyon 1	Examinateur
M. Julien Poisat	Université Paris Dauphine	Examinateur

À la mémoire de ma grand mère ...

Le destin ne nous a pas laissé le temps pour jouir ce bonheur ensemble. Puisse Dieu tout puissant t'accorder sa clémence, sa miséricorde et t'accueillir dans son saint paradis.

Remerciements

Je tiens tout d'abord à manifester ma plus profonde et sincère reconnaissance envers mon directeur de thèse Giambattista Giacomin. Il a toujours été disponible à l'écoute de mes questions malgré son planning chargé. Sa capacité d'analyse et son enthousiasme m'a montré que le monde de la recherche pouvait être un univers passionnant. Je lui suis reconnaissante de m'avoir fait bénéficier tout au long de ce travail de sa grande compétence, de sa rigueur intellectuelle et de son efficacité certaine que je n'oublierai jamais.

Je tiens ensuite à remercier Arnaud Le Ny pour avoir co-encadré cette thèse. Il a toujours fait preuve d'une totale confiance, d'un soutien et d'une gentillesse permanente à mon égard. Je me dois également de le remercier pour ses précieux conseils reflétant parfaitement sa disponibilité et sa patience et pour avoir participé à la correction de l'introduction de cette thèse. J'ai eu la chance de collaborer pendant ce travail de thèse avec Quentin Berger, je lui suis reconnaissante du temps qu'il a accordé.

Je souhaite remercier très sincèrement M. Matthias Birkner et M. Philippe Carmona pour avoir accepté d'être les rapporteurs de cette thèse. Merci également aux membres du jury d'avoir accepté d'assister à la présentation de ce travail.

Je dois remercier également les deux laboratoires qui m'ont accueillie ces trois dernières années : le LPMA à Paris 7 et le LAMA à Paris est-Créteil. Merci à tous leurs membres de faire de ces deux endroits des lieux agréables et propices au travail. Un grand merci au personnel administratif dans les deux laboratoires pour leur grande efficacité et leur sympathie : Nathalie Bergame, Pascal Chiettini, Anaïs Delgado, Elise Delos, Florence Deschamps, Valérie Juvé. Cette partie serait incomplète si je n'y adressais pas mes remerciements aux doctorants que j'ai rencontrés durant cette thèse. J'en profite aussi pour remercier toutes les personnes qui m'ont aidé pour faire mes études en France, en particulier M. Rached Mneimné.

Plus éloignées de mon parcours mathématique, certaines personnes qui me suivent depuis mon arrivée en France me sont très chères, et méritent plus qu'une ligne de remerciement dans ma thèse. Un grand merci à Ali et Ghena, Charbel, Hussein, Houssam, Mahran et Heba, Nicolas, Nadine, Rana, Rania, Rawad, Serena et Syrine. Un remerciement spécial pour les filles "Barabis": Marwa, Rana, Rania et Syrine qui m'ont soutenue de près

et ce n'est pas ici le cadre adapté pour détailler tout ce qu'elles ont été pour moi, mais il est certain que leur amitié a été déterminante pendant ces années.

Je ne pourrais jamais oublier le soutien et l'aide des personnes chères de ma nombreuse et merveilleuse famille. Je réserve une reconnaissance particulière à mon oncle Safi et sa magnifique famille Aileen, Jawad, Ayman et Ali. Sans votre soutien et votre amour sans faille, tout ceci n'aurai jamais pu aboutir et je vous serai éternellement reconnaissant d'avoir su me donner les moyens d'arriver jusqu'ici. Aucun mot ne saurait décrire ma pensée !

La thèse a parfois été un moment difficile pour mes proches. Elle est très preneuse de temps ! Et j'avoue ne pas les avoir consacré le temps qu'ils le méritent. Mes parents, mes soeurs Mona et Dima et mon frère "Haidoura", je vous remercie pour vos encouragements et votre soutien. Ils ont toujours cru en moi et j'espère être toujours à leur hauteur. Un remerciement spécial pour mon père, qui sans lui, je n'aurai jamais eu l'opportunité de continuer mes études en France et effectuer cette thèse.

Mes derniers remerciements vont à Marwan, pour l'encouragement et la confiance qu'il m'a témoigné. Tu as su me remonter le moral quand il était au plus bas, tu as su supporter mes changements d'humeurs et mes doutes... Je te dois tellement, merci.

Le modèle de Poland-Scheraga généralisé: une approche de renouvellement bidimensionnel pour la dénaturation de l'ADN.

Le modèle de Poland-Scheraga (PS) est le modèle standard pour étudier la transition de dénaturation de deux brins d'ADN complémentaires et de même longueur. Ce modèle a fait l'objet d'une attention remarquable car il est exactement résoluble dans sa version homogène. Le caractère résoluble est lié au fait que le modèle PS homogène peut être mis en correspondance avec un processus de renouvellement discret. Dans la littérature biophysique une généralisation du modèle, obtenue en considérant des brins non complémentaires et de longueurs différentes, a été considérée et le caractère résoluble s'étend à cette généralisation substantielle.

Dans cette thèse, nous présentons une analyse mathématique du modèle de Poland-Scheraga généralisé. Nous considérons d'abord le modèle homogène et nous exploitons que les deux brins de la chaîne peuvent être modélisés par un processus de renouvellement en deux dimensions. La distribution $K(\cdot)$ de l'emplacement (bidimensionnel) du premier contact entre les deux brins est supposée de la forme $K(n+m) = (n+m)^{-\alpha-2}L(n+m)$ avec $\alpha \geq 0$ et $L(\cdot)$ à variation lente et correspond à une boucle avec n bases dans le premier brin et m dans le deuxième. Nous étudions la transition de localisation-délocalisation et nous montrons l'existence des transitions à l'intérieur de la phase localisée. Nous présentons ensuite des estimations précises sur les propriétés de chemin du modèle.

Ensuite, nous étudions la version désordonnée du modèle en incluant une séquence de variables aléatoires indépendantes identiquement distribuées à deux indices. Nous nous concentrons sur l'influence du désordre sur la transition de dénaturation: nous voulons déterminer si la présence des inhomogénéités modifie les propriétés critiques du système par rapport au cas homogène. Nous prouvons que le désordre est non pertinent si $\alpha < 1$ et nous montrons que pour $\alpha > 1$, les points critiques gelés et recuits diffèrent (basant sur les techniques de coarse graining et la méthode des moments fractionnaires), ce qui prouve la présence d'un régime de désordre pertinent.

Mots clés: Modèle de Poland-Scheraga, Dénaturation de l'ADN, Modèle d'accrochage du polymère, Processus de renouvellement bidimensionnel, Propriétés de chemin, Comportement critique, Pertinence du désordre, Coarse Graining.

The generalized Poland-Scheraga model: a bivariate renewal approach to DNA denaturation.

The Poland-Scheraga (PS) model is the standard basic model to study the denaturation transition of two complementary and equally long strands of DNA. This model has enjoyed a remarkable attention because it is exactly solvable in its homogeneous version. The solvable character is related to the fact that the homogeneous PS model can be mapped to a discrete renewal process. In the bio-physical literature a generalization of the model, allowing different length and non complementarity of the strands, has been considered and the solvable character extends to this substantial generalization.

In this thesis we present a generalized version of the PS model that allows mismatches and non complementary strands (in particular, the two strands may be of different lengths). We consider first the homogeneous model and we exploit that this model can be mapped to a bivariate renewal process. The distribution $K(\cdot)$ of the location (in two dimensions) of the first contact between the two strands is assumed to be of the form $K(n+m) = (n+m)^{-\alpha-2}L(n+m)$ with $\alpha \geq 0$ and $L(\cdot)$ slowly varying and corresponds to a loop with n bases in the first strand and m in the second. We study the localization-delocalization transition and we prove the existence of transitions inside the localized regime. We then present precise estimates on the path properties of the model.

We then study the disordered version of the model by including a sequence of independent and identically distributed random variables with two indices. We focus on the influence of disorder on the denaturation transition: we want to determine whether the presence of randomness modifies the critical properties of the system with respect to the homogeneous case. We prove that the disorder is irrelevant if $\alpha < 1$. We show also that for $\alpha > 1$, the quenched and annealed critical points differ (basing on coarse graining techniques and fractional moment method), proving the presence of a relevant disorder regime.

Keywords: Poland-Scheraga model, DNA Denaturation, Polymer Pinning Model, Bivariate Renewal Processes, Path Properties, Critical Behavior, Disorder Relevance, Coarse Graining.

Contents

Chapter 1

Introduction

Résumé

Dans ce chapitre, nous parlons d'abord des motivations physiques et nous donnons un aperçu des travaux établis dans la littérature sur les différents modèles de l'ADN et plus précisément sur le modèle de Poland-Scheraga. Nous rappelons que la version homogène de ce modèle est liée au processus de renouvellement discret unidimensionnel et nous regroupons les principaux résultats du modèle d'accrochage sur un défaut unidimensionnel. Ensuite nous introduisons une généralisation du modèle de Poland-Scheraga et nous verrons qu'elle est liée à un modèle d'accrochage bidimensionnel qui constitue l'objet principal de cette thèse. Enfin nous présentons les résultats obtenus et nous citons quelques questions ouvertes.

Nous donnons quelques abréviations et notations utilisées dans la suite de la thèse:

- i.i.d.: indépendant(es) et identiquement distribué(es).
- p.s.: presque sûrement (a.s.: almost surely).
- $\mathbf{1}_A$: la fonction indicatrice d'un ensemble A, i.e $\mathbf{1}_A(x) = 1$ si x appartient à A et 0 sinon.
- $\langle x, y \rangle$: produit scalaire canonique sur $\mathbb{R}^d$.
- $\lceil (x_1, x_2) \rceil = (\lceil x_1 \rceil, \lceil x_2 \rceil)$ (resp. $\lfloor (x_1, x_2) \rfloor = (\lfloor x_1 \rfloor, \lfloor x_2 \rfloor)$) où $\lceil x \rceil$ (resp. $\lfloor x \rfloor$) est la partie entière supérieure (resp. inférieure) d'un réel x.
- $x \wedge y$: $\min(x, y)$ et $x \vee y$: $\max(x, y)$.
- $\underline{u} = (u_1, ..., u_d) \in \mathbb{R}^d$ et $|\underline{u}| = \sum_{i=1}^d u_i$.

Soient f et g deux fonctions de la variable réelle x, on présente les notations suivantes lorsque x tend vers une valeur $a \in \mathbb{R} \cup \{-\infty, +\infty\}$:

- $f(x) = o(g(x))$: $f(x)/g(x) \to 0$ lorsque $x \to a$.
- $f(x) = O(g(x))$: $f(x)/g(x)$ est borné lorsque $x \to a$.
- $f(x) \sim g(x)$: $f(x)/g(x) \to 1$ lorsque $x \to a$.
- $f(x) \asymp g(x)$: $f(x) = O(g(x))$ et $g(x) = O(f(x))$ lorsque $x \to a$.

1.1 Motivations physiques et modélisations

1.1.1 Polymères

Un polymère est une grande molécule, organique ou inorganique, constituée d'unités fondamentales appelées monomères et reliées par des liaisons covalentes. Les travaux du chimiste Hermann Staudinger dans les années 1920 constituent la base de la science des polymères ou macromolécules (Nobel en 1953). Ces monomères peuvent exister sous plusieurs types (chaîne inhomogène) et interagissent entre eux et avec leur environnement. Les polymères sont classifiés dans deux grandes familles:

- homopolymères: comportent des motifs monomères tous identiques et interviennent dans de nombreux champs industriels (alimentaire, industrie automobile, bâtiment,...).

- copolymères (hétéropolymères): admettent plusieurs sortes de monomères le long de leur chaîne (comme l'ADN et l'ARN). L'agencement des monomères peut être périodique (un même ordre de monomères qui se répète) ou désordonné (pas d'ordre précis).

Les auto-interactions entre les différents monomères de la chaîne et les interactions externes avec l'environnement sont très complexes et posent beacoup de problèmes et de questions. Grâce à leurs propriétés physiques remarquables, les polymères étaient depuis longtemps l'objet des études expérimentales des physiciens [34, 38, 43] et plus récemment, des études théoriques des mathématiciens [25, 42, 48, 50, 59].

De nombreux phénomènes physiques d'un polymère sont connus: transition de dénaturation thermique de l'ADN, comportement d'un polymère dans une solution hétérogène, accrochage d'un polymère sur une surface solide, interaction d'un polymère avec une interface entre deux solvants, etc. L'étude de ces modèles se concentre sur l'existence d'une transition de phase qui apparaît à une certaine température critique T_c et sur le comportement du système près de cette valeur. Une telle transition provient du fait qu'à haute température, les interactions chimiques avec l'environnement sont négligées devant les effets entropiques et un phénomène de délocalisation apparaît. À basse température, l'influence de l'environnement sur le système devient très importante (phénomène de localisation).

La dénaturation de l'ADN est l'exemple le plus répandu de ces phénomènes. La description de la structure statique de l'ADN admise aujourd'hui a été proposée en 1953 par Watson et Crick [92] et a ouvert la voie à l'étude théorique de la dynamique de cette molécule et, depuis 50 ans, l'ADN ne cesse d'être le sujet d'études de physique théorique

et expérimentale très poussées. Ce polymère d'origine naturelle est formé de deux brins enroulés sous forme hélicoïdale et chacun de ces brins est constitué d'un enchaînement de monomères (nucléotides) de quatre types différents: adénine (A), thymine (T), cytosine (C) et guanine (G). La complémentarité de ces bases, deux à deux ($A \leftrightarrow T$ et $C \leftrightarrow G$) permet l'association des deux brins d'ADN par une liaison chimique lâche: un pont d'atomes d'hydrogène. L'étude de l'ouverture ou de la dénaturation de la double-hélice d'ADN commence en 1952 (un an avant la résolution de la structure de l'ADN par Watson et Crick) avec les travaux de Renée Thomas [85]. L'ADN est une molécule très active aux températures biologiques, de plus des expériences effectuées montrent que lorsque la molécule d'ADN est posée dans une solution et chauffée jusqu'à une température caractéristique, les liaisons hydrogène se cassent, le double brin se sépare en deux et la molécule se dissocie: on parle alors de la dénaturation thermique de l'ADN.

1.1.2 Le modèle de Poland-Scheraga

Le modèle le plus connu pour étudier le phénomène de la dénaturation de la molécule d'ADN est celui de Poland-Scheraga (PS) introduit en 1966 [76, 77], l'une des plus célèbres applications de la physique statistique à la biologie. Basant sur les lois fondamentales de la thermodynamique utilisées en physique statistique, le modèle PS décrit la dénaturation des brins d'ADN en fonction de la température en considérant un système dans lequel une paire de bases ne peut prendre que deux états de configuration: un état fermé où les liaisons hydrogène sont intactes, et un état ouvert où ces liaisons sont cassées (boucles), en ne prenant pas en compte la forme en double hélice. Plus précisément, ce modèle est utilisé pour décrire la dénaturation d'un ADN natif où les deux brins de la chaîne sont égaux et les liaisons sont formées entre les bases de même indice. Considérons une chaîne d'ADN homogène de longueur N, une configuration du système est une collection des paires de base d'indices $(i_1, i_2, ..., i_n) \in \mathbb{N}^n$ avec $n \in \{1, 2, ...\}$, $i_1 = 1$ et $i_k < i_{k'}$ pour $1 \leq k < k' \leq n$. On attribue à chaque base liée une énergie $E_b > 0$ (énergie de liaison hydrogène) et à chaque boucle une pénalité énergétique $E_l > 0$. Une base non liée à son complémentaire dans l'autre brin se trouve alors soit

- Dans une boucle $L_k := \{i, i > i_k, i < i_{k+1}\}$ pour $k \in \{1, ..., n-1\}$. La longueur de cette boucle est alors $l_k = 2(i_{k+1} - i_k - 1)$ et de facteur d'entropie $B(l_k)$ avec

$$B(l) := s^l l^{-c}, \tag{1.1}$$

  pour $s \geq 1$ et $c \geq 1$.

- Dans la partie libre de longueur $N - i_n$ (la longueur de cette partie est la même dans les deux brins) dont le terme d'entropie associé est

$$A(l) := s^l (l+1)^{-\overline{c}}, \tag{1.2}$$

avec $\overline{c} \in \mathbb{R}$.

La fonction de partition est la somme sur toutes les configurations possibles du système

$$Z_N := \sum_{i=0}^{N-1} A(i)^2 W_{N-i}\,, \tag{1.3}$$

et W_m vérifie la relation récursive suivante

$$W_{m+1} = \exp(\beta E_b) W_m + \exp(\beta(E_b - E_l)) \sum_{0<i<m} B(2i) W_{m-i}\,, \tag{1.4}$$

où $W_1 = 1$ et $\beta = 1/(k_B T)$ (T est la température et k_B est la constante de Boltzmann).

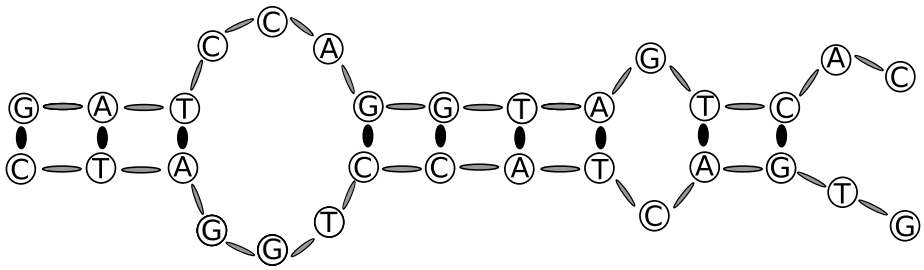

Figure 1.1. Deux brins égaux et libres aux extrémités d'une molécule d'ADN non enroulée: il s'agit des liaisons A-T ou C-G entre les bases de même indice.

L'étude du comportement critique de l'ADN homogène est très enrichissante dans la mesure où on connait le point critique et l'énergie libre sans être gêné par les complications causées par la présence d'hétérogénéité dans le système. Cependant, l'ADN se trouve naturellement sous forme hétérogène (Figure 1.1), jouant un rôle important dans les processus génétiques. Le modèle PS a été utilisé avec succès dans les études théoriques concernant les phénomènes critiques liés à la transition de dénaturation [30] et les effets de l'hétérogénéité sur la transition [45]. Notons aussi que le modèle PS constitue la base du logiciel de calcul des courbes de dénaturation thermique de l'ADN (MELTSIM [18]).

Bref, le modèle de Poland-Scheraga étudie la mécanique statistique de la liaison de deux brins complémentaires d'ADN de longueur égale à condition que seulement les bases de même indice peuvent se lier.

1.1.3 Le modèle de Poland-Scheraga généralisé

Lors de la réplication de l'ADN, une mutation peut se produire dans le brin naissant aboutissant à un désalignement de la chaîne [71]. Ce réarrangement joue un rôle important dans la variation génétique des bactéries. Pour étudier les propriétés physiques des chaînes mutées ou des polynucléotides synthétiques, une généralisation du modèle de

Poland-Scheraga (gPS) a été introduite depuis 50 ans [70, 82] en prenant en considération les deux conditions suivantes (Figure 1.2):

1. Chaque nucléotide d'un brin peut former une liaison avec n'importe quel nucléotide de l'autre brin (non complémentarité des brins). La formation des liaisons entre les nucléotides du même brin, qui constitue la base de la structure secondaire de l'ARN (Figure 1.3), est négligée dans ce modèle.

2. Les deux brins peuvent être de longueur différente.

Une configuration d'une chaîne d'ADN homogène possédant deux brins de longueur N et M est modélisée par une suite de paires de base dans $\mathbb{N}^2$: $((i_1, j_1), (i_2, j_2), ..., (i_n, j_n))$ avec $i_1 = j_1 = 1$ et $i_k < i_{k'}$, $j_k < j_{k'}$ pour $1 \leq k < k' \leq n$ et la base non liée se trouve soit

- Dans une boucle $L_k := ((i_k, i_{k+1}) \cup (j_k, j_{k+1}))$ de longueur $l_k := (i_{k+1} - i_k) + (j_{k+1} - j_k) - 2$ avec un terme d'entropie défini dans (1.1).

- Dans la partie libre de l'un de deux brins de longueur $N - i_n$ ou $M - j_n$ avec un terme d'entropie défini dans (1.2).

La fonction de partition est alors

$$Z_N^M := \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} A(i)A(j)W_{N-i}^{M-j}, \tag{1.5}$$

et W_m^r vérifie la relation récursive suivante

$$W_{m+1}^{r+1} = \exp(\beta E_b) W_m^r + \exp(\beta(E_b - E_l)) \sum_{\substack{i,i':i+i'>0 \\ i<m, i'<r}} B(i+i') W_{m-i}^{r-i'}, \tag{1.6}$$

avec $W_1^1 = 1$, $W_1^i = W_i^1 = 0$ pour tout $i > 1$.

Ce modèle possède encore une transition de dénaturation qui a été étudiée physiquement [46, 47] et peut avoir d'autres transitions dans la phase localisée [72] ce qui le rend plus riche que le modèle PS standard. Dans [72], les auteurs ont étudié la nouvelle transition (ou la phase intermédiaire) et ont trouvé une analogie avec la condensation de Bose-Einstein. Dans cette thèse, nous étudions d'abord ces transitions dans un état homogène où toutes les bases possèdent la même énergie.

Le passage au cas désordonné pose une autre question importante sur l'influence du désordre sur les différentes transitions du système. Le désordre peut être introduit comme une séquence i.i.d. sur les deux brins, une séquence périodique, séquence à deux valeurs.... Dans la molécule d'ADN, le désalignement de deux chaînes contribue à l'apparition de deux types de liaisons, les *matches* où les liaisons sont A-T et C-G (entre les bases complémentaires) et les *mismatches* qui désignent les autres liaisons produites de la non-complementarité de

deux brins (comme A-C, G-T,...). La plupart de ces modèles ont été étudiés physiquement [46, 47, 83], nous traitons dans cette thèse le cas d'un désordre introduit par une séquence i.i.d. à deux indices, où le premier (deuxième) indice correspond à l'indice du nucléotide dans le premier (deuxième) brin. Nous expliquons dans la Section 1.2.2 notre choix.

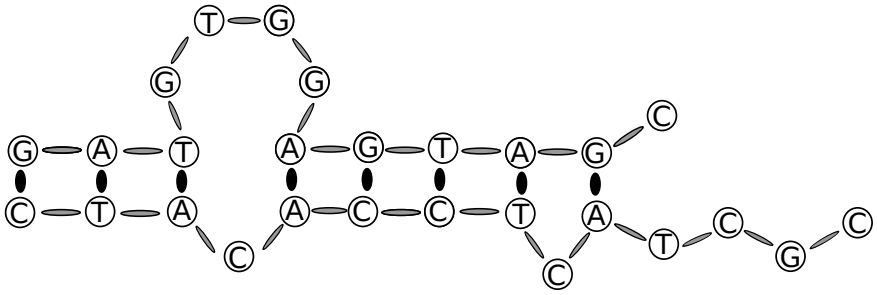

Figure 1.2. Une configuration possible de deux brins d'une molécule d'ADN de longueur différente et libres aux extrémités où les liaisons peuvent se former entre les bases d'indice différent. Les liaisons A-C, A-G, T-C ou T-G correspondent aux *mismatches*.

1.1.4 Modèles liés

On présente maintenant brièvement quelques exemples de phénomènes physiques pouvant être liés d'une manière directe ou indirecte au modèle généralisé de Poland-Scheraga:

Structure secondaire de l'ARN

L'acide ribonucléique (ARN) est une molécule très proche chimiquement de l'ADN mais elle se diffère de l'ADN de quatres manières principales: l'ARN est généralement formée d'un seul brin, plus courte de l'ADN et contient un autre type de sucre dans les nucléotides (ribose mais l'ADN contient le deoxyribose) tandis que la base Tymine trouvée dans l'ADN est remplacée par l'Uracil dans l'ARN.

La structure de l'ARN peut être décrite en plusieurs niveaux (Figure 1.3), dont le premier est tout simplement une séquence de nucléotides et le deuxième est une liste de bases liées permettant d'obtenir des boucles [80]. Un lien entre la structure secondaire de l'ARN et le modèle de gPS a été étudié dans [41]. Trouver une modélisation probabiliste de l'ARN est une question importante pour étudier son comportement et bien comprendre l'influence de la structure et du désordre sur la transition de phase [24].

Le surenroulement de l'ADN circulaire

Dans le modèle de PS (ou même dans gPS), la forme hélicoïdale de la chaîne d'ADN n'avait aucun rôle, toutefois il existe des cas où l'hélicité ne peut pas être ignorée comme le

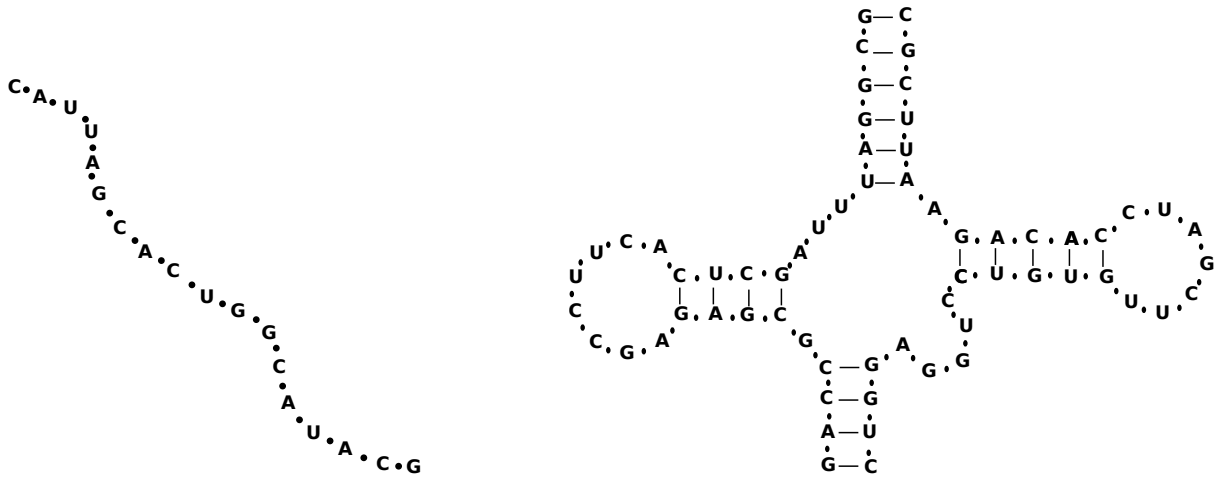

FIGURE 1.3. La structure primaire (à gauche) et la structure secondaire de l'ARN.

modèle d'ADN circulaire (un exemple d'une molécule d'ADN circulaire est le plasmide qui se trouve dans les bactéries). La formation d'une boucle dans un ADN circulaire provoque l'augmentation du nombre d'enlacement (*Linking number* L_K) et elle se manifeste par deux mécanismes: l'augmentation de la torsion (*Twist*) ou celui du vrillage (*writhe*). Le théorème de White-Fuller a été utilisé dans la description de l'ADN surenroulé [84, 86, 93] où le nombre de l'enlacement est conservé si les extrémités de l'ADN sont fixées: $L_k = Tw + Wr$.

La généralisation du modèle de PS au cas d'un ADN circulaire en prenant en considération seulement la formation des vrillages a été abordée par les physiciens dans [4, 61, 95]. Une configuration d'un double brin d'ADN circulaire est formée de trois parties différentes (Figure 1.4):

- Une partie de bases liées I^b dans laquelle les vrillages ne peuvent pas se former. Comme dans le modèle de PS, on suppose que l'énergie d'une base liée est $E_b > 0$.
- Une boucle I^l qui possède les mêmes propriétés de la boucle dans le modèle standard de PS.
- Un vrillage ou *supercoil* I^s dans lequel un brin serpente autour de l'autre. Chaque base liée dans un vrillage est énergétiquement favorisée avec une énergie E_s tel que $0 < E_s < E_b$.

Dans [4], les auteurs ont étudié la thermodynamique de ce modèle et ont montré qu'à haute température, plus précisément, à une température critique, une seule boucle macroscopique se forme dans le système, montrant l'existence d'une transition de phase au moins du second ordre. Ce phénomène de condensation est analogue à la condensation de Bose Einstein. Ce résultat a été généralisé dans [5] en prenant en considération l'augmentation de la torsion.

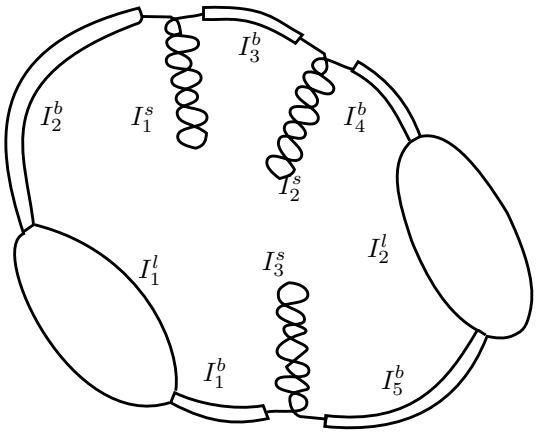

Figure 1.4. Une configuration de l'ADN circulaire composée d'une alternance de segments liés, boucles et supercoils.

1.2 Modélisation mathématique

1.2.1 Le modèle d'accrochage homogène

Nous présentons le modèle d'accrochage homogène en faisant le lien avec le modèle physique de Poland-Scheraga et sa généralisation d'une part, et avec le modèle basé sur le processus de renouvellement d'autre part. Nous gardons durant le reste de cette thèse les notations utilisées dans les parties de l'accrochage bidimensionnel de ce chapitre.

Accrochage unidimensionnel

Un lien entre le modèle PS standard et l'accrochage unidimensionnel est trouvé en observant que les deux brins d'ADN sont modélisés par deux marches aléatoires dirigées indépendantes à condition que chaque marche ne traverse pas l'autre et recevant une récompense à chaque fois que les deux marches se rencontrent. On peut aussi considérer le cas d'une seule marche aléatoire (différence de deux marches en éliminant la contrainte de non franchissement). Ce modèle est inclus dans un autre plus général en introduisant les outils du processus de renouvellement: on parle du modèle d'accrochage sur une ligne de défauts que nous présentons brièvement. Rappelons que le modèle d'accrochage homogène unidimensionnel est exactement résoluble, au sens où nous connaissons exactement le point critique ainsi que l'ordre de la transition de phase. De plus, il est important de donner un aperçu des résultats dans ce modèle afin de le comparer à la nouvelle étude liée à la généralisation du modèle de PS, sujet de cette thèse.

Pour introduire le modèle, nous considérons une marche aléatoire simple $\{S_n\}_{n=1,2,\ldots}$, i.e. $S_n = \sum_{i=1}^n X_i$ tel que $\{X_i\}_{i=1,2,\ldots}$ est une suite de variables aléatoires i.i.d. à valeurs dans $\{-1,+1\}$ dont la loi est notée $\mathbf{P}$, satisfaisant

$$\mathbf{P}(X_1 = 1) = \mathbf{P}(X_1 = -1) = \frac{1}{2}. \tag{1.7}$$

Soit $h \in \mathbb{R}$ l'énergie des contacts qui peut être un bonus ou malus selon son signe. La mesure $\mathbf{P}$ est alors modifiée en une mesure de polymère $\mathbf{P}_{N,h}$ d'une marche de longueur $N \in 2\mathbb{N}$ grâce à la densité

$$\frac{\mathrm{d}\mathbf{P}^c_{N,h}}{\mathrm{d}\mathbf{P}} := \frac{1}{Z^c_{N,h}} \exp\left(h \sum_{i=1}^N \mathbf{1}_{\{S_i=0\}}\right) \mathbf{1}_{S_N=0}. \tag{1.8}$$

La quantité $Z^c_{N,h}$ est appelée fonction de partition

$$Z^c_{N,h} := \mathbf{E}\left[\exp\left(h \sum_{i=1}^N \mathbf{1}_{\{S_i=0\}}\right) \mathbf{1}_{S_N=0}\right], \tag{1.9}$$

où l'exposant c se réfère à la contrainte $\{S_N = 0\}$. Cette contrainte impose que $N \in 2\mathbb{N}$ et elle est utilisée pour des raisons techniques.

Nous généralisons le modèle en présentant une formalisation basée sur la modélisation par un processus de renouvellement: soit $\tau = \{\tau_i\}_{i=0,1,\ldots}$ un processus de renouvellement discret de loi $\mathbf{P}$, défini par $\tau_0 = 0$ et pour tout entier $n \geq 1$

$$K(n) := \mathbf{P}(\tau_1 = n), \tag{1.10}$$

$$K(\infty) := \mathbf{P}(\tau_1 = \infty) = 1 - \sum_{n\in\mathbb{N}} K(n). \tag{1.11}$$

Les sauts ou les inter-arrivées $\{\tau_i - \tau_{i-1}\}_{i=1,2,\ldots}$ sont des variables aléatoires *i.i.d.* à valeurs dans $\mathbb{N}^*$. On dit que le renouvellement est *transient* si $K(\infty) > 0$ et alors $K(\cdot)$ est une sous-probabilité et il existe n_0 tel que pour tout $n \geq n_0$, $\tau_n = +\infty$ $\mathbf{P} - p.s.$. Par contre, le processus est dit *récurrent* si $K(\infty) = 0$. Si de plus, $\sum_n nK(n) < +\infty$, on parle d'un processus *récurrent positif* et dans le cas contraire ($\sum_n nK(n) = +\infty$), d'un processus *récurrent nul*. En analogie avec la formalisation de la marche aléatoire simple définie au début de cette partie, on définit le processus des instants de retour en 0 par $\tau_0 = 0$ et pour tout $k \in \mathbb{N}$

$$\tau_k = \inf\{i > \tau_{k-1} : S_i = 0\}. \tag{1.12}$$

De plus, par la formule de Stirling, on voit que

$$K(2n) \overset{n\to\infty}{\sim} \sqrt{\frac{1}{4\pi}} \frac{1}{n^{3/2}}. \tag{1.13}$$

Revenons au cadre plus général (processus de renouvellement), nous supposons que la loi des sauts a la forme

$$K(n) = \frac{L(n)}{n^{1+\alpha}}, \quad n \in \mathbb{N} \tag{1.14}$$

où $\alpha \geq 0$ et $L(\cdot)$ est une fonction à variation lente. Rappelons qu'une fonction $L : (0, +\infty) \mapsto (0, +\infty)$ est à variation lente si pour tous $\kappa > 0$ on a que

$$\lim_{x \to \infty} \frac{L(\kappa x)}{L(x)} = 1 . \tag{1.15}$$

L'exemple le plus simple de cette fonction est la constante. Pour plus de définitions et des exemples sur ce type de fonctions, le lecteur pourra se référer à [14]. Nous admettons que le processus est récurrent, i.e. $\sum_{n \geq 1} K(n) = 1$. Sinon on définit $\hat{K}(n) = K(n)/(1 - K(\infty))$.

Introduisons maintenant la notation

$$\delta_n = \mathbf{1}_{n \in \tau} = \sum_{i \geq 0} \mathbf{1}_{\tau_i = n} , \tag{1.16}$$

et définissons la mesure de Gibbs $\mathbf{P}_{N,h}$ sur τ de densité par rapport à la loi $\mathbf{P}$

$$\frac{\mathrm{d}\mathbf{P}^c_{N,h}}{\mathrm{d}\mathbf{P}}(\tau) := \frac{1}{Z^c_{N,h}} \exp\left(h \sum_{n=1}^{N} \delta_n \right) \delta_N , \tag{1.17}$$

et la fonction de partition par

$$Z^c_{N,h} := \mathbf{E}\left[\exp\left(h \sum_{n=1}^{N} \delta_n \right) \delta_N \right] . \tag{1.18}$$

Plus concrètement, remarquons que $Z^c_{N,h}$ est la fonction de Green d'un processus de renouvellement bien choisi, de sorte que l'on a aussi

$$Z^c_{N,h} = \sum_{n=1}^{N} \sum_{\substack{\underline{l} \in \mathbb{N}^n : \\ |\underline{l}| = N}} \prod_{i=1}^{n} \exp(h) K(l_i) . \tag{1.19}$$

Nous pourrons maintenant parler du lien entre le modèle PS et le modèle d'accrochage homogène. En se référant à la Section 1.1.2 et à la Figure 1.5 ci-dessous, nous décrivons les observations suivantes

- Une boucle contenant l nucléotides sur chaque brin apporte une contribution $\exp(\beta(E_b - E_l))B(2l)$ dans le modèle biophysique. L'interarrivée associée à ce terme dans le modèle d'accrochage que nous considérons est $\exp(h)K(l+1)$. La restriction au facteur $K(l+1)$ (au lieu de $K(2l+2)$) suit simplement du fait que les deux brins sont complémentaires et qu'il suffit de modéliser un seul brin. Notons aussi que la contribution d'une interarrivée minimale $\exp(h)K(1)$ (un saut de longueur 1) correspond à $\exp(\beta E_b)$ (la contribution d'une seule base liée).

- On peut se ramener au modèle mathématique en prenant d'abord $s = 1$: Z_N dans (1.3) coïncide avec $s^{2N}\widetilde{Z}_N$ où $\widetilde{Z}_N$ est Z_N avec $s = 1$ et $E_b - E_l$ est diminué de $\frac{2}{\beta}\log s$. Il est facile de voir le lien entre $K(\cdot)$ et $B(\cdot)$ pour $c = 1 + \alpha$. Pour le modèle avec extrémités libres, le choix est laissé au lecteur à condition qu'il soit compatible avec $A(\cdot)$, par exemple dans [48], la probabilité que la partie libre soit de longueur n est égal $\sum_{j>n} K(j)$ pour $n = 0, 1, \ldots$.

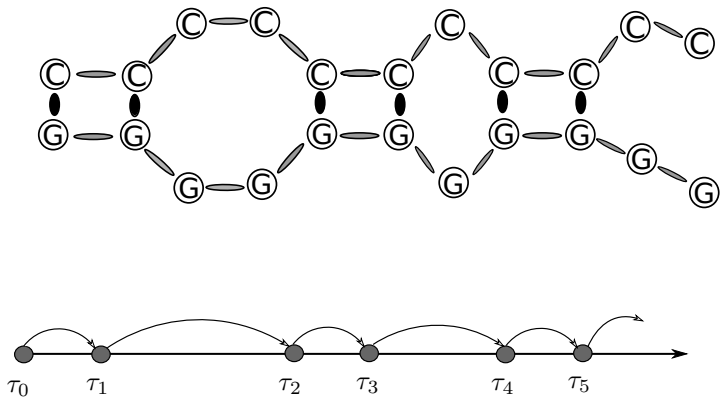

FIGURE 1.5. Une chaîne d'ADN homogène (dans laquelle toutes les liaisons sont identiques de type C-G) est liée à un processus de renouvellement unidimensionnel. Les points $\tau_0, \tau_1, \ldots$ représentent les bases liées.

Afin d'étudier le système, nous introduisons $F(h)$ comme l'unique solution de l'équation d'inconnue $x > 0$:

$$\sum_{n\geq 1} K(n) \exp(-xn) = \exp(-h), \tag{1.20}$$

si cette solution existe (autrement dit si $h > 0$) et $F(h) = 0$ sinon.

On déduit alors que le point critique $h_c := \sup\{h : F(h) = 0\}$ est égal à 0 dans le cas où $K(\infty) = 0$, et dans le cas général ($K(\infty) > 0$) le point critique vaut $-\log(1-K(\infty)) > 0$. De plus, comme la fonction $x \mapsto \sum_{n\geq 1} K(n)\exp(-xn)$ est analytique, le théorème des fonctions implicites (voir [64, théorème 8.6]) assure que $F(\cdot)$ est analytique en tout point $h \neq 0$.

Introduisons la distribution de probabilité $\widetilde{K}_h(n) = K(n)\exp(h - F(h)n)$ dans le cas où $h > h_c$, on obtient alors

$$Z^c_{N,h} = \exp(F(h)N) \sum_{n=1}^{N} \sum_{\substack{\underline{l}\in\mathbb{N}^n:\\ |\underline{l}|=N}} \prod_{i=1}^{n} \widetilde{K}_h(l_i) = \exp(F(h)N)\mathbf{P}(N \in \widetilde{\tau}_h), \tag{1.21}$$

où $\widetilde{\tau}_h$ est le processus de renouvellement de distribution $\widetilde{K}_h(\cdot)$ et de moyenne finie (récurrent positif). La fonction $F(\cdot)$ est alors l'énergie libre du système.

Proposition 1.1. *Pour tout* $h \in \mathbb{R}$

$$F(h) = \lim_{N\to\infty} \frac{1}{N} \log Z^c_{N,h}. \tag{1.22}$$

Pour $h > 0$, la preuve de (1.22) se concentre sur l'utilisation du théorème de renouvellement qui indique que $\mathbf{P}(N \in \widetilde{\tau}_h) \to 1/\left(\sum_{n\geq 1} n\widetilde{K}_h(n)\right)$ lorsque N tend vers l'infini. Donc

$$Z^c_{N,h} \sim \frac{1}{\sum_{n\geq 1} n\widetilde{K}_h(n)} \exp(F(h)N). \tag{1.23}$$

Pour le cas $h \leq 0$, on voit directement que $Z^c_{N,h}$ est une fonction de renouvellement: $Z^c_{N,h} = \mathbf{P}(N \in \widetilde{\tau}_h)$, avec $\widetilde{K}_h(\cdot)$ est une sous-probabilité pour $h < 0$. On peut dire alors que

$$\exp(h)K(N) \leq Z^c_{N,h} \leq 1\,. \tag{1.24}$$

L'équation (1.22) nous indique aussi que F est convexe et croissante, puisque la fonction $h \mapsto \frac{1}{N}\log Z^c_{N,h}$ est elle-même convexe et croissante pour N fixé. On définit maintenant la densité de contact par

$$\frac{\partial}{\partial h}\frac{1}{N}\log Z^c_{N,h} = \frac{1}{N}\mathbf{E}^c_{N,h}\left[\sum_{n=1}^{N}\delta_n\right]. \tag{1.25}$$

En passant à la limite et par convexité, on obtient

$$F'(h) = \lim_{N\to\infty}\frac{1}{N}\mathbf{E}^c_{N,h}\left[\sum_{n=1}^{N}\delta_n\right], \tag{1.26}$$

lorsque le membre de gauche existe. La dérivée de l'énergie libre a une interprétation physique importante (ainsi que sa dérivée seconde) et l'étude de sa régularité nous aide à comprendre le comportement du système.

L'équation (1.20), où l'énergie libre est exprimée implicitement en fonction de la transformée de Laplace de τ_1, permet de déduire le comportement critique de $F(h)$ près du point critique et d'avoir l'exposant critique et par suite l'ordre de la transition. Rappelons que le modèle présente une transition de phase d'ordre k si $F(h)$ est C^{k-1} mais pas C^k au point critique.

Théorème 1.2. *[[48], théorème 2.1] On a*

1. *Si* $m_K := \sum_{n\geq 1} nK(n) < +\infty$ *alors*

$$F(h) \overset{h\searrow 0}{\sim} \frac{h}{m_K}\,. \tag{1.27}$$

2. *Si* $0 < \alpha < 1$ *ou* $\alpha = 1$ *et* $m_K = +\infty$, *alors pour* $h > 0$

$$F(h) \overset{h\searrow 0}{\sim} h^{1/\alpha}\hat{L}(1/h)\,, \tag{1.28}$$

où $\hat{L}(\cdot)$ *est une fonction à variation lente.*

3. *Si* $\alpha = 0$, *alors lorsque* $h \searrow 0$, *on a* $F(h) = o(h^\kappa)$ *pour tout* $\kappa > 0$.

L'équation (1.20) permet aussi de voir que l'énergie libre est analytique en tout point $h \neq h_c$ par le théorème des fonctions implicites. Grâce au théorème 1.2, on a que si $\alpha > 1$ alors $F(h)$ n'est pas C^1 et la transition est de premier ordre, tandis qu'elle est d'ordre k, $k = 2, 3, \ldots$ si $\alpha \in [1/k, 1/(k-1))$.

Accrochage bidimensionnel

Le modèle d'accrochage bidimensionnel représente le sujet central de cette thèse. L'intérêt de ce modèle se trouve déjà dans sa version homogène et il est introduit pour étudier la généralisation du modèle physique de PS (présentée dans la Section 1.1.3) dans un cadre mathématique. La première question que l'on se pose est de savoir avec quel processus les deux brins non complémentaires de l'ADN peuvent être modélisés. Le processus de renouvellement bidimensionnel était l'outil convenable pour faire le lien physique-mathématique de ce modèle. Comme le modèle PS standard, le processus de renouvellement joue un rôle important dans la compréhension complète de la transition de délocalisation, ainsi que celles des autres transitions possibles dans la phase localisée qui rendent le modèle (PS généralisé) plus important. Nous verrons cependant que le modèle d'accrochage bidimensionnel conserve le caractère résoluble trouvé dans le cas unidimensionnel.

Un processus $\tau = \{\tau_n\}_{n=0,1,\dots} = (\tau^{(1)}, \tau^{(2)})$ de loi $\mathbf{P}$, est appelé processus de renouvellement bidimensionnel discret issu de l'origine si $\tau_0 = (0,0)$ et pour tout $n \in \mathbb{N}$, τ_n est une somme de n variables aléatoires i.i.d.. On notera pour $n, m \in \mathbb{N}$, $\mathbf{P}(\tau_1 = (n,m)) := \mathrm{K}(n,m) = K(n+m)$ où

$$K(n) = \frac{L(n)}{n^{2+\alpha}}, \tag{1.29}$$

avec $L(\cdot)$ est une fonction à variation lente et $\alpha \geq 0$ et on appellera $\mathrm{K}(\cdot,\cdot)$ la loi des inter-arrivées. Notons qu'avec cette définition, $K(n) > 0$ pour tout entier $n \geq 1$. Pour simplifier les calculs, on peut supposer que le renouvellement τ est récurrent, i.e. $\sum_{n,m\geq 1} K(n+m) = 1$.

On introduit la famille de mesures de Gibbs $\{\mathbf{P}^c_{N,M,h}\}_{N,M=1,2,\dots}$ sur τ pour $h \in \mathbb{R}$, définies par leurs densités par rapport à la loi initiale $\mathbf{P}$:

$$\frac{\mathrm{d}\mathbf{P}^c_{N,M,h}}{\mathrm{d}\mathbf{P}}(\tau) := \frac{1}{Z^c_{N,M,h}} \exp\left(h \sum_{n=1}^{N}\sum_{m=1}^{M} \delta_{n,m}\right) \delta_{N,M}, \tag{1.30}$$

où $\delta_{n,m} = 1_{(n,m)\in\tau}$. La constante de renormalisation (fonction de partition) est

$$Z^c_{N,M,h} = \mathbf{E}\left[\exp\left(h \sum_{n=1}^{N}\sum_{m=1}^{M} \delta_{n,m}\right) \delta_{N,M}\right]. \tag{1.31}$$

L'un des atouts du modèle homogène est le fait que la fonction de partition possède une représentation simple. En comparaison avec (1.19), la fonction de partition dans le modèle généralisé peut être écrite comme la somme des contributions de toutes les trajectoires accrochées en (N,M) sous la forme

$$Z^c_{N,M,h} = \sum_{n=1}^{N\wedge M} \sum_{\substack{\underline{l}\in\mathbb{N}^n:\\ |\underline{l}|=N}} \sum_{\substack{\underline{t}\in\mathbb{N}^n:\\ |\underline{t}|=M}} \prod_{i=1}^{n} \exp(h) K(l_i + t_i). \tag{1.32}$$

En ce qui concerne le modèle avec extrémités libres, la fonction de partition libre est donnée par

$$Z^f_{N,M,h} := \sum_{i=0}^{N}\sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h}, \tag{1.33}$$

où $K_f : \{0\} \cup \mathbb{N} \mapsto (0, \infty)$ est définie comme $K_f(n) := \overline{L}(n)/n^{\overline{\alpha}}$ pour tout $n \geq 1$, $\overline{\alpha} \in \mathbb{R}$, $\overline{L}(\cdot)$ une fonction à variation lente et $K_f(0) = 1$.

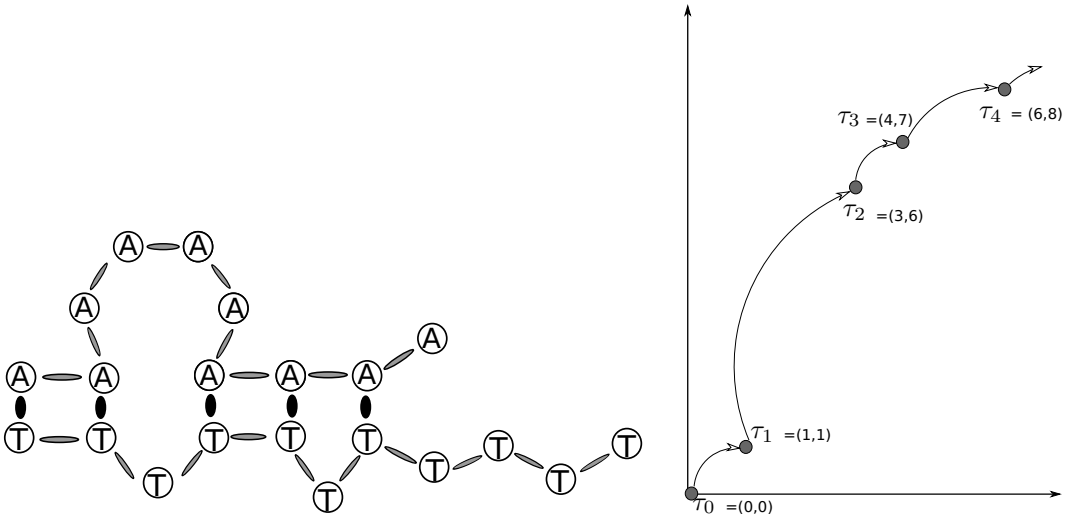

Figure 1.6. Une représentation montrant le lien entre une chaîne d'ADN homogène dont l'apparition des boucles non symétriques est autorisée et un processus de renouvellement bidimensionnel. Les coordonnées des points $\tau_0, \tau_1, \ldots$ correspondent à la position des bases liées dans la chaîne.

Pour faire un lien avec le modèle gPS, nous invitons le lecteur à se referer à la Figure 1.6, à la Section 1.1.3 et au lien entre le modèle d'accrochage homogène et le modèle PS présenté juste avant (1.20). La différence la plus importante avec le cas unidimensionnel est que la contribution d'une boucle possédant l nucléotides dans le premier brin et l' dans le deuxième $\exp(\beta(E_b - E_l))B(l + l')$ est liée à la contribution $\exp(h)K(l + l' - 2)$. La non-complémentarité des brins est à la base de cette différence. Ainsi le lien entre $K(\cdot)$ et $B(\cdot)$, $K_f(\cdot)$ et $A(\cdot)$ est facile à trouver (avec $c = 2 + \alpha$ et $\overline{c} = \overline{\alpha}$).

Le premier résultat que l'on peut donner est alors:

Théorème 1.3. *Les suites* $\{\frac{1}{N} \log Z^f_{N,M,h}\}_{N,M=1,2,\ldots}$ *et* $\{\frac{1}{N} \log Z^c_{N,M,h}\}_{N,M=1,2,\ldots}$ *convergent, lorsque* $N, M \to \infty$ *et* $M/N \to \gamma$, *vers la même limite notée* $\mathrm{F}_\gamma(h)$ *(énergie libre):*

$$\mathrm{F}_\gamma(h) := \lim_{\substack{N,M\to\infty \\ \frac{M}{N}\to\gamma}} \frac{1}{N} \log Z^f_{N,M,h} = \lim_{\substack{N,M\to\infty \\ \frac{M}{N}\to\gamma}} \frac{1}{N} \log Z^c_{N,M,h}\,. \tag{1.34}$$

Il est facile de voir que $\mathrm{F}_\gamma(h) \geq 0$ pour tout h en utilisant l'inégalité

$$Z^f_{N,M,h} \geq K_f(0)^2 Z^c_{N,M,h} = Z^c_{N,M,h} \geq \exp(h)K(N + M)\,. \tag{1.35}$$

Observons aussi que l'inégalité $\mathrm{F}_\gamma(h) \geq \mathrm{F}_1(h)$ pour tout $\gamma \geq 1$ suit simplement du fait que $Z^f_{N,M,h} \geq Z^c_{N,N,h} K_f(0) K_f(M - N)$.

Pour des raisons techniques, nous utilisons $Z^c_{N,M,h}$ pour étudier le modèle. La fraction de contact moyenne est liée à l'énergie libre dans un modèle dont les deux brins sont de taille N et M, par la relation

$$\frac{\partial}{\partial h}\frac{1}{N}\log Z^c_{N,M,h} = \frac{1}{N}\mathbf{E}^c_{N,M,h}\left[\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\right]. \tag{1.36}$$

Grâce à la convexité de la fonction $h \mapsto \frac{1}{N}\log Z^c_{N,M,h}$ et si $\mathrm{F}'_\gamma(h)$ existe, l'égalité (1.36) passe à la limite quand $N \to \infty$ et on obtient

$$\mathrm{F}'_\gamma(h) = \lim_{\substack{N,M\to\infty \\ \frac{M}{N}\to\gamma}} \frac{1}{N}\mathbf{E}^c_{N,M,h}\left[\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\right]. \tag{1.37}$$

Deux comportements possibles de la chaîne d'ADN sont donc séparés en deux phases (localisée et délocalisée) par un point critique h_c défini par

$$h_c := \sup\{h : \mathrm{F}_\gamma(h) = 0\}. \tag{1.38}$$

Nous donnons maintenant quelques points essentiels pour la résolution de ce modèle (une étude détaillée est présentée dans le chapitre 2). On définit une nouvelle distribution

$$\widetilde{K}_h(n,m) = \mathrm{K}(n,m)\exp\left(h - \mathrm{G}(h)(n+m)\right), \tag{1.39}$$

où l'introduction du facteur de normalisation $\mathrm{G} : \mathbb{R} \mapsto [0,\infty)$ vient du fait que $h > 0$ empêche $\exp(h)K(n+m)$ d'être une mesure de probabilité. Lorsque $h < 0$, la densité $\widetilde{K}_h(n,m) = \mathrm{K}(n,m)\exp(h)$ est une sous-probabilité qui peut être interprétée comme étant une probabilité sur $\mathbb{N}^2 \cup \{\infty\}$ où la probabilité d'une interarrivée infinie vaut $\widetilde{K}_h(\infty) := 1 - \sum_{n,m\geq 1}\widetilde{K}_h(n,m)$.

Contrairement à ce qui était le cas dans l'accrochage unidimensionnel pour l'énergie libre (voir (1.20)), la fonction $\mathrm{G}(\cdot)$ représente seulement une partie de l'énergie libre.

Proposition 1.4. *Si $h < 0$ alors $\mathrm{G}(h) = 0$. Si $h \geq 0$, la fonction $\mathrm{G}(h)$ est convexe et analytique sauf en 0 et définie comme l'unique solution $x \geq 0$ de*

$$\sum_{n,m}\mathrm{K}(n,m)\exp(h - x(n+m)) = 1. \tag{1.40}$$

Pour $h > 0$, le nouveau renouvellement $\widetilde{\tau}_h$ est récurrent, on a même

$$\mu_h := \mathbf{E}[(\widetilde{\tau}_h)_1] = \left(\sum_{n,m} n\widetilde{K}_h(n,m), \sum_{n,m} m\widetilde{K}_h(n,m)\right) < \infty. \tag{1.41}$$

La fonction de partition dans (1.32) possède alors une forme générale

$$Z^c_{N,M,h} = \exp\left((N+M)\mathrm{G}(h)\right)\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right). \tag{1.42}$$

Grâce à cette formule, nous avons besoin d'étudier la fonction de renouvellement bidimensionnelle $\mathbf{P}((N,M) \in \widetilde{\tau}_h)$. Le comportement asymptotique de cette fonction entre dans l'identification de la transition de localisation/délocalisation et constitue la base pour découvrir les autres transitions dans la phase localisée.

Avant de présenter un résultat de grande déviation pour la fonction de renouvellement, on introduit la seconde fonction de déviation $D(\cdot)$ par l'une de ses définitions (le lecteur pourra se référer à [20]).

Proposition 1.5 ([20], théorème 1)**.** *Pour tout $\theta \in \mathbb{R}^2$*

$$D_h(\theta) := \sup_{\lambda \in A_h} \langle \lambda, \theta \rangle = \sup_{\lambda \in \partial A_h} \langle \lambda, \theta \rangle , \tag{1.43}$$

avec $A_h = \{\lambda \in \mathbb{R}^2 : \mathbf{E}[e^{\langle \lambda, (\widetilde{\tau}_h)_1 \rangle}] \leq 1\}$ et ∂A_h est la frontière de A_h.

Nous nous concentrons sur le cas particulier que nous considérons, pour tout $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2$

$$\mathbf{E}[e^{\langle \lambda, (\widetilde{\tau}_h)_1 \rangle}] = \sum_{n,m} \exp(h) K(n+m) \exp\left(-(\mathrm{G} - \lambda_1)n - (\mathrm{G} - \lambda_2)m\right) := q_h(\lambda_1, \lambda_2) . \tag{1.44}$$

Alors l'ensemble A_h est écrit comme

$$A_h = \{\lambda \in \mathbb{R}^2 : q_h(\lambda_1, \lambda_2) \leq 1\} . \tag{1.45}$$

On mentionne une liste des observations sur l'ensemble A_h:

- Pour tout $\lambda = (\lambda_1, \lambda_2) \in A_h$, la restriction au cas $\lambda_1, \lambda_2 \leq \mathrm{G}$ provient du fait que $q_h(\lambda_1, \lambda_2) = \infty$ si $\lambda_1 \vee \lambda_2 > \mathrm{G}$.
- La convexité de A_h et les faits que $q_h(\lambda_1, \lambda_2) = q_h(\lambda_2, \lambda_1)$ et $q_h(0,0) = 1$ impliquent que $\lambda_2 \leq -\lambda_1$.
- On définit une fonction $\overline{\lambda}_1(\cdot)$ (dépendant de h) à valeurs négatives comme solution de l'équation $q_h(\overline{\lambda}_1, \mathrm{G}) = 1$. La solution de $q_h(\lambda) = 1$ est unique pour $\lambda_1 \in [\overline{\lambda}_1, \mathrm{G}]$.
- On définit la fonction $\widetilde{\lambda}_2 : [\overline{\lambda}_1, \mathrm{G}] \mapsto [\overline{\lambda}_1, \mathrm{G}]$ satisfaisant $\widetilde{\lambda}_2(\overline{\lambda}_1) = \mathrm{G}$, $\widetilde{\lambda}_2(\mathrm{G}) = \overline{\lambda}_1$ et $\widetilde{\lambda}_2(0) = 0$. On nomme par $\mathcal{W}_h$ son graphe. Cette fonction est concave, décroissante et linéaire et par conséquent on obtient que $\overline{\lambda}_1 < \mathrm{G}$.

On peut alors conclure que

$$\partial A_h = \{(\lambda_1, \mathrm{G}) : \lambda_1 < \overline{\lambda}_1\} \cup \{(\mathrm{G}, \lambda_2) : \lambda_2 < \overline{\lambda}_1\} \cup \mathcal{W}_h . \tag{1.46}$$

La seconde fonction de déviation $D(\cdot)$ définie plus haut est la fonction de grande déviation de la fonction de renouvellement et pour θ dans le premier quadrant, on a $D(\theta) < \infty$

(voir (1.43) et (1.46)). L'une des conséquences de l'analyse de $D(\cdot)$ et A_h (dans le cas où $D(\cdot)$ possède une valeur finie) est l'apparition de deux possibilités: le supremum dans (1.43) est atteint à l'intérieur de $\mathcal{W}_h$ (on dit que $\theta \in E_h$ dans ce cas) ou sur la frontière, i.e. en $\{(\overline{\lambda}_1, \mathrm{G}), (\mathrm{G}, \overline{\lambda}_1)\}$. Le résultat suivant suffit pour une étude au niveau de l'énergie libre

Théorème 1.6. *Pour tout θ, on a*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbf{P}\left(\lceil N\theta \rceil \in \widetilde{\tau}_h\right) = -D_h(\theta)\,. \tag{1.47}$$

Ce résultat est une conséquence du théorème de la limite locale (voir [20, théorème 5]) pour $\theta \in E_h$ et de la proposition 2.15 dans le chapitre 2 qui correspond à (1.47) pour $\theta \in E_h^{\mathsf{C}}$.

Notons que $D(1,1) = 0$ alors en utilisant le théorème 1.6 pour $M = N$, on obtient $\mathrm{F}_1(h) = 2\mathrm{G}(h)$ et ce cas est interprété comme le cas unidimensionnel où la proposition 1.4 et notamment (1.40) permet d'avoir le comportement critique de $\mathrm{F}_1(h)$ près du point critique.

Utilisons le théorème 1.3, théorème 1.6 et (1.42), on énonce maintenant le résultat suivant pour $M \sim \gamma N$

Proposition 1.7. *Pour tout $\gamma \geq 1$ et $h > 0$ on a*

$$\mathrm{F}_\gamma(h) = \begin{cases} 0 & \text{si } h \leq 0\,, \\ (1+\gamma)\mathrm{G}(h) - D_h(1,\gamma) & \text{si } h > 0\,. \end{cases} \tag{1.48}$$

De plus, on a

$$D_h(1,\gamma) = \max_{\lambda \in B_h} (\lambda_1 + \gamma\lambda_2)\,, \tag{1.49}$$

et B_h est l'ensemble défini par

$$B_h = \{\lambda : \overline{\lambda}_1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq \mathrm{G}, \sum_{n,m} \widetilde{\mathrm{K}}_h(n,m) e^{\lambda_1 n + \lambda_2 m} = 1\}\,. \tag{1.50}$$

Il est facile de voir que $\mathrm{F}_\gamma(h) = 0$ pour $h \leq 0$ par le fait que $Z^c_{N,M,h} \leq 1$ pour $h \leq 0$. Il est aussi clair que $h_c = 0$.

Nous sommes alors prêts à découvrir les différentes transitions du système et nous commençons par la transition fondamentale (délocalisation). La proposition 1.7 nous permet de conclure que l'étude du comportement de l'énergie libre près du point critique h_c et de l'ordre de la transition de phase consiste à étudier séparément le comportement de $\mathrm{G}(\cdot)$ grâce à (1.40) par les théorèmes taubériens, où $\mathrm{G}(\cdot)$ est exprimée implicitement en fonction de la transformée de Laplace τ_1 et celui de $D(\cdot)$ grâce au problème variationnel (1.49)-(1.50).

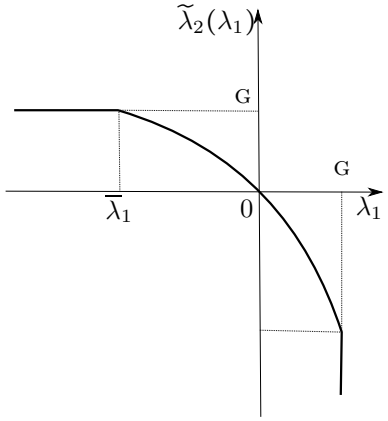

Figure 1.7. La courbe épaisse désigne ∂A_h pour h fixé. La partie qui nous intéresse est pour $\lambda_1 \in [\overline{\lambda}_1, 0]$ puisque A_h est symétrique et $q_h(\lambda_1, \lambda_2) < 1$ pour $\lambda_1, \lambda_2 < \overline{\lambda}_1$.

Théorème 1.8. *Pour tout $\alpha \geq 0$ et $\gamma \geq 1$, on a:*

- *Si $\sum_n n^2 K(n) < \infty$ (impliqué par $\alpha > 2$) alors*

$$\mathrm{F}_\gamma(h) \overset{h\searrow 0}{\sim} \mathrm{F}_1(h) \sim ch\,, \tag{1.51}$$

  avec $c^{-1} = \frac{1}{2}\sum_{n\geq 2} n(n-1)K(n)$.

- *Si $\sum_n n^2 K(n) = \infty$ (impliqué par $\alpha \in (0,1)$), alors il existe $c_{\alpha,\gamma} \geq 1$ tel que*

$$\mathrm{F}_\gamma(h) \sim c_{\alpha,\gamma} h^{1/\alpha} L_\alpha(h)\,, \tag{1.52}$$

  où $L_\alpha(\cdot)$ est une fonction à variation lente en 0.

- *Si $\alpha = 0$, alors $\mathrm{F}_\gamma(h) = O(h^{1/\epsilon})$ pour tout $\epsilon > 0$ lorsque $h \searrow 0$.*

Une conséquence de ce théorème est que la transition est de premier ordre pour $\alpha > 1$ et elle est d'ordre $k = 2, 3, \ldots$ si $\alpha \in [1/k, 1/(k-1)]$.

Comme mentionné plus haut, le modèle généralisé de PS possède des nouveautés qualitatives. La nouveauté la plus importante est l'apparition des transitions à l'intérieur du régime localisé. Nous mentionnons ici quelques résultats à ce niveau. On introduit pour $h > 0$

$$\gamma_c(h) := \frac{\sum_{n,m} mK(n+m)\exp\left(-n(\mathrm{G}(h) - \overline{\lambda}_1(h))\right)}{\sum_{n,m} nK(n+m)\exp\left(-n(\mathrm{G}(h) - \overline{\lambda}_1(h))\right)}\,. \tag{1.53}$$

La fonction $\gamma_c : (0,\infty) \longrightarrow (1,\infty)$ est analytique et l'introduction de cette fonction provient d'une interprétation du problème variationnel (1.49)-(1.50). On a le résultat suivant

Théorème 1.9. *$\mathrm{F}_\gamma(\cdot)$ est analytique sur $\{h : h > 0$ tel que $\gamma_c(h) - \gamma \neq 0\}$ et $\mathrm{F}_\gamma(\cdot)$ n'est pas analytique aux valeurs $h > 0$ en lesquels $\gamma_c(h) - \gamma$ change de signe.*

On notera par $h_{c,\gamma} > 0$ le point non analytique de $\mathrm{F}_\gamma(\cdot)$. On distingue maintenant deux régimes conduisant à l'apparition d'une (ou plusieurs) transition(s) dans la phase localisée:

- Le régime de Cramer: pour $\gamma < \gamma_c(h)$, le maximum dans (1.49) est atteint à un certain point $\big(\hat{\lambda}_1(h,\gamma), \hat{\lambda}_2(h,\gamma)\big)$ et grâce à (1.48), l'énergie libre est alors égale à

$$\hat{c}_\gamma(h) := \big(\mathrm{G}(h) - \hat{\lambda}_1(h,\gamma)\big) + \gamma\big(\mathrm{G}(h) - \hat{\lambda}_2(h,\gamma)\big)\,. \tag{1.54}$$

- L'extérieur du régime de Cramer: pour $\gamma \geq \gamma_c(h)$, dans ce cas le maximum est atteint au point $(\overline{\lambda}_1(h), \mathrm{G}(h))$ et l'énergie libre devient

$$\mathrm{N}(h) := \mathrm{G}(h) - \overline{\lambda}_1(h)\,. \tag{1.55}$$

Dans le théorème 1.9, on a présenté un résultat général sur l'analyticité de l'énergie libre dans la phase localisée. Mais que sait-on des points $\{h : \gamma_c(h) - \gamma = 0 \,\text{et}\, \gamma_c(h) - \gamma \,\text{ne change pas de signe}\}$? et quel est le type de la singularité lorsqu'elle existe?

La réponse à ces questions demande une étude cas par cas, notamment des points h_0 vérifiant $\gamma_c(h_0) = \gamma$ and $\gamma_c'(h_0) = 0$, i.e. des maximums, des minimums et des points selles.

Proposition 1.10. *Soit h_0 tel que $\gamma_c(h_0) = \gamma$. La fonction $\mathrm{F}'_\gamma(\cdot)$ est continue en h_0, alors la transition n'est pas de premier ordre. Si $\sum_m m^2 K(m) < \infty$, la transition en h_0 est de second ordre si et seulement si $\gamma_c'(h_0) \neq 0$. Si $\sum_m m^2 K(m) = \infty$, $\mathrm{F}''_\gamma(\cdot)$ est continue en h_0 montrant que la transition est d'ordre supérieure ou égale à trois.*

L'étude du comportement asymptotique de la fonction de partition dans le cas contraint ou dans le cas libre est réduit à l'étude de la fonction de renouvellement $\mathbf{P}((N,M) \in \widetilde{\tau}_h)$ grâce à (1.42). Le comportement asymptotique de la fonction de renouvellement en dimension $d \geq 2$ est très intéressant en soi mais l'absence d'une étude complète dans la littérature pour ce cas général (contrairement au cas unidimensionnel) nous incite à ne pas tenter d'être exhaustif. Avec ce qui est disponible, nous mentionnons quelques propriétés de trajectoires dans le cas libre et pour $M \sim \gamma N$.

- Dans la phase délocalisée ($h < 0$): le processus $\widetilde{\tau}_h$ est transient dans ce cas et la fonction de partition (dans le cas contraint) est égale à $\mathbf{P}((N,M) \in \widetilde{\tau}_h)$. Si $\overline{\alpha} < \alpha/2$, les deux brins sont libres à l'exception de $O(1)$ contacts proche de l'origine et si $\overline{\alpha} > \alpha/2$ les deux brins sont détachés après un nombre fini de contacts proche de l'origine et ils se rencontrent de nouveau à $O(1)$ distance de (N,M) avec deux brins libres de longueurs $O(1)$.

- Dans la phase localisée ($h > 0$): si γ est dans la région de Cramer, le processus de renouvellement converge vers un processus récurrent dont la loi des interarrivées est donnée par la fonction de $\mathbb{N}^2$ à $[0,1)$

$$(i,j) \mapsto K(i+j)\,\exp\left(-i(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)) - j\partial_\gamma \mathrm{F}_\gamma(h)\right)\,. \tag{1.56}$$

Il reste alors le cas où γ n'est pas dans la région de Cramer et le cas critique ($h = 0$). En se basant sur des résultats dans [6] sur l'étude du comportement asymptotique de la fonction de renouvellement dans les cas non traités, notamment le cas correspondant à la région hors Cramer, nous conjecturons de manière fiable ce qui suit:

- Dans le cas contraint, une grande boucle apparait et peut être n'importe où le long de la chaîne et la fonction de partition est de l'ordre $N\exp(N\mathrm{N}(h))K(M-\gamma_c(h)N)$ lorsque $N \to \infty$.
- Dans le cas avec extrémités libres, si $\overline{\alpha} > 1+\alpha$ alors le système se comporte comme dans le cas contraint (une boucle distribuée uniformément sur toute la longueur de la chaîne). Par contre si $\overline{\alpha} < 1+\alpha$, les extrémités libres deviennent macroscopiques et la fonction de partition est de l'ordre $\exp(N\mathrm{N}(h))K_f(M-\gamma_c(h)N)$.

Nous mentionnons aussi la conjecture suivante: pour $h > 0$ et $\gamma > \gamma_c(h)$, la loi de τ sous $\mathbf{P}^c_{N,M,h}$ et $\mathbf{P}^f_{N,M,h}$ converge vers un processus récurrent bidimensionnel dont le comportement est décrit par la fonction de $\mathbb{N}^2$ à $[0,1)$

$$(i,j) \longmapsto K(i+j)\exp(h - i\mathrm{N}(h))\,. \tag{1.57}$$

1.2.2 Le modèle d'accrochage désordonné

L'influence du désordre sur la transition de phase et les phénomènes critiques est une question centrale en mécanique statistique (voir par exemple [50, Chapitre 5] et [90]). Etant considéré le modèle PS et sa généralisation comme sujet principal dans cette thèse et l'observation fondamentale de la non homogénéité de l'ADN nous amènent à étudier la version désordonnée en ajoutant une perturbation aléatoire au modèle homogène. On se propose d'abord de faire une étude sommaire du modèle d'accrochage désordonné unidimensionnel.

Accrochage unidimensionnel

Soit $\omega = \{\omega_n\}_{n=1,2,\ldots}$ une suite de variables aléatoires i.i.d. centrée de variance unitaire, de loi $\mathbb{P}$ et vérifiant

$$Q(\beta) := \mathbb{E}[\exp(\beta\omega_1)] < \infty,\ \forall \beta > 0\,. \tag{1.58}$$

Pour tout $N \in \mathbb{N}$, $h \in \mathbb{R}$ et $\beta \geq 0$, nous définissons les mesures de Gibbs:

$$\frac{\mathrm{d}\mathbf{P}^{\beta,h}_{N,\omega}}{\mathrm{d}\mathbf{P}} = \frac{1}{Z^{\beta,h}_{N,\omega}}\exp\left(\sum_{n=1}^{N}(\beta\omega_n + h)\delta_n\right)\delta_N\,, \tag{1.59}$$

où

$$Z^{\beta,h}_{N,\omega} = \mathbf{E}\left[\exp\left(\sum_{n=1}^{N}(\beta\omega_n + h)\delta_n\right)\delta_N\right], \tag{1.60}$$

est la fonction de partition du système avec contrainte.

Considérons θ l'opérateur de translation de l'environnement, i.e. $\theta(\omega_n) = \omega_{n+1}$, alors pour tout $N, N' \in \mathbb{N}$, on a la propriété suivante

$$\log Z^{\beta,h}_{N+N',\omega} \geq \log Z^{\beta,h}_{N,\omega} + \log Z^{\beta,h}_{N',\theta^N\omega}\,. \tag{1.61}$$

Appliquons le théorème ergodique sur-additif de Kingman [65] avec la propriété sur-additive ci-dessus, l'énergie libre gelée ou *quenched* du système est donnée par

Théorème 1.11. *La limite*

$$F(\beta,h) := \lim_{N\to\infty}\frac{1}{N}\log Z^{\beta,h}_{N,\omega} = \sup_{N\geq 1}\frac{1}{N}\mathbb{E}\left[\log Z^{\beta,h}_{N,\omega}\right], \tag{1.62}$$

existe et est constante $\mathbb{P}$*-p.s.*

La fonction $h \mapsto F(\beta,h)$ est convexe et croissante en h en tant que la limite des fonctions convexes et croissantes. On définit le point critique *quenched* par

$$h_c(\beta) := \inf\{h : F(\beta,h) > 0\}\,. \tag{1.63}$$

On définit aussi l'énergie libre du système recuit ou *annealed* par

$$F^a(\beta,h) := \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\left[Z^{\beta,h}_{N,\omega}\right], \tag{1.64}$$

et le point critique *annealed* $h^a_c(\beta)$, tel que $F^a(\beta,h) > 0$ si et seulement si $h > h^a_c(\beta)$. L'inégalité de Jensen et le théorème de Fubini donnent

$$\mathbb{E}\left[\log Z^{\beta,h}_{N,\omega}\right] \leq \log\mathbb{E}\left[Z^{\beta,h}_{N,\omega}\right] = \log\mathbf{E}\left[\exp\left((\log Q(\beta)+h)\sum_{n=1}^{N}\delta_n\right)\delta_N\right], \tag{1.65}$$

en divisant par N et en passant à la limite, on obtient

$$F(\beta,h) \leq F(0,h+\log Q(\beta))\,, \tag{1.66}$$

et par conséquent $h^a_c(\beta) \leq h_c(\beta)$. Puisque $\beta \mapsto \mathbb{E}[\log Z^{\beta,h}_{N,\omega}]$ est convexe et croissante alors $\beta \mapsto F(\beta,h)$ l'est aussi et $F(0,h) \leq F(\beta,h)$ ce qui conduit à voir que $h_c(\beta) \leq h_c(0) = 0$. Notons que dans [2] il est prouvé que cette dernière inégalité est stricte. On se trouve alors devant deux questions importantes: Les points critiques *quenched* et *annealed* coïncident-ils? L'exposant critique *quenched* est-il égale à l'exposant critique *annealed*?

Nous parlons de pertinence du désordre lorsque le désordre rend le comportement critique du modèle différent de celui du modèle non désordonné. Le désordre est non pertinent si les deux comportements critiques coïncident.

Pour étudier le changement du comportement critique après l'introduction d'un désordre à un modèle homogène, le physicien A.B. Harris donne une méthode générale développée pour le modèle d'Ising dans son article original [58]. Il a prédit que la pertinence ou

la non pertinence du désordre ne dépend que de l'exposant ν^{hom} de la transition de phase homogène où $\nu^{hom} = \max(1, 1/\alpha)$ (théorème 1.2): si $\nu^{hom} < 2$ ($\alpha > 1/2$) alors le désordre est pertinent et si $\nu^{hom} > 2$ ($\alpha < 1/2$) alors le désordre est non pertinent. Par contre, le critère de Harris ne donne pas des informations sur le cas marginal ($\alpha = 1/2$).

Nous présentons une explication heuristique de ce critère dans notre modèle. Comparer le comportement des énergies libres *quenched* et *annealed* revient à comparer le comportement de $Z_{N,\omega}^{\beta,h}$ et de son espérance lorsque $N \to \infty$ (ils ont le même comportement si $\mathbb{E}[(Z_{N,\omega}^{\beta,h})^2]$ reste bornée lorsque $N \to \infty$). Nous calculons alors le second moment de $Z_{N,\omega}^{\beta,h}$ pour $h = h_c^a(\beta)$ (car pour $h > h_c^a(\beta)$, nous avons $\mathbb{E}[(Z_{N,\omega}^{\beta,h})^2] = +\infty$ pour tout $\beta > 0$). Utilisons le théorème de Fubini

$$\mathbb{E}\left[\left(Z_{N,\omega}^{\beta,h_c^a(\beta)}\right)^2\right] = \mathbf{E}^{\otimes 2}\left[\exp\left(\sum_{n=1}^{N}(\log Q(2\beta) - 2\log Q(\beta))\mathbf{1}_{n\in\tau\cap\tau'}\right)\right], \tag{1.67}$$

où τ et τ' sont deux copies indépendantes de τ et $\mathbf{P}^{\otimes 2}$ est leur loi produit. Le second moment de $Z_{N,\omega}^{\beta,h_c^a(\beta)}$ correspond alors à une fonction de partition d'un modèle homogène de paramètre $\log Q(2\beta) - 2\log Q(\beta)$. Si $\tau \cap \tau'$ est récurrent alors $\mathbb{E}[(Z_{N,\omega}^{\beta,h_c^a(\beta)})^2]$ diverge, on obtient alors le critère suivant

$$\text{Le désordre est pertinent} \Leftrightarrow \sum_{n\geq 1}\mathbf{P}(n\in\tau)^2 = +\infty\,. \tag{1.68}$$

La question de pertinence du désordre a été résolue dans le cadre du modèle d'accrochage pour toute valeur de α:

- La non pertinence du désordre pour $\alpha < 1/2$ a été démontrée par plusieurs méthodes: la méthode des moments secondaires [1, 87], la technique des martingales [69] et l'approche des grandes déviations [31].

- Plusieurs travaux ont étudié le cas $\alpha > 1/2$ [1, 3, 35] et ont utilisé la méthode des moments fractionnaires pour trouver l'écart entre les points critiques *quenched* et *annealed*. Cette méthode part de l'observation suivante: pour tout $\delta \in (0,1)$

$$F(\beta,h) = \lim_{N\to\infty}\mathbb{E}\frac{1}{\delta N}\log(Z_{N,\omega}^{\beta,h})^\delta \overset{Jensen}{\leq} \lim_{N\to\infty}\frac{1}{\delta N}\log\mathbb{E}[(Z_{N,\omega}^{\beta,h})^\delta]\,. \tag{1.69}$$

  Alors si pour un certain $h_0 > h_c^a(\beta)$, on a $\frac{1}{N}\log\mathbb{E}[(Z_{N,\omega}^{\beta,h_0})^\delta] \to 0$ lorsque $N \to \infty$, (1.69) montre que $F(\beta,h_0) \leq 0$ et par suite $h_c(\beta) \geq h_0 > h_c^a(\beta)$. L'estimation du moment fractionnaire se fait par un changement de mesure (en introduisant une nouvelle mesure $\widetilde{\mathbb{P}}$ absolument continue par rapport à $\mathbb{P}$) et en utilisant l'inégalité de Hölder pour obtenir

$$\mathbb{E}[(Z_{N,\omega}^{\beta,h_0})^\delta] \leq \widetilde{\mathbb{E}}[Z_{N,\omega}^{\beta,h_0}]^\delta\,\mathbb{E}\left[\left(\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right)^{-\frac{\delta}{1-\delta}}\right]^{1-\delta}. \tag{1.70}$$

Le choix de la mesure $\widetilde{\mathbb{P}}$ est important pour que le premier terme dans la partie droite de (1.70) soit petit et que le deuxième terme ne soit pas trop grand. Appliquant la procédure de *coarse graining*, qui consiste à diviser le système en blocs et de faire les estimations sur chaque bloc, il suffit de montrer que $\mathbb{E}[(Z_{N,\omega}^{\beta,h_0})^\delta]$ est très petit pour N de l'ordre de la longueur de corrélation du système *annealed*. Dans les travaux mentionnés en début de ce paragraphe, il a été prouvé que

$$h_c(\beta) - h_c^a(\beta) \overset{\beta\searrow 0}{\asymp} \beta^{\frac{2\alpha}{2\alpha-1}\vee 2}. \tag{1.71}$$

Une asymptotique précise de $h_c(\beta) - h_c^a(\beta)$ quand $\beta \to 0$ et a été démontrée dans [7] pour $\alpha > 1$ et dans [28] pour $\alpha \in (1/2, 1)$.

- Le cas marginal $\alpha = 1/2$ a été un sujet de conflit notamment dans [36, 44] qui ont donné des prédictions contradictoires. Ce cas a été résolu dans plusieurs travaux [1, 53, 54, 89] et récemment dans [10] avec des estimations précises.

Une étude complémentaire sur l'influence du désordre est de voir comment l'exposant critique de l'énergie libre se modifie. Basant sur des idées de grandes déviations, il a été montré dans [31, 56] que le désordre (sous certaines conditions) lisse la courbe de l'énergie libre, c'est à dire que la transition de phase est au moins d'ordre 2. Par exemple, dans le cas d'un désordre gaussien [48, Section 5.4] l'énergie libre vérifie la relation suivante: pour tout $\beta > 0$, $h \geq h_c(\beta)$ et $\alpha \geq 0$, on a

$$0 \leq F(\beta, h) \leq \frac{1+\alpha}{2\beta^2}(h - h_c(\beta))^2. \tag{1.72}$$

C'est le phénomène de lissage de la transition de phase par le désordre.

Accrochage bidimensionnel

On s'intéresse de nouveau au modèle désordonné mais cette fois pour le modèle généralisé. L'étude d'un tel modèle a été faite dans un cadre physique (numériquement) pour différents types du désordre [41, 46, 47, 83]. Il est sans doute logique de penser à introduire le désordre au modèle gPS par une suite $\omega_{i,j} = f(\omega_i, \omega_j)$ pour un choix convenable de la fonction f et de la suite $\{\omega_i\}_{i=1,2,\ldots}$. On peut par exemple les choisir en prenant en considération les *matches* et les *mismatches* qui apparaissent uniquement dans ce modèle généralisé et/ou penser à la différence entre les liaisons A-T et C-G. On peut aussi attribuer à chaque brin une séquence de potentiels $\{\omega_i^{(1)}\}_{i=1,2,\ldots}$ et $\{\omega_i^{(2)}\}_{i=1,2,\ldots}$. Cependant, les corrélations dans le désordre rendent le traitement de ces modèles très difficile. Dans cette thèse, nous introduisons le désordre par une séquence de variables aléatoires i.i.d. à deux indices $\omega = \{\omega_{n,m}\}_{n,m\in\mathbb{N}}$ qui correspond à un modèle physique artificiel. Ce choix est fait pour son aspect résoluble et sa proximité du cas unidimensionnel d'une part, et pour comprendre la théorie de Harris de la (non) pertinence du désordre d'autre part.

Nous supposons que cette séquence est de moyenne nulle et de variance unitaire dont la loi est notée par $\mathbb{P}$ et vérifiant

$$Q(\beta) := \mathbb{E}[\exp(\beta\omega)] < \infty, \, \forall \beta > 0 \, . \tag{1.73}$$

Pour tout $N, M \in \mathbb{N}$, la famille de mesures de probabilité $\mathbf{P}^{\beta,h}_{N,M,\omega}$ pour $h \in \mathbb{R}$ et $\beta \in \mathbb{R}_+$ est définie par

$$\frac{\mathrm{d}\mathbf{P}^{\beta,h}_{N,M,\omega}}{\mathrm{d}\mathbf{P}}(\tau) := \frac{1}{Z^{\beta,h}_{N,M,\omega}} \exp\left(\sum_{n=1}^{N} \sum_{m=1}^{M} (\beta\omega_{n,m} + h) \mathbf{1}_{(n,m)\in\tau} \right) \mathbf{1}_{(N,M)\in\tau} \, , \tag{1.74}$$

avec

$$Z^{\beta,h}_{N,M,\omega} = \mathbf{E}\left[\exp\left(\sum_{n=1}^{N} \sum_{m=1}^{M} (\beta\omega_{n,m} + h) \mathbf{1}_{(n,m)\in\tau} \right) \mathbf{1}_{(N,M)\in\tau} \right] . \tag{1.75}$$

Une autre représentation de la fonction de partition en termes de la séquence τ est

$$Z^{\beta,h}_{N,M,\omega} = \mathbf{E}\left[\exp\left(\sum_{i=1}^{\mathcal{N}_{N,M}} (\beta\omega_{(\tau_i^{(1)},\tau_i^{(2)})} + h) \right), \tau_{\mathcal{N}_{N,M}} = (N, M) \right] , \tag{1.76}$$

avec $\mathcal{N}_{N,M}$ est le nombre de renouvellement dans $(0, N] \times (0, M]$.

En analogie avec le modèle homogène, nous nous intéressons à l'analyse de l'énergie libre pour étudier le comportement du modèle désordonné. Nous définissons l'énergie libre *quenched*

Théorème 1.12. *Pour tout $\gamma > 0$, $h \in \mathbb{R}$, $\beta \geq 0$ et pour tout $\{M(N)\}_{N=1,2,\ldots}$ tel que $\lim_{N\to\infty} M(N)/N = \gamma$, on a*

$$\lim_{N\to\infty} \frac{1}{N} \log Z^{\beta,h}_{N,M(N),\omega} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z^{\beta,h}_{N,M(N),\omega} := \mathrm{F}_\gamma(\beta, h) \, , \tag{1.77}$$

où la première limite existe $\mathbb{P}(d\omega)$-p.s. et dans $L^1(\mathbb{P})$.

La preuve de ce théorème est divisée en plusieurs étapes dont la première consiste à appliquer directement l'argument de super-additivité à $\log Z^{\beta,h}_{N,\gamma N,\omega}$ seulement lorsque $\gamma := p/q \in \mathbb{Q}$ (où p et q sont deux nombres entiers premiers et positifs) et $N/q \in \mathbb{N}$. Les étapes suivantes consistent à étudier séparément – en comparant les fonctions de partition convenables – les cas $\gamma \in \mathbb{Q}$ (sans la condition $\gamma N \in \mathbb{N}$), $\gamma \in \mathbb{R}_+$ et enfin le cas de toute suite vérifiant $M \sim \gamma N$ pour $N \to \infty$.

L'assertion (1.77) est valide aussi dans le cas avec extrémités libres, où on utilise la fonction de partition de la forme

$$Z^{f,\beta,h}_{N,M,\omega} = \sum_{n=0}^{N} \sum_{m=0}^{M} Z^{\beta,h}_{n,m,\omega} \sum_{i=N-n+1}^{\infty} \sum_{j=M-m+1}^{\infty} K(i+j) \, . \tag{1.78}$$

Cette forme particulière (en comparaison avec sa forme dans (1.33)) est utilisée pour des raisons techniques dans la preuve de la non pertinence du désordre.

Nous citons quelques propriétés de l'énergie libre:

- La fonction $(\beta, h) \mapsto \mathrm{F}_\gamma(\beta, h)$ est convexe car c'est la limite de fonctions convexes et la fonction $h \mapsto \mathrm{F}_\gamma(\beta, h)$ est croissante aussi comme étant la limite de fonctions croissantes.

- La fonction $\gamma \mapsto \mathrm{F}_\gamma(\beta, h)$ est localement lipschitizienne et cette propriété suit de la comparaison suivante

$$\mathrm{F}_{\gamma_1}(\beta, h) \leq \mathrm{F}_{\gamma_2}(\beta, h) \leq \frac{\gamma_2}{\gamma_1} \mathrm{F}_{\gamma_1}(\beta, h), \tag{1.79}$$

  pour tout $\gamma_2 \geq \gamma_1 > 0$. Nous utilisons des arguments élémentaires pour comparer les fonctions de partition convenables et par suite obtenir (1.79) (voir la preuve du théorème 3.1 dans le chapitre 3).

- La convexité de l'énergie libre et le fait que $\partial_\beta \mathrm{F}_\gamma(\beta, h)|_{\beta=0} = 0$ impliquent que $\beta \mapsto \mathrm{F}_\gamma(\beta, h)$ est croissante pour $\beta > 0$ et décroissante pour $\beta \leq 0$.

Comme dans le modèle désordonné unidimensionnel, l'étude de la fonction de partition *annealed* est l'un des principaux outils utilisés dans l'analyse de la pertinence du désordre. L'énergie libre *annealed* est donnée par $\mathrm{F}^a_\gamma(\beta, h) := \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[Z^{\beta,h}_{N,M(N),\omega}]$ et on a un point critique $h^a_c(\beta)$, tel que $\mathrm{F}^a_\gamma(\beta, h) > 0$ si et seulement si $h > h^a_c(\beta)$.

Les mêmes arguments utilisés pour trouver les bornes inférieures et supérieures du point critique et de l'énergie libre dans le cas unidimensionnel sont encore établis ici pour obtenir

$$h^a_c(\beta) \leq h_c(\beta) \leq h_c(0). \tag{1.80}$$

Pour montrer que la deuxième inégalité est stricte pour tout $\beta \neq 0$, nous utilisons la procédure du [48, théorème 5.2] dans un cadre plus général et approprié au cas bidimensionnel. En ce qui concerne la première inégalité, la question de savoir s'il y a effectivement égalité ou non est liée à la question de la pertinence et la non pertinence du désordre.

Dans le chapitre 3, nous montrons que le désordre est non pertinent pour $\alpha < 1$ et pertinent pour $\alpha > 1$. Pour le cas $\alpha = 1$, une difficulté au niveau du comportement asymptotique de la fonction de renouvellement est la raison principale au choix fait de ne pas traiter ce cas. Nous donnons un aperçu sur les résultats obtenus ainsi des idées principales utilisées pour les prouver.

Une idée importante dans l'étude du régime non pertinent est le contrôle du moment d'ordre deux de la fonction de partition au point critique *annealed*. Le théorème de Fubini donne (prenons pour la suite $M = M(N)$)

$$\mathbb{E}\left[\left(Z^{f,\beta,h^a_c(\beta)}_{N,M,\omega}\right)^2\right] = \mathbf{E}^{\otimes 2}\left[\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}(\log Q(2\beta) - 2\log Q(\beta))\mathbf{1}_{(n,m)\in\nu}\right)\right], \tag{1.81}$$

où $\nu := \tau \cap \tau'$ et $\mathbf{P}^{\otimes 2}$ est la loi jointe de deux processus de renouvellement indépendants τ et τ' (deux copies indépendantes de τ). On obtient alors une fonction de partition d'un

modèle d'accrochage homogène de paramètre $\log Q(2\beta) - 2\log Q(\beta)$ associé au processus ν. La divergence du second moment dépend alors de la récurrence du renouvellement ν. On a

$$\alpha < 1 \Rightarrow \sum_{n\geq 1} \mathbf{P}^{\otimes 2}\left((n,m) \in \nu\right) < \infty\,. \tag{1.82}$$

On reprend la technique utilisée dans [69] pour montrer le résultat suivant:

Théorème 1.13. *Si ν est transient, alors il existe $\beta_1 > 0$ tel que pour tout $\beta \in (0,\beta_1)$, on a $h_c(\beta) = h_c^a(\beta)$ et*

$$\lim_{h\searrow h_c(\beta)} \frac{\log \mathrm{F}_\gamma(\beta,h)}{\log(h-h_c(\beta))} = \frac{1}{\alpha}\,. \tag{1.83}$$

Il n'y a donc aucun déplacement du point critique ni de modification de l'exposant critique. Un résultat plus précis sur le comportement asymptotique de $\mathrm{F}_\gamma(\beta,h)$ est obtenu en appliquant la méthode des répliques utilisée dans [88]. Rappelons que nous exclurons de la discussion le cas marginal $\alpha = 1$ à cause d'une difficulté au niveau du processus de renouvellement.

Pour les résultats concernant le régime pertinent, on utilise la méthode des moments fractionnaires et le changement de la mesure pour estimer quantitativement le déplacement du point critique à toute température. Prenant en compte la structure bidimensionnelle des trajectoires des renouvellements, plus compliquée de celle du cas unidimensionnel (linéaire), des nouveautés quantitatives et qualitatives apparaissent: nous utilisons plus des estimations qui dépendent de la distance entre les renouvellements et la diagonale.

Théorème 1.14. *Pour $\alpha > 1$, on a $h_c(\beta) > h_c^a(\beta)$ pour tout $\beta > 0$. De plus, pour tout $\epsilon > 0$, il existe $\beta_\epsilon > 0$ tel que pour tout $\beta \leq \beta_\epsilon$ on a*

$$h_c(\beta) - h_c^a(\beta) \geq \begin{cases} \beta^{\frac{2\alpha}{\alpha-1}+\epsilon} & si\ \alpha \in (1,2] \\ \beta^4 |\log\beta|^{-(1+\epsilon)} & si\ \alpha > 2. \end{cases} \tag{1.84}$$

L'idée de la preuve consiste à observer d'abord pour tout $\delta \in (0,1)$

$$\mathrm{F}_\gamma^q(\beta,h) = \lim_{\substack{N\to\infty \\ M/N\to\gamma}} \frac{1}{\delta N}\mathbb{E}\log\left(Z_{N,M,\omega}^{\beta,h}\right)^\delta \leq \lim_{\substack{N\to\infty \\ M/N\to\gamma}} \frac{1}{\delta N}\log\mathbb{E}\left[\left(Z_{N,M,\omega}^{\beta,h}\right)^\delta\right]. \tag{1.85}$$

L'étape principale (analogue à celle présentée dans le cas unidimensionnel) est de choisir un point h près du point critique *annealed* et de montrer que le dernier membre dans (1.85) est nul en ce point et cela exige un changement de mesure pour obtenir l'analogue de (1.70). Nous appliquons de nouveau la procédure de *coarse graining* qui consiste à découper le système en blocs afin de montrer que $\mathbb{E}\left[\left(Z_{N,M,\omega}^{\beta,h}\right)^\delta\right]$ est très petit. Comme dans le cas unidimensionnel, cette procédure permet de ramener l'étude du système lorsque $N \to \infty$ à des estimées sur des systèmes pour N de l'ordre de la longueur de corrélation

du système *annealed* et pour borner le moment fractionnaire de la fonction de partition dans ce modèle, nous appliquons une procédure adaptée à la structure bidimensionnelle et basée sur un contrôle des termes hors de la diagonale. Les estimations de la fonction de renouvellement bidimensionnelle vont jouer un rôle important dans cette preuve surtout dans le choix de la nouvelle mesure. Plus précisément, il est plus probable que le processus de renouvellement visite les sites proches de la diagonale et un *tilting* de l'environnement par une valeur convenable λ est fait seulement pour $\omega_{n,m}$ tel que (n,m) proche de la diagonale.

De plus, basant sur des méthodes identiques à celles utilisées pour le cas $\alpha < 1$, on obtient

$$h_c(\beta) - h_c^a(\beta) \le \widetilde{L}(1/\beta)\,\beta^{\frac{2\alpha}{\alpha-1}\vee 4}\,, \tag{1.86}$$

où $\widetilde{L}(\cdot)$ est une fonction à variation lente.

En comparaison avec le cas unidimensionnel où l'exposant de l'énergie libre ne peut être plus petit que deux, le théorème 1.13 montre que cet exposant dans le cas généralisé est plus grand ou égale à un. Pour étudier l'influence du désordre sur l'exposant critique, la première idée était d'appliquer l'argument utilisé pour prouver (1.72) dans notre modèle. Cet argument n'était pas suffisant et le modèle nécessite une généralisation appropiée présentée dans le Chapitre 3 conduisant à la conjecture suivante: pour tout $\alpha > 0$ et $\beta > 0$

$$\limsup_{h\searrow h_c(\beta)} \frac{\log \mathrm{F}_\alpha(\beta,h)}{\log\left(h-h_c(\beta)\right)} \ge \begin{cases} \frac{2\alpha}{\alpha+1} & \text{for } \alpha \in (1,2)\,, \\ \frac{4}{3} & \text{for } \alpha \ge 2\,. \end{cases} \tag{1.87}$$

1.3 Perspectives et questions ouvertes

Cette thèse est donc consacrée à l'étude mathématique des différents transitions dans le modèle de Poland-Scheraga généralisé et elle est divisée en deux grandes parties:

- Dans le cas homogène où toutes les bases dans les deux brins sont de même type, les points de contact sont modélisés par un processus de renouvellement bidimensionnel. Dans le chapitre 2, nous étudions le comportement de l'énergie libre du système et les différentes transitions de phase. Ainsi nous nous intéressons à l'étude du comportement asymptotique de la fonction de partition et aux propriétés des trajectoires dans la région de Cramer.

- Dans le chapitre 3, en ajoutant un désordre au modèle, nous montrons que le désordre est non pertinent pour $\alpha < 1$. Nous utilisons aussi la méthode des moments fractionnaires pour calculer la différence des points critiques et par suite montrer la pertinence du désordre pour $\alpha > 1$.

De nombreuses questions restent ouvertes. Nous citons:

- **Région hors Cramer et estimations au point critique**. Obtenir des comportements asymptotiques précis pour $h > 0$ et $\gamma \notin (1/\gamma_c(h), \gamma_c(h))$ (resp. $h = 0$) est important pour comprendre le comportement du système lors de la transition de phase dans la phase localisée (resp. lors de la transition de dénaturation). Cela revient, comme mentionné plus haut, à montrer des résultats au niveau de la fonction de renouvellement bidimensionnelle.

- **Nombre de transitions dans la phase localisée**. La question de trouver le nombre de ces transitions est ouverte. Dans le chapitre 2, nous donnons des exemples avec zero, une ou deux transitions. Peut-il y en avoir plus? Sont-elles toujours en nombre fini?

- **Choix du désordre**. Nous traitons dans le chapitre 3 le cas d'un désordre introduit par une séquence de variables aléatoire i.i.d. à deux indices $\{\omega_{n,m}\}_{n,m=1,2,\ldots}$, tandis que la façon la plus logique pour introduire le désordre au système est d'établir un désordre correspondant à la complexité du système surtout l'apparition des liaisons *anormales* (*mismatches*). Par exemple, considérer une séquence $\{\omega_{n,m}\}_{n,m=1,2,\ldots}$ à deux valeurs u et v tel que la première valeur correspond à un *match* (A-T ou C-G) et la deuxième à un *mismatch* [83]. Il y a aussi l'exemple d'attribuer à chaque brin une séquence de potentiels $\{\omega_n^{(1)}\}_{n=1,2,\ldots}$ et $\{\omega_n^{(2)}\}_{n=1,2,\ldots}$ (indépendants ou corrélés) et de supposer que $\omega_{n,m} = \omega_n^{(1)}\omega_m^{(2)}$ ou $\omega_{n,m} = \omega_n^{(1)} + \omega_m^{(2)}$. La difficulté est due aux corrélations du désordre et ces modèles peuvent être traités dans certains cas (voir [8] pour le modèle PS standard).

- **Désordre I.I.D.** Dans le cas le plus simple considéré dans cette thèse, plusieurs questions restent ouvertes.

  1. Montrer qu'il y a un phénomène de lissage de la transition de phase par le désordre pour $\alpha > 1$.
  2. Résoudre le cas marginal (cas $\alpha = 1$) qui correspond à une compréhension poussée de la fonction de renouvellement bidimensionnelle couplée avec une stratégie de *coarse graining* et changement de mesure non triviales.
  3. Trouver les estimations précises pour étudier les propriétés des trajectoires en présence du désordre et comme dans le modèle PS [50, Chapitre 8], cette question n'est pas triviale.

- **Influence du désordre sur les transitions dans la phase localisée**. Cette question semble être difficile (même dans le cas d'un désordre i.i.d.) mais importante et indispensable pour arriver à traiter le modèle le plus proche à la réalité et comprendre le comportement du système.

- **Modèles liés**. Nous soulignons aussi qu'un lien se trouve entre le modèle gPS et les modèles présentés au début de l'introduction (ARN, ADN circulaire). Ces derniers ont été étudiés par les physiciens [61, 95] et il sera intéressant d'étudier ce lien dans un cadre mathématique. Ainsi le modèle désordonné est lié au modèle des polymères dirigés dans un environnement aléatoire [32, 94].

Chapter 2

The generalized Poland-Scheraga model: homogeneous case

Abstract

The Poland-Scheraga model describes the denaturation transition of two complementary – in particular, equally long – strands of DNA, and it has enjoyed a remarkable success both for quantitative modeling purposes and at a more theoretical level. The solvable character of the homogeneous version of the model is one of features to which its success is due. In the bio-physical literature a generalization of the model, allowing different length and non complementarity of the strands, has been considered and the solvable character extends to this substantial generalization. We present a mathematical analysis of the homogeneous generalized Poland-Scheraga model. Our approach is based on the fact that such a model is a homogeneous pinning model based on a bivariate renewal process, much like the basic Poland-Scheraga model is a pinning model based on a univariate, i.e. standard, renewal. We present a complete analysis of the free energy singularities, which include the localization-delocalization critical point and (in general) other critical points that have been only partially captured in the physical literature. We obtain also precise estimates on the path properties of the model. This chapter is work in collaboration with G. Giacomin (arXiv:1510.07996, submitted to Stoch. Proc. Appl.).

2.1 Introduction and main results

2.1.1 General overview

The localization-delocalization phenomenon in various polymer models has been the object of much attention in the physics, biophysics and mathematics literature [42, 48, 59, 63, 78]. One of the main biological and physical phenomenon that motivates this work is DNA denaturation, that is the separation of the two DNA strands at high temperature and, more generally, the fluctuation phenomena observed at lower temperatures, when the two strands are tied together. The most basic and studied model in this field is the Poland-Sheraga (PS) model [78] which is limited to the case of sharp complementarity of two equal length strands: only bases with the same index can form pairs. From the theoretical physics and mathematical viewpoint what is most remarkable in the homogeneous version of the model is its solvable character and the fact that at the delocalization (or denaturation) transition the behavior – i.e. the critical behavior – can be fully captured. A mathematical viewpoint on this solvable character and on the solution itself is that the Poland-Scheraga model is a Gibbs measure with only one body potentials and built on a one-dimensional process [48].

In [46, 47, 72] (see also [70, 82] for a primitive version), a generalization of the Poland-Sheraga (gPS) model has been introduced and the novelties are:

- The possibility of formation of non-symmetrical loops in the two strands (i.e., the contribution to a loop, in terms of number of bases, from the two strands is not necessarily the same).

- The two strands may be of different lengths.

These novelties are very substantial (we invite the reader to compare Figure 2.1 with [42, Fig. 6] or [50, Fig. 2.5]). Nevertheless, as already pointed out in [47, 72, 82], the solvable character is preserved. However, that the novelties are really substantial is witnessed by a richer phenomenology (partially captured and understood in [41, 72]): in addition to the expected denaturation transition, the gPS model undergoes other transitions.

Here we develop a mathematical analysis of the gPS model based on the observation that it is a pinning model based on a two-dimensional renewal process. Much like for the original PS model, tools from Renewal Theory allow going far toward a complete understanding of the model. Nevertheless, as we will explain, some important questions are still open and they correspond to open problems in the theory of two and higher dimensional renewal processes.

2.1.2 The gPS model: biophysics version

This subsection, as well as § 2.1.4, can be skipped if one is not focusing on the biophyisics set-up. The model we consider has been introduced in [47]. The two DNA strands, of lengths M and $N \geq 1$ – the length of course corresponds to the number of bases – interact by forming some base pairs. We talk of N-strand, M-strand and of base i of the $N-$ or $M-$strand with the obvious meaning. An allowed configuration of our system is a collection of base pairs

$$((i_1, j_1), (i_2, j_2), \dots, (i_n, j_n)) \in \mathbb{N}^{2n}, \quad \text{with } \mathbb{N} = \{1, 2, \dots\}, \tag{2.1}$$

where $n \in \mathbb{N} := \{1, 2, \dots, N\}$ and

1. $(i_1, j_1) = (1, 1)$;
2. $i_k < i_{k'}$, as well as $j_k < j_{k'}$, for $1 \leq k < k' \leq n$.

The first condition is simply saying that the first two bases form a pair and the second condition is imposed by the geometric constraint (see Figure 2.1). The weight of every configuration is assigned by the following rules:

1. Each base pair is energetically favored and carries an energy $-E_b < 0$;
2. A base which is not in a pair is either in a loop or in the free ends:
   - It is in a loop if it is in $L_k := ((i_k, i_{k+1}) \cup (j_k, j_{k+1})) \cap \mathbb{N}$ for some $k \in \{1, \dots, n-1\}$: the loop L_k has length $\ell_k := (i_{k+1} - i_k) + (j_{k+1} - j_k) - 2$ and we associate to L_k an entropy factor $B(\ell_k)$ with
   $$B(\ell) := s^\ell \ell^{-c}, \tag{2.2}$$
   where $s \geq 1$ and $c > 2$. There is also an energetic penalty $E_l > 0$ penalty associated to a loop.
   - The free ends have length $N - i_n$ and $M - j_n$ and to each free end we associate the entropy term $A(\ell) := s^\ell (\ell + 1)^{-\bar{c}}$ where $\bar{c}$ is another positive constant.

As we will see the value of s is irrelevant. The value of $\bar{c}$, chosen equal to 0.1 in [72], is somewhat more relevant, but what is very relevant is the value of c: in [72] it is chosen equal to 2.15.

These rules easily lead to a formula for the partition function, i.e. the sum of the weights over all allowed configurations, of our system

$$Z_N^M := \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} A(i) A(j) W_{N-i}^{M-j}, \tag{2.3}$$

where W_l^r obeys the recursion relation ($\beta \geq 0$ is proportional to the inverse of the temperature)

$$W_{m+1}^{r+1} = \exp(\beta E_b) W_m^r + \exp(\beta(E_b - E_l)) \sum_{\substack{i,i' : i+i'>0 \\ i<m, i'<r}} B(i+i') W_{m-i}^{r-i'}, \tag{2.4}$$

with $W_1^1 = 1$ and $W_1^i = W_i^1 = 0$ for $i > 1$.

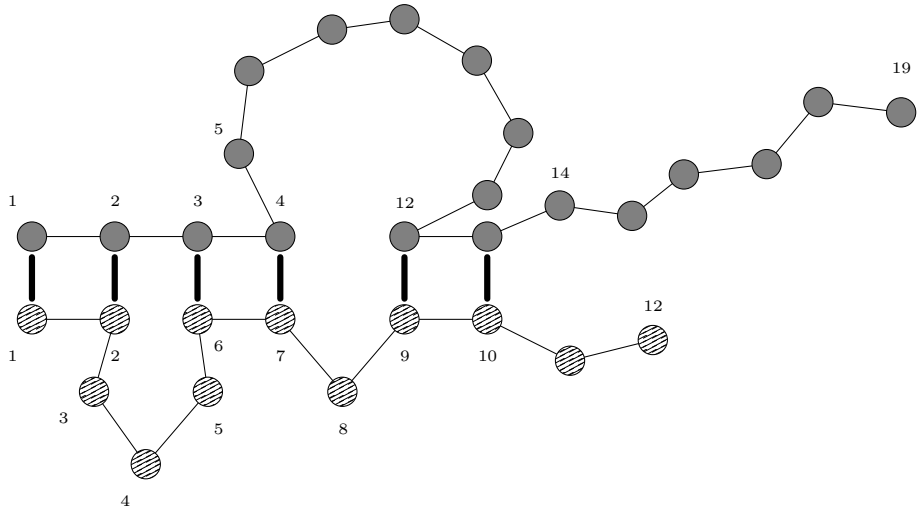

FIGURE 2.1. A representation of a trajectory of the gPS model in the biophysics representation. The first strand contains 12 bases, the second strand 19. The six base pairs determining the configuration are $(1,1)$, $(2,2)$, $(6,3)$, $(7,4)$, $(9,12)$ and $(10,13)$.

2.1.3 The gPS model: renewal process viewpoint

From a mathematical perspective we take a more general viewpoint and we introduce a two-dimensional renewal pinning model. A discrete two-dimensional renewal issued from the origin is a random walk $\tau = \{\tau_n\}_{n=0,1,\ldots} = (\tau^{(1)}, \tau^{(2)}) = \{(\tau_n^{(1)}, \tau_n^{(2)})\}_{n=0,1,\ldots}$ where $\tau_0 = (0,0)$ and, for $n \in \mathbb{N} := \{1,2,\ldots\}$, τ_n is a sum of n independent identically distributed random variables taking values in $\mathbb{N}^2$. So if we set $\mathrm{K}(n,m) := \mathbf{P}(\tau_1 = (n,m))$ then given $\{(i_n, j_n)\}_{n=0,1,2,\ldots}$, with $(i_0, j_0) = (0,0)$, for every $k \in \mathbb{N}$

$$\mathbf{P}\left(\tau_n = (i_n, j_n) \text{ for } n = 1,2,\ldots,k\right) = \prod_{n=1}^{k} \mathrm{K}\left(i_n - i_{n-1}, j_n - j_{n-1}\right), \tag{2.5}$$

and, by construction, such a probability is zero unless the i's and j's form strictly increasing sequences.

We can then introduce for given N and $M \in \mathbb{N}$ a pinning model of length (N,M) by forcing, i.e. conditioning, τ to visit (N,M) and by penalizing ($h \leq 0$) or rewarding ($h \geq 0$) the number of renewals up to (N,M). More formally, we introduce the probability measure $\mathbf{P}_{N,M,h}$ by setting for every $k \in \mathbb{N}$ such that $k \leq \min(N,M) =: N \wedge M$ and for every

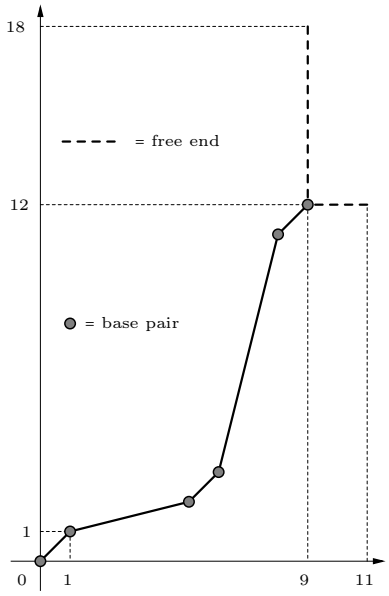

Figure 2.2. A representation of a trajectory of the gPS model in the renewal process representation, corresponding to Figure 2.1. The base pairs are now renewal points of a two-dimensional discrete renewal process: these points correspond to the (six) base pairs of Figure 2.1, except that the first base pair is now $(0,0)$ and also all the other ones are translated down of $(1,1)$ with respect to Figure 2.1. The renewal trajectory is drawn up to the renewal point $(9,12)$ and the trajectory up to this point corresponds to one of the possible trajectories of $Z^{c}_{9,12,h}$. The free ends, of lengths 2 and 6, are then represented as straight lines that go till the boundary of the rectangle with opposite vertices $(0,0)$ and $(11,18)$. $Z^{f}_{11,18,h}$ is obtained by summing up with respect to the position of the last renewal point – $(9,12)$ in this example – the contribution of the constrained partition function times the contribution due to the two free ends.

$\{(i_n,j_n)\}_{n=0,1,\dots,k}$ with $(i_0,j_0)=(0,0)$, $i_n-i_{n-1}>0$ as well as $j_n-j_{n-1}>0$ for $n=1,\dots,k$ and $(i_k,j_k)=(N,M)$

$$\frac{\mathbf{P}_{N,M,h}\left(\tau_n=(i_n,j_n) \text{ for } n=1,2,\dots,k\right)}{\mathbf{P}\left(\tau_n=(i_n,j_n) \text{ for } n=1,2,\dots,k\right)} := \frac{1}{\mathcal{Z}_{N,M,h}}\exp\left(hk\right), \tag{2.6}$$

where $\mathcal{Z}_{N,M,h}$ is the partition function (or normalization constant):

$$\mathcal{Z}_{N,M,h} := \mathbf{E}\left[\exp\left(h|\tau\cap([1,N]\times[1,M])|\right)\mathbf{1}_{(N,M)\in\tau}\right], \tag{2.7}$$

in which we are interpreting τ as a random subset of $\mathbb{N}^2$ and $|\cdot|$ denotes the cardinality. Note that $\mathcal{Z}_{0,0,h}=1$, as well as $\mathcal{Z}_{0,M,h}=\mathcal{Z}_{N,0,h}=0$. Of course $\mathbf{P}_{N,M,h}$ requires $\mathcal{Z}_{N,M,h}\neq 0$ and whether this is the case or not depends on the inter-arrival distribution $\mathrm{K}(n,m)$ and, possibly, on N and M.

We have used the atypical notation $\mathcal{Z}$ instead of Z because the latter is going to be employed for the model on which we really focus: we consider in fact a very special choice of the *inter-arrival distribution* $\mathrm{K}(\cdot,\cdot)$, namely $\mathrm{K}(n,m)=K(n+m)$ where $K:\{2,3,\dots\}\to(0,\infty)$ is not a probability distribution and

$$K(n) := \frac{L(n)}{n^{2+\alpha}}, \tag{2.8}$$

where $L(\cdot)$ is a slowly varying function and $\alpha\geq 0$ (see Appendix 2.A for the properties of slowly varying functions). With this definition, we have that $K(n)>0$ for every n: all statements generalize to the case in which $K(n)=0$ for finitely many n, but we make this choice aiming at conciseness of some proofs. We require $\sum_{n,m\in\mathbb{N}}K(n+m)=1$ and of

course $\sum_{n,m\in\mathbb{N}} K(n+m) = \sum_{m=1}^{\infty} mK(m+1)$. We introduce then the *constrained partition function* $Z^c_{N,M,h}$ which coincides with $\mathcal{Z}_{N,M,h}$ once the specific choice of the inter-arrival is made. In an alternative explicit fashion

$$Z^c_{N,M,h} := \sum_{n=1}^{N\wedge M} \sum_{\substack{\underline{l}\in\mathbb{N}^n:\\ |\underline{l}|=N}} \sum_{\substack{\underline{t}\in\mathbb{N}^n:\\ |\underline{t}|=M}} \prod_{i=1}^{n} \exp(h) K\left(l_i + t_i\right) , \tag{2.9}$$

where $|\underline{l}| = \sum_{i=1}^{n} l_i$. The *free partition function* is then defined by

$$Z^f_{N,M,h} := \sum_{i=0}^{N} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h} , \tag{2.10}$$

where $K_f : \{0\} \cup \mathbb{N} \to (0,\infty)$ is defined as $K_f(n) := \overline{L}(n)/n^{\overline{\alpha}}$ for every $n \geq 1$ and $K_f(0) = 1$ (an arbitrary choice: there is no loss of generality with respect to requiring just $K_f(0) > 0$ and, once again, one can even allow $K_f(n) = 0$ for finitely many n, but we choose positivity for conciseness) with $\overline{\alpha} \in \mathbb{R}$. The free partition function is the normalization associated to the probability $\mathbf{P}^f_{N,M,h}$ defined by setting for every $k \in \mathbb{N}$ such that $k \leq \min(N,M)$ and for every $\{(i_n, j_n)\}_{n=0,1,\ldots,k}$ with $(i_0, j_0) = (0,0)$, $i_n - i_{n-1} > 0$ as well as $j_n - j_{n-1} > 0$ for $n = 1,\ldots,k$ and i_k (respectively, j_k) that does not exceed N (respectively, M)

$$\begin{aligned} \mathbf{P}^f_{N,M,h}\left(\tau \cap [1,N]\times[1,M] = \{(i_1,j_1),(i_2,j_2),\ldots,(i_k,j_k)\}\right) := \\ \frac{\exp(hk)}{Z^f_{N,M,h}} K_f(N-i_k) K_f(M-j_k) \mathbf{P}\left(\tau_n = (i_n,j_n) \text{ for } n = 1,2,\ldots,k\right) . \end{aligned} \tag{2.11}$$

Here we introduce also the free energy density (the existence of the limit is proven in Section 2.4)

$$\mathrm{F}_\gamma(h) := \lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N} \log Z^f_{N,M,h} , \tag{2.12}$$

for $\gamma > 0$ or, as we will often do without loss of generality, for $\gamma \geq 1$. The limit in (2.12) means: for every $\{(N_n, M_n)\}_{n=1,2,\ldots}$ with $\lim_n M_n/N_n = \gamma$ and $\lim_n N_n = \infty$.

2.1.4 Matching biophysics and renewal process viewpoints

The first remark to make on the biophysics model is that the dependence on s of Z^M_N (cf. (2.3)) is trivial: Z^M_N for a given value of s coincides with $s^{N+M}\widetilde{Z}^M_N$, where $\widetilde{Z}^M_N$ is Z^M_N with $s = 1$ and both E_b and E_l are decreased by $\frac{2}{\beta}\log s$. We can therefore set $s = 1$ without true loss of generality. We then remark that we can match $\exp(-\beta E_b) Z^{M+1}_{N+1}$ and $Z^f_{N,M,h}$ by an appropriate choice of h, $K(\cdot)$ and $K_f(\cdot)$: the fact that we consider the biophysics model with lengths augmented by 1 and renormalized by the factor $\exp(-\beta E_b)$ is just because in parallel with one-dimensional renewal pinning works we have chosen to start the renewal from *time* (or *renewal epoch*) zero, but without giving an energetic reward to the base pair 0. Matching is then made by observing that:

1. The match for the free end terms $K_f(\cdot)$ and $A(\cdot)$ is easily made.

2. A *minimal* inter-arrival step, that is $(1,1)$, in the renewal model corresponds to a base pair and contributes $K(2)\exp(h)$. It would then be matched to $\exp(\beta E_b)$ in the biophysics model.

3. All other inter-arrival steps (i,i'), i.e. $i+i' = 3,4,\dots$, which give a contribution $K(i+i')\exp(h)$ correspond to loops with $i+i'-2$ unpaired bases and the contribution in the biophysics model is $\exp(\beta(E_b - E_l))B(i+i'-2)$.

Then of course $c = 2+\alpha$ and $\bar{c} = \bar{\alpha}$ and, at last, the matching between the two models is done thanks to the ample freedom that we have in choosing $K(\cdot)$, except that we have required that $\sum_{n,m} K(n+m) = 1$. This *probability constraint* in reality just corresponds to a shift in the parameter h. In a nutshell, renewal models include the biophysics ones: from a qualitative view point the matching is immediate, form the quantitative one it requires some bookkeeping care. Explicit matchings are presented in Section 2.3.4.

2.1.5 The free energy and the localization transition

We start from:

Proposition 2.1. *Recall (2.12). We have*

$$\lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N}\log Z^c_{N,M,h} \;=\; \lim_{\substack{N,M\to\infty:\\ \frac{M}{N}\to\gamma}} \frac{1}{N}\log Z^f_{N,M,h}\,. \tag{2.13}$$

In practice, $Z^c_{N,M,h}$ is a more fundamental quantity for our computations and we will first identify the free energy density by looking at the exponential growth of this quantity and only after, in Section 2.4, we will match it with the exponential growth rate in the free case.

Note that we are just speaking of exponential *growth* and not of *decrease*. In fact $\mathrm{F}_\gamma(h) \ge 0$ simply because $Z^f_{N,M,h} \ge K_f(0)^2 Z^c_{N,M,h} = Z^c_{N,M,h} \ge \exp(h)K(N+M)$. As a matter of fact, this is a very important issue because it is natural to set

$$h_c := \inf\{h : \mathrm{F}_\gamma(h) > 0\} = \max\{h : \mathrm{F}_\gamma(h) = 0\}\,, \tag{2.14}$$

where the equality on the right comes from the fact that $\mathrm{F}_\gamma(\cdot)$ is locally bounded, convex (hence continuous) and non-decreasing. These facts are evident from the definitions, like the following two preliminary observations:

1. $Z^c_{N,M,h} \le 1$ for $h \le 0$: hence $h_c \ge 0$;

2. we will see just below that $h_c = 0$, but it is worth pointing out that $h_c < \infty$ by elementary arguments. For example: $h_c \le -\log K(2)$ because $Z^f_{N,M,h} \ge K_f(0)K_f(M-N)Z^c_{N,N,h}$ and $Z^c_{N,N,h} \ge (\exp(h)K(2))^N$.

From (2.14) we readily see that h_c is a non analyticity point of $\mathrm{F}_\gamma(\cdot)$ and there is a phase transition of the system. By standard arguments based on convexity (see for example [48, Section 1.2] for the univariate pinning model) one realizes that this transition is the denaturation, or localization/delocalization, transition: $\partial_h \mathrm{F}_\gamma(h)$ – in case to be interpreted as, say, left derivative, but we will soon see that $\partial_h \mathrm{F}_\gamma(h)$ exists except, in some cases, at $h = h_c$ – is the density of base pairs (or contact fraction), which is therefore positive, respectively zero, for $h > h_c$, respectively $h < h_c$.

The next result is much more quantitative about this transition: since $Z^f_{N,M,h} \ge K_f(0)K_f(M-N)Z^c_{N,N,h}$, we get that $\mathrm{F}_\gamma(h) \ge \mathrm{F}_1(h)$ for every $\gamma \ge 1$. All asymptotic statements in the next theorem are for $h \searrow 0$:

Theorem 2.2. *For every $\alpha \ge 0$ and $\gamma \ge 1$ we have $h_c = 0$ and there exists $h_{c,\gamma} \in (0,\infty]$ such that $\mathrm{F}_\gamma(\cdot)$ is real analytic in $(-\infty,0) \cup (0,h_{c,\gamma})$ and $h_{c,\gamma}$ is a non analyticity point (when $h_{c,\gamma} < \infty$). Moreover if $\sum_n n^2 K(n) < \infty$, a condition implied by $\alpha > 1$, we have as $h \searrow 0$*

$$\mathrm{F}_\gamma(h) \sim \mathrm{F}_1(h) \sim c\,h\,, \tag{2.15}$$

with $c^{-1} := \frac{1}{2}\sum_{n=2}^{\infty} n(n-1)K(n)$. If instead $\sum_n n^2 K(n) = \infty$, implied by $\alpha \in [0,1)$, there exists $c_{\alpha,\gamma} \ge 1$ such that

$$\mathrm{F}_\gamma(h) \sim c_{\alpha,\gamma}\mathrm{F}_1(h) \quad \text{and} \quad \mathrm{F}_1(h) \sim L_\alpha(h)h^{1/\alpha} \tag{2.16}$$

where $L_\alpha(\cdot)$ is slowly varying at 0. In the case $\alpha = 0$, (2.16) should be interpreted as $\mathrm{F}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$.

In Section 2.3.3 $c_{\alpha,\gamma}$ and $L_\alpha(\cdot)$ are determined. The expression of $c_{\alpha,\gamma}$ implicitly contains nontrivial information on the system, see Proposition 2.22.

We will see that it may be that $h_{c,\gamma} = \infty$, for example $h_{c,1} = \infty$ in full generality, but when $h_{c,\gamma} < \infty$, $h_{c,\gamma}$ may not be the only critical point inside the localized regime (Theorem 2.5). This means that there is more than one localized phase in the system: this is what we treat next, but we need to introduce more concepts and definitions. By doing so we will start outlining the proof of Theorem 2.2.

2.1.6 Transitions in the localized regime

A crucial elementary observation is that (2.9) can be written as

$$\begin{aligned} Z^c_{N,M,h} &= \exp\left((N+M)\mathrm{G}\right) \sum_{n=1}^{N\wedge M} \sum_{\substack{\underline{l}\in\mathbb{N}^n:\\ |\underline{l}|=N}} \sum_{\substack{\underline{t}\in\mathbb{N}^n:\\ |\underline{t}|=M}} \prod_{i=1}^{n} \widetilde{K}_h(l_i,t_i) \\ &= \exp\left((N+M)\mathrm{G}\right) \mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right), \end{aligned} \tag{2.17}$$

where

- $\widetilde{K}_h(n,m) = \exp(h-(n+m)\mathrm{G})K(n+m)$ – note that $\widetilde{K}_h(n,m)$ is just a function of $n+m$ – and $\mathrm{G} = \mathrm{G}(h)$ is the only solution to

$$\sum_{n,m} K(n+m)\exp(h-(n+m)\mathrm{G}) = 1, \tag{2.18}$$

  when such a solution exists (that is, when $h \geq 0$), and $\mathrm{G} = 0$ otherwise. We have therefore defined a function $\mathrm{G}:\mathbb{R}\to[0,\infty)$.

- $\widetilde{\tau}_h$ is the two-dimensional renewal issued from $(0,0)$ with inter-arrival distribution $\widetilde{K}_h$: if $h<0$ then $\widetilde{K}_h:\mathbb{N}^2\to[0,1)$ is a sub-probability that we make a probability by defining $\widetilde{K}_h(\infty) := 1-\sum_{(n,m)\in\mathbb{N}^2}\widetilde{K}_h(n,m)$ and, in this case, $\widetilde{\tau}_h$ contains a.s. a finite number of points (and $\{\infty\}$) and we refer to it as a *terminating* renewal.

$\mathrm{G}(h)$ accounts for part of the free energy of the system and some basic features are straightforward:

Lemma 2.3. *The function $\mathrm{G}(\cdot)$ is convex and real analytic except at $h=0$.*

Proof. First of all the function is well defined for $h\geq 0$ because the function $\mathrm{G}\mapsto\sum_{n,m}K(n+m)\exp(-(n+m)\mathrm{G})$ is real analytic on the positive semi-axis and decreases from 1 to 0 as G goes from 0 to ∞. In particular, $\mathrm{G}(0)=0$. Analyticity of $\mathrm{G}(\cdot)$ on the positive semi-axis follows directly by the Inverse Function Theorem (for analytic functions: see for example [66, Sec. 2.5]). Convexity can be proven directly by differentiating the implicit expression (2.18). A less tedious proof of convexity can be achieved by recognizing that $\mathrm{G}(h)$ is a free energy (see (2.64) below). □

In view of (2.17) we see that (limits are with $M\sim\gamma N$)

$$\lim\frac{1}{N}\log Z^c_{N,M,h} = (1+\gamma)\mathrm{G}(h) + \lim\frac{1}{N}\log\mathbf{P}\left((N,M)\in\widetilde{\tau}_h\right). \tag{2.19}$$

Let us focus on $h>0$. The inter-arrivals of $\widetilde{\tau}_h$ are exponentially integrable, in particular they have finite mean:

$$\mu_h = \left(\mu_h^{(1)},\mu_h^{(2)}\right) := \mathbf{E}[(\widetilde{\tau}_h)_1] = \left(\sum_{n,m} n\widetilde{K}_h(n,m), \sum_{n,m} m\widetilde{K}_h(n,m)\right). \tag{2.20}$$

Evidently $\mu_h^{(1)} = \mu_h^{(2)}$ and the Law of Large Numbers directly implies that $\widetilde{\tau}_h$ stays close to the diagonal of the first quadrant. But then the event $\{(N,M) \in \widetilde{\tau}_h\}$ is a Large Deviation event for $\gamma \neq 1$ and it contributes to the free energy: the singularities of the free energy for $h > 0$ come from this extra Large Deviation contribution and the key word to understand these new transitions is *Cramér regime*. The point in fact is whether or not the Large Deviation event can be made typical by an exponential change of measure (a *tilt*): the larger γ is the more possible it is that a tilt does not suffice and the typical Large Deviation trajectories will not correspond to a tilt of the measure (in this case we say that we are outside of the Cramér regime). On the other hand, the interaction strength directly impacts whether or not the process is in the Cramér regime. The formulas that follow precisely characterize the switching between Cramér and non Cramér regimes.

We introduce the convex function $q_h : \mathbb{R}^2 \to (0,\infty]^2$

$$q_h(\lambda) = q_h(\lambda_1,\lambda_2) := \sum_{n,m} e^h K(n+m) \exp\left(-(\mathrm{G}(h)-\lambda_1)n - (\mathrm{G}(h)-\lambda_2)m\right), \tag{2.21}$$

which is bounded in $(-\infty, \mathrm{G}(h)]^2$ and it is analytic in the interior of this domain. We set for $h > 0$

$$\overline{\lambda}_1(h) := \sup\{\lambda_1 < 0 : q_h(\lambda_1, \mathrm{G}(h)) \le 1\}, \tag{2.22}$$

and, since $q_h(\lambda_1, \mathrm{G}(h))$ increases continuously in λ_1 from $q_h(-\infty, \mathrm{G}(h)) = 0$ to $q_h(0, \mathrm{G}(h)) > 1$, $\overline{\lambda}_1(h)$ is negative and it is characterized by $q_h(\overline{\lambda}_1, \mathrm{G}(h)) = 1$. Finally, we set, always for $h > 0$

$$\gamma_c(h) := \frac{\sum_{n,m} mK(n+m)\exp\left(-n\left(\mathrm{G}(h)-\overline{\lambda}_1(h)\right)\right)}{\sum_{n,m} nK(n+m)\exp\left(-n\left(\mathrm{G}(h)-\overline{\lambda}_1(h)\right)\right)}, \tag{2.23}$$

which is the ratio of averaged loop sizes on the different strands. Both denominator and numerator are bounded because for $c > 0$ and for every $\alpha \ge 0$

$$\sum_{n,m}(n+m)K(n+m)e^{-cn} = \sum_{t=2}^{\infty} tK(t) \sum_{n=1}^{t-1} \exp(-cn) \le \frac{1}{e^c-1}\sum_{t=2}^{\infty} tK(t) < \infty. \tag{2.24}$$

Here are some properties (see Section 2.3.3 for the proof):

Lemma 2.4. *Choose $\alpha > 0$. The function $\gamma_c : (0,\infty) \longrightarrow (1,\infty)$ is real analytic and*

$$\gamma_c(0) := \lim_{h \searrow 0} \gamma_c(h) = \frac{1}{\alpha} \vee 1 \quad \text{and} \quad \gamma_c(\infty) = \frac{\sum_m mK(1+m)}{\sum_m K(1+m)}. \tag{2.25}$$

The examples worked out in Section 2.3.4 show that $\gamma_c(\cdot)$ can have various behaviors: in particular, in general it is not monotonic.

Theorem 2.5. *Fix $\gamma \ge 1$. The function $\mathrm{F}_\gamma(\cdot)$ is analytic on $\{h : h > 0$ such that $\gamma_c(h) - \gamma \neq 0\}$ and $\mathrm{F}_\gamma(\cdot)$ is not analytic for the values $h > 0$ at which $\gamma_c(h) - \gamma$ changes sign. However, $\mathrm{F}_\gamma(\cdot)$ is C^1 on the positive semi-axis*

Theorem 2.5 is just a sample of the results we have and that can be gotten on these transitions that are transitions between localized regimes, because, by convexity of the free energy, the expected number of contacts does not decrease in h. In particular the *tangential case* – when $h \mapsto \gamma_c(h) - \gamma$ touches zero without changing sign – is treated in detail and while we can deal with most of the cases we are unable to produce a concise statement that says which zeros of $\gamma_c(h) - \gamma$ are critical points (some are not!) and, in general, what is the precise order of the transition. This is due to the fact that these transitions, unlike the denaturation transition, do depend on the details of $K(\cdot)$ and, to a certain extent, one needs to do a case by case study. Examples and more considerations on all these issues are developed at the end of the introduction and in Section 2.3.4.

2.1.7 Outline of the approach, sharp estimates and limit path properties

As we already mentioned, the cornerstone is (2.17). In fact (2.17) reduces sharp, respectively Laplace, estimates on $Z^c_{N,M,h}$ to sharp, respectively Laplace, estimates on the renewal function $\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right)$. A quick overview of the behavior of $\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right)$ is

1. If $h < 0$, so $\widetilde{\tau}_h$ is terminating, we will show that there exists $C_h > 0$ such that
$$\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right) \overset{N,M\to\infty}{\sim} C_h \widetilde{K}_h(N,M)\,. \tag{2.26}$$

2. For $h > 0$, recall (2.20) and the discussion right after that formula, one can show that
$$\lim_{\varepsilon \searrow 0} \lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(t\{x \in \mathbb{R}^2 : |x - v| \le \varepsilon\} \cap \widetilde{\tau}_h \neq \varnothing\right) = -D_h\left(v\right)\,, \tag{2.27}$$
where $D_h(\cdot)$ is a non-negative function defined in $\mathbb{R}^2$, but equal to $+\infty$ outside of the first quadrant. We shall see that $D_h(\cdot)$ is linear along rays, that is $D_h(sv) = sD_h(v)$, and $D_h(v) = 0$ if and only if $v \propto \mu_h$ (for us $v \propto (1,1)$). By the symmetry $\mathrm{K}(n,m) = \mathrm{K}(m,n)$, we get that $D_h(v_1,v_2) = D_h(v_2,v_1)$ so there is no loss of generality in sticking to $D(1,\gamma)$, $\gamma \ge 1$. Moreover we will then see that $\gamma \mapsto D(1,\gamma)$ is affine for γ larger than a critical value $\gamma_c(h) > 1$ or smaller than another critical value that coincides with $1/\gamma_c(h)$. On the other hand $\gamma \mapsto D(1,\gamma)$ is strictly convex in the interval $(1/\gamma_c(h), \gamma_c(h))$ – the *Cramér region* – and, when γ is in this interval, sharp asymptotic estimates are known. Namely, we have that
$$\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right) \sim \frac{c_v}{\sqrt{t}} \exp\left(-tD_h(v)\right)\,, \tag{2.28}$$
where $t = \sqrt{N^2+M^2}$, $v = (N,M)/t$, c_v is a positive constant (which depends of course also on h and $K(\cdot)$) and the asymptotic statement is for $t \to \infty$ and it is uniform provided that the unit vector v is in a compact arc of circle subset contained in an open arc of the unit circle that goes from $(\gamma_c(h),1)/\sqrt{1+(\gamma_c(h))^2}$ to $(1,\gamma_c(h))/\sqrt{1+(\gamma_c(h))^2}$.

We remark that by putting (2.17) and (2.28) together we readily see that we can make a substantial step ahead with respect to (2.19):

$$\lim_{\substack{N,M\to\infty:\\ M\sim\gamma N}} \frac{1}{N}\log Z^c_{N,M,h} \;=\; (1+\gamma)\mathrm{G} - D_h(1,\gamma)\,. \tag{2.29}$$

From this formula, since G is (implicitly) determined by (2.18) and since $D_h(\cdot)$ has a variational formulation, we will be able to use it to establish Theorem 2.2 and Theorem 2.5.

But with what we just outlined we can go beyond Laplace type estimates: (2.26) and (2.28) yield sharp estimates on $Z^c_{N,M,h}$ as $N\to\infty$ for $h<0$, with $M\sim\gamma N$, any $\gamma>0$. The same holds for $h>0$, but only for γ in the Cramér region. We state here the result for the free case, which is less immediate than the constrained one:

Theorem 2.6. *We have the following sharp estimates for $M\sim\gamma N$ and $\alpha>0$:*

1. *For $h>0$ and $\gamma\in(1/\gamma_c(h),\gamma_c(h))$ there exists $c_{\gamma,h}>0$ such that*
$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{c_{\gamma,h}}{\sqrt{N}}\exp\Big(N\mathrm{F}_{\frac{M}{N}}(h)\Big)\,. \tag{2.30}$$
2. *For $h<0$, if $\overline{\alpha}<1+\alpha/2$ we have*
$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{K_f(N)K_f(M)}{1-\exp(h)}\,. \tag{2.31}$$
Moreover, if $\overline{\alpha}>1+\alpha/2$
$$Z^f_{N,M,h} \overset{N\to\infty}{\sim} \frac{\exp(h)\big(\sum_{n\geq 0}K_f(n)\big)^2}{(1-\exp(h))^2}\mathrm{K}(N,M)\,. \tag{2.32}$$

Sharp estimates on the partition function lead to sharp control on path properties:

Theorem 2.7. *Choose $\gamma>0$ and consider the case $M\sim\gamma N$ and $\alpha>0$.*

1. *Let $(\mathcal{F}_1,\mathcal{F}_2):=\max\{\tau\cap[0,N]\times[0,M]\}$ be the last renewal epoch in $[0,N]\times[0,M]$. For $h<0$ and $\overline{\alpha}<1+\alpha/2$, the law of $(\mathcal{F}_1,\mathcal{F}_2)$ under $\mathbf{P}^f_{N,M,h}$ – a probability measure on $(\{0\}\cup\mathbb{N})^2$ – converges for $N\to\infty$ to the probability distribution that assigns to (i,j) probability*
$$(1-\exp(h))\mathbf{P}\left((i,j)\in\widetilde{\tau}_h\right)\,. \tag{2.33}$$
Set $\mathcal{L}_1:=N-\mathcal{F}_1$ and $\mathcal{L}_2:=M-\mathcal{F}_2$. For $h<0$ and $\overline{\alpha}>1+\alpha/2$, the law of $(\mathcal{L}_1,\mathcal{L}_2)$ under $\mathbf{P}^f_{N,M,h}$ – a probability measure on $(\{0\}\cup\mathbb{N})^2$ – converges for $N\to\infty$ to the probability distribution that assigns to (i,j) probability
$$\frac{1}{\big(\sum_{n\geq 0}K_f(n)\big)^2}K_f(i)K_f(j)\,. \tag{2.34}$$
Moreover, for $h<0$ and $\overline{\alpha}>1+\alpha/2$, we have
$$\lim_{L\to\infty}\lim_{N\to\infty}\mathbf{P}^f_{N,M,h}(\tau\cap[L,N-L]\times[L,M-L]=\varnothing)=1\,. \tag{2.35}$$

2. *For $h > 0$ and $\gamma \in (1/\gamma_c(h), \gamma_c(h))$ we have that both $\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)$ and $\partial_\gamma \mathrm{F}_\gamma(h)$ are positive and the law of $(\mathcal{L}_1, \mathcal{L}_2)$ under $\mathbf{P}^f_{N,M,h}$ – a probability measure on $(\{0\} \cup \mathbb{N})^2$ – converges for $N \to \infty$ to the probability distribution that assigns to (i, j) probability*

$$\frac{1}{C_{\gamma,h}} K_f(i) \exp\left(-i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)\right) K_f(j) \exp\left(-j\partial_\gamma \mathrm{F}_\gamma(h)\right) \tag{2.36}$$

with $C_{\gamma,h} > 0$ the normalization constant. Moreover the law of τ under $\mathbf{P}^f_{N,M,h}$ – a probability on the subsets of $(\{0\} \cup \mathbb{N})^2$ – converges in the same limit to the law of a positive recurrent two-dimensional renewal with inter-arrival law given by the function from $\mathbb{N}^2$ to $[0, 1)$

$$(i, j) \mapsto K(i + j) \exp\left(-i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right) - j\partial_\gamma \mathrm{F}_\gamma(h)\right) . \tag{2.37}$$

We observe that exploiting the symmetry of the model under the exchange of N and M, from the definition (2.12) of $\mathrm{F}_\gamma(h)$ we directly obtain $\mathrm{F}_\gamma(\cdot) = \gamma \mathrm{F}_{1/\gamma}(\cdot)$. The same symmetry can be appreciated in (2.37) if we make the replacement (i, j, γ) with $(j, i, 1/\gamma)$.

Theorem 2.7 can be summed up as:

1. In the delocalized phase, $h < 0$ (and $\overline{\alpha} \neq 1 + \alpha/2$, see below), there is no contact in the bulk of the system and, according to whether the $K_f(\cdot)$ exponent $\overline{\alpha}$ is larger or smaller than $1 + \alpha/2$ the two strands are free except for $O(1)$ contacts all close to the origin, or the two strands get detached after finitely many contacts (all close to the origin) and they meet again at a $O(1)$ distance from (N, M), terminating with two free ends of length $O(1)$. In the case $\overline{\alpha} = 1 + \alpha/2$ the slowly varying corrections $L(\cdot)$ and $\overline{L}(\cdot)$ matter and we leave out this rather cumbersome analysis.

2. In the localized phase ($h > 0$) and for γ in the Cramér region the process converges to a persistent renewal that we determine: this is similar to what happens in the one-dimensional case, but in this new set-up the limit process has the expression (2.37) which is much less straightforward than the corresponding one-dimensional case. A number of other results can be proven, in the spirit of the one-dimensional analogs (see [27] and [48, Ch. 2]), but we have chosen to limit ourselves to Theorem 2.7(2) and we signal that the proof of (2.35), see § 2.4.3, is much richer than (2.35).

2.1.8 Open issues and perspectives

We do not treat a number of natural issues: we list and discuss them here.

The non Cramér regime

For $h > 0$ and $\gamma \notin (1/\gamma_c(h), \gamma_c(h))$ we do not give sharp estimates. To our knowledge sharp estimates on the renewal function in this regime are for the moment not available (the most advanced reference available appears to be [22]). The issue is not a secondary one: it is at the heart of understanding the transitions and the different phases that one observes in the localized regime. And what one expects is rather clear: for γ in the Cramér region we have seen that the free ends are microscopic, i.e. $O(1)$, and the limit process is just a recurrent renewal; for γ in the interior of the complementary of the Cramér region, and for the free case, instead a big loop should appear (showing that there is only one) or the free ends, probably only one, should become macroscopic; which of these two phenomenologies prevail should depend on the exponents α and $\overline{\alpha}$. The analysis is certainly different for the constrained case, because the expected big loop can only be along the chain (and, by exchangeability, its location is going to be uniformly distributed along the chain). This and very similar issues are widely discussed in the physical literature ([61, 72]) and the analogy with Bose condensation is regularly invoked, but the analysis is far from being rigorous.

Counting the transitions in the localized regime

We present examples with zero, one or two transitions. Can there be more than two? Are they always finitely many?

Sharp estimates at criticality

If $\sum_n n^3 K(n) < \infty$ sharp estimates for $h = 0$ are covered by (2.28). We have not treated this case here because it would be natural to consider the complete spectrum of loop exponents, but we meet again with the limitations of multivariate renewal theory. The gPS model demands control only on the special class of renewals with inter-arrivals $\mathrm{K}(n,m) = K(n+m)$ and we hope that an ad-hoc treatment will lead to progress. And of course the gPS model is one more motivation for a more systematic study of multivariate renewals.

Disordered interactions

Here the issues are several: we stick to the one of disorder relevance at criticality (see the review of the literature in [50, Ch. 4]), but there are most probably intriguing questions also away from criticality (in analogy with [55]). The effect of disorder on the critical point is directly related to obtaining sharp estimates on the renewal function of the underlying renewal process, at least if the disorder is introduced via an IID family $\{h_{n,m}\}_{(n,m)\in\mathbb{N}^2}$ of random variables. In fact the tools developed for the basic disordered PS model ([50] and references therein) can be applied, but the problem is that sharp estimates are available

only for the very particular case of $\sum_n n^3 K(n) < \infty$. And there is the issue that such an IID disorder is not the most suited for DNA modeling, but if the aim is understanding the effect of noise on critical behaviors this way of introducing the disorder is certainly acceptable (and it is what has been done also in the bio-physical literature, even sticking to the DNA/RNA set-up! See for example [24]). A more natural disorder is however obtained by assigning to each strand a sequence of, possibly IID, potentials $\{h_j^{(1)}\}_{j\in\mathbb{N}}$ and $\{h_j^{(2)}\}_{j\in\mathbb{N}}$ – one can imagine the case in which the two sequences are independent or the case in which they are (strongly) correlated – and $h_{n,m} = h_n^{(1)} h_m^{(2)}$: note that correlations are introduced with this product choice even if the two sequences are independent. This appears to be a very challenging model (see [12] and references therein for the issues that arise when correlations are introduced in the disorder sequence for the basic PS model).

Related models

The gPS model is intimately related to the more complex RNA models for secondary structure: [41], where the vast literature is cited, is particularly interesting for us because RNA models are linked with the gPS model. Models for circular DNA [95] are also very much related to gPS, as pointed out for example in [46]. In [61] the authors focus on an issue (existence of one macroscopic loop) for circular DNA that is precisely the one that we face outside of the Cramér regime. Finally, the gPS model can be seen as a toy model for interacting self-avoiding walks. In the related direction of simplified models for a self-interacting self-avoiding walk we signal the Zwanzig-Lauritzen model that has been tackled first by generating functions techniques (e.g. [23]) and recently by probabilistic methods in [29, 73].

2.1.9 Organization of the paper

In Section 2.2, we present the results on Large and Sharp Deviations for bivariate renewal processes that we use. In Section 2.3, we introduce the constrained model and study the free energy in the localized and delocalized regime proving Theorem 2.2, Lemma 2.4 and Theorem 2.5. At the end of this section, we work out explicitly some examples. Finally in Section 2.4, we prove Proposition 2.1 and we compute the sharp estimates of the free partition function and the path properties proving Theorem 2.6 and Theorem 2.7. For the Dominated Converge Theorem we use the shortcut (DOM).

2.2 Large Deviations and Local Limit Theorems for bivariate renewals

We now give a number of results on the renewal $\widetilde{\tau}_h$, $h > 0$, defined at the beginning of Section 2.1.6. As it will be clear, they follow directly from various results that one can find in [20, 21] where a more general case is treated (starting from the fact that we limit ourselves to the two-dimensional case). Let us start by introducing the exponential moment generating function of $\xi := (\widetilde{\tau}_h)_1$ (we recall: $\mathbf{P}(\xi = (n,m)) = \widetilde{K}_h(n,m)$)

$$q_h(\lambda) := \mathbf{E}\left[\exp\left(\langle \lambda, \xi \rangle\right)\right], \tag{2.38}$$

where $\lambda \in \mathbb{R}^2$, $\langle \lambda, \xi \rangle = \lambda_1\xi_1 + \lambda_2\xi_2$. From the definition of $\widetilde{K}_h(\cdot,\cdot)$ one readily sees that $q_h(\lambda) < \infty$ if and only if both $\lambda_1 \leq G$ and $\lambda_2 \leq G$. The Large Deviation function $\Lambda(\cdot)$ corresponding to the random vector ξ is the Legendre transform of the function $\log q_h(\cdot)$

$$\Lambda(\theta) = \sup_{\lambda} \left\{ \langle \lambda, \theta \rangle - \log q_h(\lambda) \right\}, \tag{2.39}$$

where $\theta \in \mathbb{R}^2$. Since we are after the renewal function of $\widetilde{\tau}_h$ the Large Deviation function of the inter-arrival random vector is just an intermediate step. The asymptotic behavior of the renewal function is directly related to the so-called *second deviation function*, introduced and investigated in [20]:

$$D_h(\theta) = D(\theta) = \inf_{s>0} \frac{\Lambda(s\theta)}{s}, \tag{2.40}$$

where $\theta \in \mathbb{R}^2$. The notation with the subscript h will be useful further on to remind the dependence on the parameter but at this stage it is rather superfluous.

Remark 2.8. *From (2.40) one can see that $D(s\theta) = sD(\theta)$ for every $s \geq 0$. In [20, pp. 652-653] a detailed analysis of $D(\cdot)$ is given, notably the fact that it is convex: for $p \geq 0, q \geq 0, p+q = 1$ and $\theta, \eta \in \mathbb{R}^2$,*

$$D(p\theta + q\eta) \leq pD(\theta) + qD(\eta). \tag{2.41}$$

We can immediately deduce from these properties that for every $\theta, \eta \in \mathbb{R}^2$, we have

$$D(\theta + \eta) \leq D(\theta) + D(\eta). \tag{2.42}$$

Here is an important step:

Proposition 2.9. *[[20], Theorem 1]. For every $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$*

$$D(\theta) = \sup_{\lambda \in A} \langle \lambda, \theta \rangle = \sup_{\lambda \in \partial A} \langle \lambda, \theta \rangle, \tag{2.43}$$

where A is the closed convex set $\left\{\lambda \in \mathbb{R}^2 : q_h(\lambda) \leq 1\right\}$ and ∂A is the boundary of A.

It is now practical to focus on the specific case we are considering, notably the fact that $D(\theta) = \infty$ if θ is not in the first quadrant is an intuitive consequence of the fact that our process has increments that have positive components and can be read out of the structure of A. Let us make A more explicit

$$A = \left\{\lambda \in \mathbb{R}^2 : q_h(\lambda_1, \lambda_2) \le 1\right\}, \tag{2.44}$$

where $q_h(\cdot)$, defined in (2.38), corresponds to (2.21) in our special case and we recall that $h > 0$ and $G = G(h) > 0$ is chosen so that $(0,0) \in \partial A$. Note that $q_h(\lambda_1, \lambda_2) = q_h(\lambda_2, \lambda_1)$ and that $q_h(\cdot)$ is convex (this of course implies the convexity of A) and it is symmetric with respect to the diagonal of the first and third quadrant.

Lemma 2.10. *We have*

$$A \subset \left\{\lambda \in \mathbb{R}^2 : \lambda_1 \le G \text{ and } \lambda_2 \le G\right\} \cap \left\{\lambda \in \mathbb{R}^2 : \lambda_2 \le -\lambda_1\right\}. \tag{2.45}$$

Moreover $\overline{\lambda}_1 \overset{(2.22)}{=} \sup\{\lambda_1 < 0 : q_h(\lambda_1, G) = 1\} < -G$ *and the equation* $q_h(\lambda_1, \lambda_2) = 1$ *is uniquely solvable for* $(\lambda_1, \lambda_2) \in [\overline{\lambda}_1, G]^2$*, defining the curve* $\mathcal{W}_h$*, symmetric with respect to the diagonal of the first and third quadrant.* $\mathcal{W}_h$ *is the graph of a concave and decreasing function* $\widetilde{\lambda}_2 : [\overline{\lambda}_1, G] \to [\overline{\lambda}_1, G]$ *which satisfies* $\widetilde{\lambda}_2(\overline{\lambda}_1) = G$*,* $\widetilde{\lambda}_2(G) = \overline{\lambda}_1$ *and* $\widetilde{\lambda}_2(0) = 0$*. Moreover* $\widetilde{\lambda}_2$ *is analytic in the interior of its domain and*

$$\partial A = \left\{(\lambda_1, G) : \lambda_1 < \overline{\lambda}_1\right\} \cup \left\{(G, \lambda_2) : \lambda_2 < \overline{\lambda}_1\right\} \cup \mathcal{W}_h. \tag{2.46}$$

Finally, $q_h(\lambda_1, \lambda_2) < 1$ *for* $(\lambda_1, \lambda_2) \in \partial A \setminus \mathcal{W}_h$*.*

Proof. For this proof it is practical to keep at hand part (left) of Figure 2.3. The fact that $A \subset \left\{\lambda \in \mathbb{R}^2 : \lambda_1 \le G \text{ and } \lambda_2 \le G\right\}$ is just the fact that $q_h(\lambda_1, \lambda_2) = \infty$ if $\lambda_1 \wedge \lambda_2 > G$. On the other hand this last observation, coupled with $q_h(\lambda_1, \lambda_2) = q_h(\lambda_2, \lambda_1)$, $q_h(0,0) = 1$ and convexity of A, tells us that A does not go above the line $\lambda_2 = -\lambda_1$ and (2.45) is established. Now, since $q_h(\cdot,\cdot)$ is a separately non-decreasing function (and even increasing where it is bounded), ∂A contains $\{\lambda : q_h(\lambda) = 1\}$. So the issue is the solvability of $q_h(\lambda) = 1$ and it is straightforward to see that $q_h(\lambda) = 1$ has a (unique, by monotonicity) solution if $\lambda_1 \in [\overline{\lambda}_1, G]$ and this way we define a function $\widetilde{\lambda}_2(\cdot) : [\overline{\lambda}_1, G] \mapsto [\overline{\lambda}_1, G]$ (note that at this stage it is already clear that $\overline{\lambda}_1 \le -G(h)$) which is analytic in the interior of its domain, by the analytic Implicit Function Theorem. It is actually immediate to check that this function is not linear (for example, compute the second derivative at the origin), so $\overline{\lambda}_1 < -G(h)$. Finally $q_h(\lambda_1, G) < 1$ for $\lambda_1 < \overline{\lambda}_1$ and, by symmetry, $q_h(G, \lambda_2) < 1$ for $\lambda_2 < \overline{\lambda}_1$. The proof is therefore complete. □

Remark 2.11. *Lemma 2.10 is given for fixed $h > 0$, but $\overline{\lambda}_1$ depends also on h and when we need to make this dependence explicit we write $\overline{\lambda}_1(h)$. The same is true for the function $\widetilde{\lambda}_2(\cdot)$ and we write $\widetilde{\lambda}_{2,h}(\cdot)$. Note that the analyticity of $\overline{\lambda}_1(\cdot)$ on the positive semi-axis is a direct consequence of the Analytic Implicit Function Theorem ([66, Sec. 2.3], and it is just a matter of analyticity in one variable). If instead we consider the function $(\lambda_1, h) \mapsto \widetilde{\lambda}_{2,h}(\lambda_1)$, with $\widetilde{\lambda}_{2,h}(\lambda_1)$ which is obtained by solving for λ_2 the equation $q_h(\lambda_1, \lambda_2) = 1$ and we have seen that this requires $\lambda_1 \in [\overline{\lambda}_1(h), \mathrm{G}(h)]$. Therefore, by the Analytic Implicit Function Theorem [66, Sec. 2.3], the function $(\lambda_1, h) \mapsto \widetilde{\lambda}_{2,h}(\lambda_1)$ is analytic (in two variables this time) in the domain $h > 0$ and $\lambda_1 \in (\overline{\lambda}_1(h), \mathrm{G}(h))$.*

From Proposition 2.9 and Lemma 2.10 we can derive a number of consequences, like the fact that $D(\theta) < \infty$ for every θ in the first quadrant (that here includes the two axes) and that, when $D(\theta) < \infty$, that is in the first quadrant, there are only two possibilities: either the supremum in the rightmost term in (2.43) is reached in the interior of $\mathcal{W}_h$ or *at the boundary*, that is in $\{(\overline{\lambda}_1, \mathrm{G}), (\mathrm{G}, \overline{\lambda}_1)\}$. We observe by direct inspection that if $\theta_2 > \theta_1$ then if the supremum is not achieved in the interior, then it is achieved at $(\overline{\lambda}_1, \mathrm{G})$. Moreover if $\theta_1 = \theta_2 > 0$ the supremum is always achieved in the interior and, more precisely, at $(0, 0)$ and therefore $D(\theta_1, \theta_1) = 0$. This induces a partition of the first quadrant: $\theta \in E_h$ if the supremum is achieved in the interior and $\theta \in E_h^{\mathsf{C}}$ if it is achieved at the boundary. It is also useful to remark that E_h is an open sector ([20, p. 653], or it can be seen directly in our specific set-up): in our case it is also symmetric with respect to the diagonal of the first quadrant, that is there exists $\varphi \in (0, \pi/4)$ such that $E_h = \{\theta \in [0, \infty)^2 : \theta_2/\theta_1 \in (\tan(-\varphi + \pi/4), \tan(\varphi + \pi/4))\}$.

Remark 2.12. If $\theta \in E_h$ then [20, Theorem 2] tells us that the infimum in (2.40) is attained at a unique point $s(\theta)$, so $D(\theta) = \Lambda(\theta s(\theta))/s(\theta)$, and $\Lambda(\cdot)$ is analytic and strictly convex at $s(\theta)$. This is what we may call the *Cramér region* of parameters: such a region being the set of θ's for which the large deviations trajectories can be made typical by a suitable change of measure (*tilting*).

We are now ready to state the result that links $D(\cdot)$ and the renewal function. Once again it is obtained by restricting to our context a result – [20, Theorem 5] – see however Remark 2.14 below. We employ the notation $\Lambda''(\theta)$ for the Hessian matrix of Λ.

Proposition 2.13. *For every $v = (v_1, v_2) = (N, M)/t \in E_h \cap \{\theta : |\theta| = 1\}$ with $t = \sqrt{N^2 + M^2}$, the following representation holds:*

$$\mathbf{P}\left((N, M) \in \widetilde{\tau}_h\right) = \frac{1}{t^{1/2}} \left(\sqrt{\frac{\det\left(\Lambda''(s(v)v)\right)}{2\pi s(v) \left\langle v, \Lambda''(s(v)v)v \right\rangle}} + \varepsilon(N, M) \right) \exp\left(-tD(v)\right) , \qquad (2.47)$$

and for every compact $C \subseteq E_h \cap \{\theta : |\theta| = 1\}$

$$\lim_{t\to\infty} \sup_{\substack{N,M:\sqrt{N^2+M^2}=t \\ v\in C}} \varepsilon(N,M) = 0\,. \tag{2.48}$$

Recalling (2.17), we see that Proposition 2.13 implies the sharp estimate

$$Z^c_{N,M,h} = \frac{A(M/N)+\widetilde{\varepsilon}(N,M)}{\sqrt{N}} \exp\left((N+M)\mathrm{G}(h) - ND_h(1,M/N)\right)\,, \tag{2.49}$$

where $A(M/N)$ is equal to $(1+(M/N)^2)^{-1/4}$ times the square root term in (2.47) (which depends only on v, which in turn is just a function of M/N) and $\widetilde{\varepsilon}(N,M) = (1+(M/N)^2)^{-1/4}\varepsilon(N,M)$.

Remark 2.14. In [20] the factor $s(v)$ that appears just after 2π in (2.47) has the exponent $d+3 = 5$. This formula appears also in [21, p. 11], with an additional oversight. The formula we give is in agreement with [39] who covers only the case $v \propto \mu_h$ (cf. (2.20)), but in greater generality than [20, 21]. Formula (2.47) a priori requires some exponential decay of the inter-arrival law to make sure that the Hessian is computed at an analyticity point of $\Lambda(\cdot)$. Actually, the analysis in [39] shows that this is not necessary and (2.47) still holds true for $h = 0$ if $\sum_{n,m\geq 1}(n+m)^2K(n+m) = \sum_{t\geq 2} t^2(t-1)K(t) < \infty$. With the help of the notation in [39] we remark that (2.47) can be made slightly more readable if we observe that, as it is well known, the inverse $B(\theta)$ of $\Lambda''(\theta)$ is the covariance matrix of the *tilted* random vector $X = (X^1, X^2)$ with $\mathbf{P}(X = (n,m)) = \widetilde{K}_h(n,m)\exp(n\lambda_1(\theta)+m\lambda_2(\theta))/C_\theta$ and $C_\theta = \sum_{n,m}\widetilde{K}_h(n,m)\exp(n\lambda_1(\theta)+m\lambda_2(\theta))$ and $\lambda(\theta)$ is the optimal point of (2.39). Therefore with $v = \theta/|\theta|$

$$\begin{aligned}\frac{\det\left(\Lambda''(\theta)\right)}{\langle v, \Lambda''(\theta)v\rangle} &= \frac{1}{\det\left(B(\theta)\right)\langle v, (B(\theta))^{-1}v\rangle} \\ &= \left\langle v, \begin{pmatrix} \mathbf{E}[(X^2)^2]-\mathbf{E}[X^2]^2 & -\mathbf{E}[X^1X^2]+\mathbf{E}[X^1]\mathbf{E}[X^2] \\ -\mathbf{E}[X^1X^2]+\mathbf{E}[X^1]\mathbf{E}[X^2] & \mathbf{E}[(X^1)^2]-\mathbf{E}[X^1]^2 \end{pmatrix} v\right\rangle^{-1}\,.\end{aligned} \tag{2.50}$$

A Local Limit Theorem, analogous to Proposition 2.13, for $v \in E_h^{\mathsf{C}}$ is available at the moment (see [22, Theorem 2.1]) only if the entries of $\Lambda''(v)$ are all finite, and this is not always the case in our set-up, notably it is not the case if the exponent α entering the definition of $K(\cdot)$ is smaller than two. Nevertheless, the following weaker result will suffice for our purposes:

Proposition 2.15. *For every* θ

$$\limsup_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(\lceil t\theta\rceil \in \widetilde{\tau}_h\right) \leq -D\left(\theta\right)\,, \tag{2.51}$$

and if $\theta \in E_h^{\mathsf{C}}$, *with* $\theta_1 > 0$ *and* $\theta_2 > 0$, *then*

$$\lim_{t\to\infty} \frac{1}{t}\log \mathbf{P}\left(\lceil t\theta\rceil \in \widetilde{\tau}_h\right) = -D\left(\theta\right)\,. \tag{2.52}$$

Of course (2.52) holds also for $\theta \in E_h$ as an immediate consequence of Proposition 2.13. In Proposition 2.15, we have used the notation $\lceil (x,y) \rceil = (\lceil x \rceil, \lceil y \rceil)$.

Proof. The upper bound is a direct consequence of the Large Deviations Principle [20, Theorem 4]):

$$\lim_{\varepsilon \searrow 0} \lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(t\{v \in \mathbb{R}^2 : |v-\theta| \le \varepsilon\} \cap \widetilde{\tau}_h \neq \varnothing \right) = -D(\theta) \,. \tag{2.53}$$

For the lower bound we assume without loss of generality that $\theta_2 > \theta_1$ and we observe that if $\theta \in E_h^{\mathsf{C}}$ – a closed set (recall Remark 2.12 and the explanation that precedes it) – then either θ is in the boundary or in the interior of E_h^{C}. If it is in the interior then there exists $\theta_2^\star < \theta_2$ with $(\theta_1, \theta_2^\star)$ in the boundary of E_h^{C}, so that for every $\varepsilon > 0$ small we have $(\theta_1, \theta_2^\star - \varepsilon) \in E_h$. If θ is in the boundary of E_h^{C} we directly set $\theta_2^\star = \theta_2$. By Proposition 2.13

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left((\lceil t\theta_1 \rceil - 1, \lceil t(\theta_2^\star - \varepsilon) \rceil) \in \widetilde{\tau}_h \right) = -D_h\left((\theta_1, \theta_2^\star - \varepsilon) \right) . \tag{2.54}$$

But

$$\mathbf{P}\left(\lceil t\theta \rceil \in \widetilde{\tau}_h \right) \ge \mathbf{P}\left((\lceil t\theta_1 \rceil - 1, \lceil t(\theta_2^\star - \varepsilon) \rceil) \in \widetilde{\tau}_h \right) \widetilde{K}_h\left(1, \lceil t\theta_2 \rceil - \lceil t(\theta_2^\star - \varepsilon) \rceil \right) , \tag{2.55}$$

and therefore we have

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(\lceil t\theta \rceil \in \widetilde{\tau}_h \right) \ge -D_h\left((\theta_1, \theta_2^\star - \varepsilon) \right) - \mathrm{G}(h)\left(\theta_2 - \theta_2^\star + \varepsilon \right) , \tag{2.56}$$

and, by continuity of $D_h(\cdot)$, we get to

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbf{P}\left(\lceil t\theta \rceil \in \widetilde{\tau}_h \right) \ge -D_h\left((\theta_1, \theta_2^\star) \right) - \mathrm{G}(h)\left(\theta_2 - \theta_2^\star \right) . \tag{2.57}$$

Since $(\theta_1, \theta_2^\star) \in E_h^{\mathsf{C}}$ we have that

$$D_h\left((\theta_1, \theta_2^\star) \right) + \mathrm{G}(h)\left(\theta_2 - \theta_2^\star \right) = \overline{\lambda}_1 \theta_1 + \mathrm{G}(h)\theta_2^\star + \mathrm{G}(h)\left(\theta_2 - \theta_2^\star \right) = \overline{\lambda}_1 \theta_1 + \mathrm{G}(h)\theta_2 \,, \tag{2.58}$$

and since $\theta \in E_h^{\mathsf{C}}$ the rightmost term is $D_h(\theta)$ and we are done. □

2.3 The constrained model

2.3.1 The free energy of the constrained model

We assume $M \ge N$. The first result we present is:

Proposition 2.16. *For every $\gamma \ge 1$*

$$\widetilde{\mathrm{F}}_\gamma(h) := \lim_{\substack{N,M\to\infty: \\ \frac{M}{N}\to\gamma}} \frac{1}{N} \log Z^c_{N,M,h} = \begin{cases} 0 & \text{if } h \le 0 \,, \\ (1+\gamma)\mathrm{G}(h) - D_h(1,\gamma) & \text{if } h > 0 \,, \end{cases} \tag{2.59}$$

and in fact we have also that for every $L > 1$

$$\lim_{\varepsilon \searrow 0} \lim_{N\to\infty} \sup_{\gamma\in[1,L]} \sup_{M:|(M/N)-\gamma|\le\varepsilon} \left| \frac{1}{N} \log Z^c_{N,M,h} - \widetilde{\mathrm{F}}_\gamma(h) \right| = 0 . \tag{2.60}$$

Moreover we have

$$D_h(1,\gamma) = \max_{\lambda\in B_h} (\lambda_1 + \gamma\lambda_2) , \tag{2.61}$$

with

$$B_h = \left\{ \lambda : \overline{\lambda}_1 \le \lambda_1 \le 0,\ 0 \le \lambda_2 \le \mathrm{G} , \sum_{n,m} \widetilde{K}_h(n,m) \exp(\lambda_1 n + \lambda_2 m) = 1 \right\} . \tag{2.62}$$

Finally $\widetilde{\mathrm{F}}_\gamma(h) > 0$ *for* $h > 0$ *and in fact*

$$2\mathrm{G}(h) \le \widetilde{\mathrm{F}}_\gamma(h) \le (1+\gamma)\mathrm{G}(h) . \tag{2.63}$$

Proof. Let us start by observing that

$$\mathrm{G}(h) = \lim_{N\to\infty} \frac{1}{2N} \log Z^c_{N,N,h} = \frac{1}{2}\widetilde{\mathrm{F}}_1(h) . \tag{2.64}$$

This can be seen, for $h \le 0$, by the elementary argument presented just after (2.14) and, for $h > 0$, by (2.17) and Proposition 2.13: in fact $D(1,1) = 0$. It is slightly more practical to establish first a result which is a priori weaker than (2.59). Consider first the limit for $N \to \infty$ and $M = \lfloor \gamma N \rfloor$ and let us establish the rightmost equality in (2.59) with this notion of limit: call $\breve{\mathrm{F}}_\gamma(h)$ this expression. Again, the case $h \le 0$ of (2.59) is treated by elementary methods just after the statement of Proposition 2.1. For the case $h > 0$ we observe that by (2.17) it suffices to show that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbf{P}\left((N, \lfloor \gamma N \rfloor) \in \widetilde{\tau}_h \right) = -D_h(1,\gamma) . \tag{2.65}$$

But (2.65) is a direct consequence of Proposition 2.13 and Proposition 2.15, so the weaker version of (2.59) is established. Now we step to (2.60): if we establish (2.60) with $\widetilde{\mathrm{F}}_\gamma(h)$ replaced by $\breve{\mathrm{F}}_\gamma(h)$, then the original version of (2.59) holds, which implies that $\breve{\mathrm{F}}_\gamma(h) = \widetilde{\mathrm{F}}_\gamma(h)$ and, in turn, that also the original version of (2.60) holds. It is therefore a matter of comparing (uniformly) the limits along all sequences of (M,N), with $|(M/N) - \gamma| \le \varepsilon$, with the special case $M = \lfloor \gamma N \rfloor$. But this is obtained for example by exploiting that if $M' > M$ we have that for any $C_1 > 2 + \alpha$ there exists $C_2 > 0$ such that for every $N \ge 1$

$$Z^c_{N,M,h} \le C_2 {M'}^{C_1} Z^c_{N,M',h} . \tag{2.66}$$

(2.66) can be established by using $K(n)/K(n+m) \le 1/K(n+m) \le C_2 m^{C_1}$ for $m \ge n$. More precisely we apply this inequality to bound $1/K(l_n + t_n + M' - M)$ (using the fact that $l_n + t_n + M' - M \le N + M' \le (1 + 1/\gamma)M'$) in (2.9) to stretch the last renewal so that it matches the boundary constraint (N, M'), and then we allow n to go up to $N \wedge M'$,

which in this case is N anyways, and the new constraint on $\underline{t}$ is $|\underline{t}| = M'$. Inequality (2.66) can then be used to sandwich the partition functions $Z^c_{N,M,h}$, with $|(M/N) - \gamma| \le \varepsilon$ and N sufficiently large, between $Z^c_{N,\lfloor(\gamma-2\varepsilon)N\rfloor,h}/N^{2C_1}$ and $Z^c_{N,\lfloor(\gamma+2\varepsilon)N\rfloor,h}N^{2C_1}$. The continuity of $\gamma \mapsto \breve{\mathrm{F}}_\gamma(h)$, and hence the uniform continuity and boundedness on compact sets – in our case $[1, L]$ – completes the argument and the proof of (2.59) and (2.60).

Let us check (2.61): it is of course a matter of replacing ∂A in (2.43), Proposition 2.9, by B_h. Much of the work has been done in Lemma 2.10: we are just left with showing that we can restrict the supremum to B_h. First of all the symmetry of ∂A tells us that $\gamma \ge 1$ implies that $\sup_{\partial A}(\lambda_1 + \gamma\lambda_2)$ does not change if we restrict ∂A to $\lambda_1 \le 0$. Morevover, by Lemma 2.10, if $\lambda_1 < \overline{\lambda}_1$, then $(\lambda_1, \mathrm{G}) \in \partial A$, but since $\lambda_1 + \gamma\mathrm{G} < \overline{\lambda}_1 + \gamma\mathrm{G}$ for $\lambda_1 < \overline{\lambda}_1$ we can actually neglect these points in taking the supremum.

We are left with the positivity of $\widetilde{\mathrm{F}}_\gamma(h)$ for $h > 0$. This follows directly by observing that

$$Z^c_{N,M,h} \ge Z^c_{N-1,N-1,h} \exp(h) K(M - N + 2)\,, \tag{2.67}$$

which implies more than the positivity, that is $\widetilde{\mathrm{F}}_\gamma(h) \ge 2\mathrm{G}(h)$. Finally, by exploiting also that $D_h(\cdot) \ge 0$ (for $h > 0$) and that in any case $\widetilde{\mathrm{F}}_\gamma(h) = 0$ for $h \le 0$, we see that (2.63) holds and the proof of Proposition 2.16 is therefore complete. □

We can go beyond Proposition 2.16 by exploiting the variational problem (2.61)-(2.62). Note that since $\widetilde{\mathrm{F}}_1(h) = 2\mathrm{G}(h)$ we know that $\widetilde{\mathrm{F}}_1(\cdot)$ is analytic except at the origin, but for $\gamma > 0$ the situation is more involved.

Proposition 2.17. *The function $\gamma_c : (0, \infty) \to (1, \infty)$, defined in (2.23), is real analytic. Moreover $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is analytic on the positive semi-axis out of the set $\{h : \gamma_c(h) - \gamma = 0\}$, but $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is not analytic at the values h at which $\gamma_c(h) - \gamma$ changes sign.*

Remark 2.18. *Of course the regularity issue is not completely resolved by Proposition 2.17, both because it does not make clear whether or not all the points in the discrete set $\{h : \gamma_c(h) - \gamma = 0\}$ are non-analyticity points and because it does not specify the type of singularities.*

Proof. Let us first prove the first statement (which is also the first statement in Lemma 2.4). Recall the function $q_h(\cdot)$ and $\overline{\lambda}_1$ from (2.21) and (2.22). Recall also the definition of $\lambda_1 \mapsto \widetilde{\lambda}_2(\lambda_1)$ in Lemma 2.10: $\widetilde{\lambda}_2(\lambda_1)$ is the only solution in λ_2 to $q_h(\lambda_1, \lambda_2) = 1$, for $\lambda_1 \in [\overline{\lambda}_1, \mathrm{G}]$ (see Figure 2.3). Recall also that $\widetilde{\lambda}_2(\cdot)$ is concave – i.e. concave down – and analytic in $(\overline{\lambda}_1, 0)$. We have that $\widetilde{\lambda}_2'(\lambda_1)$ equals $-\partial_{\lambda_1} q_h(\lambda_1, \lambda_2)/\partial_{\lambda_2} q_h(\lambda_1, \lambda_2)$ evaluated at $\lambda_2 = \widetilde{\lambda}_2(\lambda_1)$. Therefore

$$\lim_{\lambda_1 \searrow \overline{\lambda}_1} \widetilde{\lambda}_2'(\lambda_1) = \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right) = -\frac{\sum_{n,m} nK(n+m)\exp\left(h - n\left(\mathrm{G}(h) + |\overline{\lambda}_1|\right)\right)}{\sum_{n,m} mK(n+m)\exp\left(h - n\left(\mathrm{G}(h) + |\overline{\lambda}_1|\right)\right)} \in (-1, 0)\,. \tag{2.68}$$

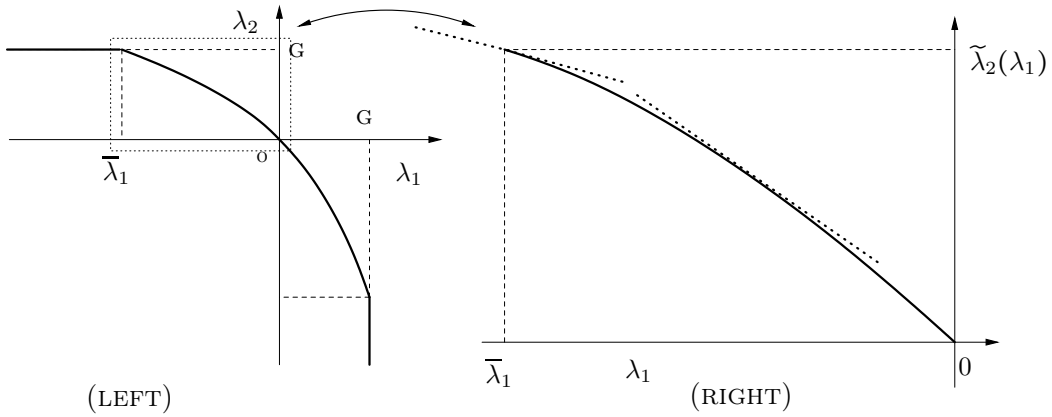

FIGURE 2.3. In part (LEFT) ∂A, with $A = \{\lambda : q_h(\lambda_1, \lambda_2) \le 1\}$, is represented by the thick line. Notice the symmetries and the convexity of A. In part (RIGHT) we zoom into the relevant part for our analysis.

The denominator is bounded because of (2.24) and the fact that the ratio is bounded below by -1 is just a consequence of concavity and the fact that $\widetilde{\lambda}_2'(0) = -1$. Moreover $\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ is a function of h: with the notations in Remark 2.11 we write rather $\widetilde{\lambda}_{2,h}'\left(\overline{\lambda}_1(h)^+\right)$ and a look at the right-hand side of (2.68), recalling that $\overline{\lambda}_1(\cdot)$ is analytic (cf. Remark 2.11), suffices to see that $h \mapsto \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ is analytic. Since $\gamma_c(h) = -1/\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$, the properties of $\gamma_c(\cdot)$ claimed in the statement are proven.

We are now at the heart of the argument: since the function to be maximized, $\lambda_1 + \gamma\lambda_2$ is constant on lines with slope $-1/\gamma$, the maximum is achieved

1. at a value $\lambda_1 \in (\overline{\lambda}_1, 0)$ if $1/\gamma > -\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$: the value of λ_1 is of course found by solving $\widetilde{\lambda}_2'(\lambda_1) = -1/\gamma$: this is what we call the Cramér regime (we have a tilted measure that makes typical the Large Deviation event that the renewal follows the slope γ);
2. at $\lambda_1 = \overline{\lambda}_1$ if $1/\gamma \le -\widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$: we are outside of the Cramér regime.

The situation is therefore that varying h one may switch in and out of the Cramér regime. Out of the Cramér regime the free energy is actually equal to $\mathrm{G}(h) - \overline{\lambda}_1(h)$, and the function $h \mapsto \mathrm{G}(h) - \overline{\lambda}_1(h)$ is analytic for every $h > 0$ (but it is not necessarily equal to the free energy $\widetilde{\mathrm{F}}_\gamma(h)$!). In the Cramér regime the free energy is strictly smaller than $\mathrm{G}(h) - \overline{\lambda}_1(h)$, just because the maximum of $\lambda_1 + \gamma\lambda_2$ is achieved on the boundary and $D_h(1, \gamma)$ contributes to the free energy with a negative sign, cf. (2.59). This explains why the changes of sign of $h \mapsto 1/\gamma + \widetilde{\lambda}_2'\left(\overline{\lambda}_1^+\right)$ are non-analyticity points. □

2.3.2 Analysis of the free energy singularities in the localized regime

We now go deeper into the analysis of the singularities for the non-analyticity points we have found, that is the values of $h > 0$ for which $\gamma_c(h) - \gamma$ (cf. (2.23)) changes sign. At the

same time we tackle also the case in which $\gamma_c(h) - \gamma$ hits zero without crossing it.

For this and referring to the list of two items of the previous proof we introduce some notations.

1. In the Cramér regime we introduce the notation $\big(\hat\lambda_1(h,\gamma), \hat\lambda_2(h,\gamma)\big) \in B_h$ for the optimal point (for the sake of brevity, we omit the dependence on γ). It is of course found by solving $\widetilde\Lambda_2'(\lambda_1) = -1/\gamma$ that yields $\hat\lambda_1(h)$ and $\hat\lambda_2(h) = \widetilde\Lambda_2\big(\hat\lambda_1(h)\big)$. Note that the values of h corresponding to the Cramér regime is an union of open disjoint intervals: we call these intervals $\mathcal{I}_1, \mathcal{I}_2, \dots$. The analyticity of $\hat\lambda_1(\cdot)$ and $\hat\lambda_2(\cdot)$ in the intervals $\mathcal{I}_j$ follows by Remark 2.11 and by repeating the arguments in the same remark. We also introduce

$$\hat c_\gamma(h) := \big(\mathrm{G}(h) - \hat\lambda_1(h)\big) + \gamma\big(\mathrm{G}(h) - \hat\lambda_2(h)\big)\,, \tag{2.69}$$

for $h \in \cup_j \mathcal{I}_j$ and of course $\widetilde{\mathrm{F}}_\gamma(h) = \hat c_\gamma(h)$ on this set. Note that $\cup_j \mathcal{I}_j$ can be alternatively characterized as $\{h : \gamma_c(h) - \gamma > 0\}$.

2. Out of the Cramér regime, that is for $h \in \mathcal{I}_{\mathrm{b}} := (0,\infty) \setminus \cup_j \mathcal{I}_j$, we introduce instead

$$\mathrm{N}(h) := \mathrm{G}(h) - \overline{\lambda}_1(h)\,, \tag{2.70}$$

and $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{N}(h)$ if and only if $h \in \mathcal{I}_{\mathrm{b}}$. Note that $\mathcal{I}_{\mathrm{b}}$ can be alternatively characterized as $\{h : \gamma_c(h) - \gamma \le 0\}$. Note also that (cf. the end of the proof of Proposition 2.17) $\mathrm{N}(h) > \hat c_\gamma(h)$ in the interior on the intervals on which $\hat c_\gamma(\cdot)$ is defined.

Observe moreover that $\partial \cup_j \mathcal{I}_j = \partial \mathcal{I}_{\mathrm{b}}$, here ∂A denotes the boundary of A seen as a subset of $(0,\infty)$, and if $h \in \partial\mathcal{I}_{\mathrm{b}}$ we have $\hat c_\gamma(h) = \mathrm{N}(h)$ and $\gamma_c(h) = \gamma$. By differentiating (2.23) we obtain that if $\gamma_c(h) = \gamma$

$$\gamma_c'(h) = -\frac{\mathrm{N}'(h)}{\sum_{n,m} nK(n+m)\exp(-\mathrm{N}(h)n)} \sum_{n,m} n(m-\gamma n)K(n+m)\exp(-n\mathrm{N}(h))\,, \tag{2.71}$$

hence (with the convention $\mathrm{sign}(0) = 0$)

$$\mathrm{sign}\big(\gamma_c'(h)\big) = -\,\mathrm{sign}\left(\sum_{n,m} n(m-\gamma n)K(n+m)\exp(-n\mathrm{N}(h))\right), \tag{2.72}$$

which is saying in particular that the sum in the right-hand side is zero if and only if $\gamma_c'(h)$ is zero.

In preparation of the next result, that investigates the regularity of the critical points in the positive semi-axis, it is useful to go through what may happen in the *tangential cases*, namely when $\gamma_c(h_0) = \gamma$ and $\gamma_c'(h_0) = 0$, for $h_0 > 0$. There are three different scenarios

1. $\gamma_c(h_0 \pm \epsilon) - \gamma < 0$ for every $\epsilon > 0$ small, that is h_0 is a maximum, and in this case at h_0 there is no phase transition simply because $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{N}(h)$ in a neighborhood of h_0, and $\mathrm{N}(\cdot)$ is analytic in the positive semi-axis;

2. $\gamma_c(h_0 \pm \epsilon) - \gamma > 0$ for every $\epsilon > 0$ small, that is h_0 is a minimum, and, as we will see in the next statement, $\widetilde{\mathrm{F}}_\gamma(\cdot)$ is at least C^2 at h_0, but we are not sure in general that h_0 is a critical point. In fact higher derivatives may exist, see Remark 2.20, but this depends on fine details and we cannot exclude that all of them exist and even that there is no singularity at all in some very special case;

3. $\gamma_c(\cdot) - \gamma$ changes sign at h_0, that is h_0 is a saddle: in this case, as we have seen in Proposition 2.17, there is a transition at h_0 and, as we will point out in Remark 2.20, this transition is smoother than in the case in which the derivative of $\gamma_c'(\cdot)$ is not zero.

Proposition 2.19. *Consider $h_0 \in \partial\mathcal{I}_b$, that is $h_0 > 0$ such that $\gamma_c(h_0) = \gamma$. The function $\widetilde{\mathrm{F}}_\gamma'(\cdot)$ is continuous at h_0, so a transition at $h > 0$ is not of first order. If $\sum_m m^2 K(m) < \infty$, $\widetilde{\mathrm{F}}_\gamma''(\cdot)$ has a jump discontinuity at h_0 (second order transition) if and only if $\gamma_c'(h_0) \neq 0$. If $\sum_m m^2 K(m) = \infty$, $\widetilde{\mathrm{F}}_\gamma''(\cdot)$ is continuous at h_0 (so the transition is of third order or more).*

Note that this statement says in particular that $\widetilde{\mathrm{F}}_\gamma''(\cdot)$ is continuous in full generality if $\gamma_c'(h_0) = 0$.

Proof. Let us first prove that $\widetilde{\mathrm{F}}_\gamma'(h)$ is continuous at h_0. For $\mathrm{N}(h)$, recall from Lemma 2.10 that $q_h\left(\overline{\lambda}_1(h), \mathrm{G}(h)\right) = 1$ (see (2.21) for the definition of $q_h(\cdot)$), therefore

$$\frac{\partial}{\partial h} q_h\left(\overline{\lambda}_1(h), \mathrm{G}(h)\right) = 0\,, \tag{2.73}$$

which directly implies that

$$\mathrm{N}'(h) \sum_{n,m} nK(n+m) \exp\left(-n\mathrm{N}(h)\right) = \exp(-h)\,. \tag{2.74}$$

For $\hat{c}_\gamma(h)$, first replace $\mathrm{G}(h) - \hat{\lambda}_1(h)$ by $\hat{c}_\gamma(h) - \gamma\left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)$ (from (2.69)) in $q_h(\cdot)$ to obtain $q_h\left(\hat{\lambda}_1(h), \hat{\lambda}_2(h)\right) = \sum_{n,m} K(n+m) \exp\left(h - \hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right)$. Recall that $\widetilde{\lambda}_2'\left(\hat{\lambda}_1(h)\right) = -1/\gamma$, which can be rewritten as (using the argument above (2.68))

$$\sum_{n,m} (m - \gamma n) K(n+m) \exp\left(h - \hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right) = 0\,, \tag{2.75}$$

for every $h > 0$. Keeping in mind this equality in evaluating $\frac{\partial}{\partial h} q_h(\hat{\lambda}_1(h), \hat{\lambda}_2(h)) = 0$, we obtain

$$\hat{c}_\gamma'(h) \sum_{n,m} nK(n+m) \exp\left(-\hat{c}_\gamma(h)n - \left(\mathrm{G}(h) - \hat{\lambda}_2(h)\right)(m - \gamma n)\right) = \exp(-h)\,. \tag{2.76}$$

Since we have that $\big(\hat\lambda_1(h_0),\hat\lambda_2(h_0)\big) = \big(\overline\lambda_1(h_0),\mathrm{G}(h_0)\big)$ and $\mathrm{N}(h_0) = \hat c_\gamma(h_0)$ (from the fact that $h_0 \in \partial\mathcal{I}_{\mathrm{b}}$), by evaluating (2.74) and (2.76) at h_0 we get

$$\widetilde{\mathrm{F}}'_\gamma(h_0) = \mathrm{N}'(h_0) = \hat c'_\gamma(h_0) = \left(\sum_{n,m} nK(n+m)\exp\left(h_0 - n\mathrm{N}(h_0)\right)\right)^{-1}, \tag{2.77}$$

and the continuity of $\widetilde{\mathrm{F}}'_\gamma(h)$ at h_0 is proven.

Now by differentiating once again (2.74) we have

$$\mathrm{N}''(h) = \frac{-\exp(-h) + \big(\mathrm{N}'(h)\big)^2 \sum_{n,m} n^2 K(n+m)\exp\left(-n\mathrm{N}(h)\right)}{\sum_{n,m} nK(n+m)\exp\left(-n\mathrm{N}(h)\right)}, \tag{2.78}$$

and by differentiating (2.76)

$$\begin{aligned}\hat c''_\gamma(h) = {} & \frac{-e^{-h} + \big(\hat c'_\gamma(h)\big)^2 \sum_{n,m} n^2 K(n+m) e^{-\hat c_\gamma(h)n - (\mathrm{G}(h)-\hat\lambda_2(h))(m-\gamma n)}}{\sum_{n,m} nK(n+m) e^{-\hat c_\gamma(h)n - (\mathrm{G}(h)-\hat\lambda_2(h))(m-\gamma n)}} \\ & + \frac{\hat c'_\gamma(h)\big(\mathrm{G}'(h) - \hat\lambda'_2(h)\big)\sum_{n,m} n(m-\gamma n)K(n+m) e^{-\hat c_\gamma(h)n - \big(\mathrm{G}(h)-\hat\lambda_2(h)\big)(m-\gamma n)}}{\sum_{n,m} nK(n+m) e^{-\hat c_\gamma(h)n - \big(\mathrm{G}(h)-\hat\lambda_2(h)\big)(m-\gamma n)}}.\end{aligned} \tag{2.79}$$

We now observe that $\mathrm{N}''(h_0)$ coincides with the first term in the right-hand side of (2.79) evaluated at $h = h_0$. Therefore $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ is continuous at h_0 if and only if the second term in the right-hand side of (2.79) vanishes at $h = h_0$. To clarify this issue we rewrite $\mathrm{G}'(\cdot) - \hat\lambda'_2(\cdot)$ by exploiting the fact that by differentiating (2.75) with respect to h we get

$$\begin{aligned}&\mathrm{G}'(h) - \hat\lambda'_2(h) = \\ &\quad - \hat c'_\gamma(h)\frac{\sum_{n,m} n(m-\gamma n)K(n+m)\exp\left(-\hat c_\gamma(h)n - \big(\mathrm{G}(h) - \hat\lambda_2(h)\big)(m-\gamma n)\right)}{\sum_{n,m}(m-\gamma n)^2 K(n+m)\exp\left(-\hat c_\gamma(h)n - \big(\mathrm{G}(h) - \hat\lambda_2(h)\big)(m-\gamma n)\right)}.\end{aligned} \tag{2.80}$$

Therefore going back to (2.77), (2.78) and (2.79), we see that if $\sum_m m^2 K(m) < \infty$ then

$$\hat c''_\gamma(h_0) - \mathrm{N}''(h_0) = -\big(\widetilde{\mathrm{F}}'_\gamma(h_0)\big)^3 \frac{\big(\sum_{n,m} n(m-\gamma n)K(n+m)\exp\left(h_0 - n\mathrm{N}(h_0)\right)\big)^2}{\sum_{n,m}(m-\gamma n)^2 K(n+m)\exp\left(h_0 - n\mathrm{N}(h_0)\right)}, \tag{2.81}$$

and if $\sum_m m^2 K(m) = \infty$ then $\hat c''_\gamma(h_0) - \mathrm{N}''(h_0) = 0$. So, if $\sum_m m^2 K(m) = \infty$ then $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ is continuous at h_0. If $\sum_m m^2 K(m) < \infty$ then (2.81) *generically* tells us that $\widetilde{\mathrm{F}}''_\gamma(\cdot)$ has a jump discontinuity at h_0, but the jump is zero if $\sum_{n,m} n(m-\gamma n)K(n+m)e^{h_0 - n\mathrm{N}(h_0)} = 0$ and, by (2.72), this is equivalent to $\gamma'_c(h_0) = 0$. The proof of Proposition 2.19 is therefore complete. □

Remark 2.20. *A sharper analysis of the singularity at $h_0 \in \mathcal{I}_b$ is possible and one sees that the closer α is to zero the more the transition is regular. The general analysis however is cumbersome due also to the fact that the transition can gain regularity from cancellations that may appear and that depend on fine details: we have already found an instance of*

this when in the proof of Proposition 2.19 we have seen that if $\sum_m m^2 K(m) < \infty$ *then the second derivative of the free energy has a jump at* h_0 *unless* $\gamma_c'(h_0) = 0$. *These cancellations are at the origin of the difficulties in resolving the issue in item (2) of the list right before Proposition 2.19 .*

2.3.3 Free energy analysis of the delocalization transition

We complete now the proof of Theorem 2.2 by studying the asymptotic behavior near h_c of $\widetilde{\mathrm{F}}_\gamma(h)$: we will show in Section 2.4 that $\widetilde{\mathrm{F}}_\gamma(h) = \mathrm{F}_\gamma(h)$.

We start by observing that, by Proposition 2.16, (2.59) reduces to study the critical behavior of $\mathrm{G}(h)$ and $D_h(1,\gamma)$. In the case $\gamma = 1$, since $D_h(1,1) = 0$, $\mathrm{F}(h) = 2\mathrm{G}(h)$ and, since we have seen that the only singularity of $\mathrm{G}(\cdot)$ is at the origin, Theorem 2.2 reduces in this case to:

Lemma 2.21. *For* $\alpha > 0$ *we have*

$$\mathrm{G}(h) \overset{h\searrow 0}{\sim} \begin{cases} \frac{1}{2}ch & \text{if } \sum_n n^2 K(n) < \infty\,, \\ \frac{1}{2}L_\alpha(h)h^{1/\alpha} & \text{if } \sum_n n^2 K(n) = \infty\,, \end{cases} \tag{2.82}$$

where c *is the same as for (2.15) and* $L_\alpha(\cdot)$ *is a slowly varying function. For* $\alpha = 0$, $\mathrm{G}(h)$ *vanishes faster than any power of* h.

Implicit expressions for $L_\alpha(\cdot)$ as well as $\mathrm{G}(\cdot)$ in terms of the inverse of a suitable slowly varying function in the case $\alpha = 0$ can be found in the proof.

Proof. Actually since (2.18) can be written as

$$\sum_{n\ge 2}(n-1)K(n)\exp(h - n\mathrm{G}(h)) = 1\,, \tag{2.83}$$

we remark that $\mathrm{G}(h)$ is the free energy of the pinning model based on a one-dimensional renewal process and the proof is therefore just a revisitation of [48, Theorem 2.1]. We give in any case a substantial part of the arguments here.

If $\sum_n n^2 K(n) < \infty$, by (DOM) we have that $\sum_{n\ge 2}(n-1)K(n)\,(1-\exp(-n\mathrm{G}(h))) \sim \mathrm{G}(h)\sum_{n\ge 2} n(n-1)K(n)$ as $h \searrow 0$. Therefore by (2.83)

$$\mathrm{G}(h)\sum_{n\ge 2} n(n-1)K(n) \sim 1 - \exp(-h) \sim h\,, \tag{2.84}$$

and this of course proves the statement (2.15) for $\widetilde{\mathrm{F}}(h)$.

If $\sum_n n^2 K(n) = \infty$ and $\alpha \in (0,1)$, by (2.83) and by Riemann sum approximation, one has

$$1 - \exp(-h) = \sum_{n\ge 2}(n-1)K(n)\,(1-\exp(-n\mathrm{G}(h))) \overset{h\searrow 0}{\sim} (\mathrm{G}(h))^\alpha L\left(\frac{1}{\mathrm{G}(h)}\right)\int_0^\infty \frac{1-\exp(-x)}{x^{1+\alpha}}\,\mathrm{d}x\,, \tag{2.85}$$

and therefore

$$\frac{1}{\alpha}\Gamma(1-\alpha)(\mathrm{G}(h))^{\alpha}L\left(1/\mathrm{G}(h)\right) \overset{h\searrow 0}{\sim} h\,, \tag{2.86}$$

and $\mathrm{G}(h) \sim \frac{1}{2}L_\alpha(h)\,h^{1/\alpha}$ where $L_\alpha(h) = 2(\alpha/\Gamma(1-\alpha))^{1/\alpha}h^{-1/\alpha}\hat{L}_\alpha(h)$ and $\hat{L}_\alpha(\cdot)$ is asymptotically equivalent to the inverse of $x \longmapsto x^\alpha L(1/x)$.

For the case $\alpha = 1$ we restart with the right-hand side of the first equality in (2.85) which equals, up to an additive term $O(\mathrm{G})$ for $\mathrm{G} \searrow 0$), to

$$\sum_{n\geq 2}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right) = \sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right) + \sum_{n>\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right)\,, \tag{2.87}$$

with $\varepsilon > 0$. By performing a Riemann sum approximation we see that the second term in the right-hand side is $O(\mathrm{G}L(1/\mathrm{G}))$. For the first one instead we use that, by Taylor formula, for every $\delta > 0$ there exists $\varepsilon > 0$ such that

$$(1-\delta)\mathrm{G}\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n} \leq \sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n^2}\left(1-\exp(-n\mathrm{G})\right) \leq (1+\delta)\mathrm{G}\sum_{n=2}^{\lfloor \varepsilon/\mathrm{G}\rfloor}\frac{L(n)}{n}\,, \tag{2.88}$$

but the sum on the leftmost and rightmost term is asymptotic to $\int_1^{1/\mathrm{G}}(L(t)/t)\,\mathrm{d}t =: \breve{L}(1/\mathrm{G})$, which is slowly varying and $L(x) = o(\breve{L}(x))$ for $x \to \infty$ [14, Proposition 1.5.9a]. At this point we go back to the first equality in (2.85) and we have

$$\mathrm{G}(h)\breve{L}\left(1/(\mathrm{G}(h))\right) \overset{h\searrow 0}{\sim} h\,. \tag{2.89}$$

Since the right-hand side of the first equality in (2.85) is an increasing function of G, (2.89) can be asymptotically inverted and the case $\alpha = 1$ is complete.

For the case $\alpha = 0$ the computation is similar. Again we replace the term $(n-1)$ with n in the right-hand side of the first equality in (2.85): the error is $O(\mathrm{G})$. Then we are left with $\sum_{n\geq 2}(L(n)/n)(1-\exp(-n\mathrm{G}))$ and we split the sum into n smaller and larger than $1/(\varepsilon\mathrm{G})$. The first sum can be treated by Riemann approximation yielding a term $O(L(1/\mathrm{G}))$. The other term instead is asymptotic to $\breve{L}(x) := \int_x^\infty (L(t)/t)$, which is slowly varying and $L(x) = o(\breve{L}(x))$ [14, Proposition 1.5.9b]. The right-hand side of the first equality in (2.85) is an increasing function of G that vanishes as G tends to zero. So $\breve{L}(x)$ can be chosen decreasing (to zero) and the slowly varying property implies that $\breve{L}(x) \geq x^{-\varepsilon}$ for every $\varepsilon > 0$ and every x sufficiently large. Hence $\breve{L}(1/\mathrm{G}(h)) \geq \mathrm{G}(h)^\varepsilon$ for h sufficiently small and $\breve{L}(1/\mathrm{G}(h)) \sim h$ readily implies that $\mathrm{G}(h) = O(h^{1/\varepsilon})$ which is what we wanted to prove. The proof of Lemma 2.21 is therefore complete. □

Proof of Lemma 2.4. The analyticity of $\gamma_c(\cdot)$ has been proven in Proposition 2.17 and the second statement of (2.25) is trivial. Let us prove the first statement of (2.25) (keeping in mind that $\alpha > 0$).

Recall the definition of $\gamma_c(\cdot)$ from (2.23). If $\sum_{n\geq 1} n^2 K(n) < \infty$ then it is easy to see that

$$\gamma_c(0) = \frac{\sum_{n,m\geq 1} mK(n+m)}{\sum_{n,m\geq 1} nK(n+m)} = 1\,. \tag{2.90}$$

Now if $\sum_{n\geq 1} n^2 K(n) = \infty$ and $\alpha \in (0,1)$, set $\mathrm{G}(h) - \overline{\lambda}_1(h) = x = o(1)$ (as $h \searrow 0$) and remark that

$$\sum_{n,m} nK(n+m)\exp(-xn) = \sum_{t\geq 2} K(t)\,\frac{e^x(1-e^{-xt}) + te^{x(1-t)}(1-e^x)}{(e^x-1)^2}\,. \tag{2.91}$$

By Riemann sum approximation, the right-hand side of (2.91) is equivalent to (as $x \searrow 0$)

$$x^{\alpha-1}L(1/x)\int_0^\infty \frac{1-\exp(-y)(1+y)}{y^{2+\alpha}}\,\mathrm{d}y\,, \tag{2.92}$$

therefore

$$\sum_{n,m\geq 1} nK(n+m)\exp(-xn) \overset{x\searrow 0}{\sim} x^{\alpha-1}\frac{\Gamma(1-\alpha)}{1+\alpha}L(1/x)\,. \tag{2.93}$$

Repeating the same argument leading to (2.93), we see that

$$\sum_{n,m\geq 1} mK(n+m)\exp(-xn) \overset{x\searrow 0}{\sim} x^{\alpha-1}\Gamma(-\alpha-1)L(1/x)\,. \tag{2.94}$$

By (2.23),(2.93) and (2.94), we get

$$\gamma_c(h) \overset{h\searrow 0}{\sim} \frac{1}{\alpha}\,. \tag{2.95}$$

For the case $\alpha = 1$, it suffices to show that

$$\sum_{n,m} nK(n+m)\exp(-xn) \overset{x\searrow 0}{\sim} \sum_{n,m} mK(n+m)\exp(-xn)\,. \tag{2.96}$$

In fact, both terms are asymptotic to $\breve{L}(x) = \int_1^{1/x} L(t)/t$, slowly varying by [14, Proposition 1.5.9b] and diverging at ∞ because $\sum_{n,m} nK(n+m) = \infty$. This can be seen by restarting from (2.91): the left-hand side of (2.96) is equal to

$$\frac{e^x}{(e^x-1)^2}\sum_{t\geq 2}\frac{L(t)}{t^3}\left((1-e^{-xt}) + te^{-xt}(1-e^x)\right) = \\ \frac{1+O(x)}{x^2}\sum_{t\geq 2}\frac{L(t)}{t^3}\left((1-e^{-xt}) - txe^{-xt}\right) + O(1)\,. \tag{2.97}$$

A Riemann sum approximation shows that if the sum is limited to $t > \varepsilon/x$, for an arbitrary $\varepsilon > 0$, yields an $O(1)$ contribution. For the term that is left we use that, by Taylor formula, for every $\delta > 0$ one finds a $\varepsilon > 0$ such that

$$\frac{1}{x^2}\sum_{t=2}^{\lfloor \varepsilon/x\rfloor}\frac{L(t)}{t^3}\left((1-e^{-xt}) - txe^{-xt}\right) \leq \frac{1+\delta}{2x^2}\sum_{t=2}^{\lfloor \varepsilon/x\rfloor}\frac{L(t)}{t^3}(tx)^2 \overset{x\searrow 0}{\sim} \frac{(1+\delta)}{2}\sum_{t=2}^{\lfloor \varepsilon/x\rfloor}\frac{L(t)}{t}\,, \tag{2.98}$$

and analogous lower bound with $1-\delta$. Since $\delta > 0$ is arbitrary, the claimed asymptotic behavior for the left-hand side of (2.96) is established. The computation for the right-hand side is very similar and left to the reader. Therefore the proof is complete. □

Recall now Proposition 2.16 and in particular the fact that $D_h(1,\gamma)$ is the result of an optimization problem, cf. (2.61), over the set B_h (cf. (2.62)). As widely used and discussed in and right after Proposition 2.17, the maximum can be achieved in the interior of B_h (Cramér regime) or at the boundary (out of Cramér regime).

Proposition 2.22. *Choose $\gamma > 1$. If $\sum_n n^2 K(n) < \infty$ we have that for h small the system is outside of the Cramér regime and*

$$\widetilde{\mathrm{F}}_\gamma(h) \overset{h\searrow 0}{\sim} \widetilde{\mathrm{F}}_1(h) = 2\mathrm{G}(h)\,. \tag{2.99}$$

If instead $\sum_n n^2 K(n) = \infty$ for h small the system is in the Cramér regime if $\gamma < 1/\alpha$ and there exists $c_{\alpha,\gamma} \in (1, \frac{1}{2}((1+\alpha)^{1/\alpha} \wedge (1+\gamma)))$ such that

$$\widetilde{\mathrm{F}}_\gamma(h) \sim c_{\alpha,\gamma}\widetilde{\mathrm{F}}_1(h)\,. \tag{2.100}$$

If $\gamma > 1/\alpha$, the system is outside of this regime and

$$\widetilde{\mathrm{F}}_\gamma(h) \sim \frac{(1+\alpha)^{1/\alpha}}{2}\widetilde{\mathrm{F}}_1(h)\,. \tag{2.101}$$

If $\alpha = 0$ the system is in the Cramér regime for every γ and $\widetilde{\mathrm{F}}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$. The asymptotic behavior of $\mathrm{G}(h)$ is given in Lemma 2.21.

Proof. Since $\gamma > 1$ we have that $D_h(1,\gamma) > 0$. We recall Proposition 2.16 and for $(\lambda_1,\lambda_2) \in B_h$, we make the change of variables $\mathrm{G}(h) - \lambda_1 = a_1\,\mathrm{G}(h)$, so $a_1 \geq 1$, and $\mathrm{G}(h) - \lambda_2 = a_2\,\mathrm{G}(h)$, so $a_2 \in [0,1]$. With these new variables we have

$$D_h(1,\gamma) = \mathrm{G}(h) \max_{a\in\mathcal{B}_h} (1 - a_1 + \gamma(1-a_2)) = \mathrm{G}(h)\left(1+\gamma - \min_{a\in\mathcal{B}_h}(a_1+\gamma a_2)\right), \tag{2.102}$$

hence

$$\widetilde{\mathrm{F}}_\gamma(h) = \min_{a\in\mathcal{B}_h}(a_1+\gamma a_2)\,\mathrm{G}(h)\,, \tag{2.103}$$

with

$$\mathcal{B}_h := \left\{a : 1 \leq a_1 \leq 1 - \frac{\overline{\lambda}_1(h)}{\mathrm{G}(h)},\, 0 \leq a_2 \leq 1\,,\, \sum_{n,m} K(n+m) e^{h-(a_1 n + a_2 m)\mathrm{G}(h)} = 1\right\}. \tag{2.104}$$

We set $\Psi_h(a,\mathrm{G}) := \sum_{n,m} K(n+m)\exp\left(-(a_1 n + a_2 m)\mathrm{G}\right)$ and for every $a \in \mathcal{B}_h$, we have that $\Psi_h(a,\mathrm{G}(h)) = \exp(-h)$, which we will use asymptotically as

$$1 - \Psi_h(a,\mathrm{G}(h)) \overset{h\searrow 0}{\sim} h\,. \tag{2.105}$$

If $\sum_n n^2 K(n) < \infty$, since $\gamma_c(0) = 1$ (from Lemma 2.4) and $\gamma > 1$, the system is outside of the Cramér regime for h small (recall that the system is in the Cramér regime if and only if $\gamma_c(h) > \gamma$) and in terms of the new variable it means that the minimizer satisfies $a_2 = 0$ (see the proof of the Proposition 2.17). Set $\overline{a} = (\overline{a}_1, 0)$ with $\overline{a}_1 \mathrm{G}(h) = \mathrm{G}(h) - \overline{\lambda}_1(h)$. By (DOM) we have

$$1 - \Psi_h(\overline{a}, \mathrm{G}(h)) = 1 - \sum_{n,m\geq 1} K(n+m) \exp\left(-\overline{a}_1 n \mathrm{G}(h)\right) \overset{h\searrow 0}{\sim} \frac{1}{2}\overline{a}_1 \mathrm{G}(h) \sum_{n\geq 2} n(n-1)K(n)\,, \tag{2.106}$$

which, together with (2.84) and (2.105), implies $\overline{a}_1 \overset{h\searrow 0}{\sim} 2$. Therefore , by (2.103), $\widetilde{\mathrm{F}}_\gamma(h) \sim 2\mathrm{G}(h)$.

When $\sum_n n^2 K(n) = \infty$, we first treat the case $\alpha \in (0,1)$: we have seen in Lemma 2.4 that in this case $\gamma_c(0) = 1/\alpha$ which implies that we are in the Cramér regime for h small if $\gamma < 1/\alpha$ and outside if $\gamma > 1/\alpha$ (and $a_2 = 0$). But let us consider the two regimes together for now and observe that for $a_2 \in [0,1]$, $a_1 \geq 1$ and $a_1 \neq a_2$

$$1 - \Psi_h(a, \mathrm{G}(h)) = \sum_{t\geq 2} K(t) \left((t-1) - \left(\frac{\exp\left(-t a_1 \mathrm{G}(h)\right) - \exp\left(-t a_2 \mathrm{G}(h) - (a_1 - a_2)\mathrm{G}(h)\right)}{\exp\left(-(a_1 - a_2)\mathrm{G}(h)\right) - 1} \right) \right), \tag{2.107}$$

and by Riemann sum approximation, the right-hand side is equivalent to (as $h \searrow 0$)

$$\mathrm{G}(h)^\alpha L\left(1/\mathrm{G}(h)\right) \int_0^\infty \frac{x - \left(\exp(-a_1 x) - \exp(-a_2 x)\right)/(a_2 - a_1)}{x^{2+\alpha}} \,\mathrm{d}x\,. \tag{2.108}$$

The integral can be made explicit:

$$1 - \Psi_h(a, \mathrm{G}) \sim b_\alpha(a) \Gamma(-\alpha - 1)\, \mathrm{G}(h)^\alpha L\left(1/\mathrm{G}(h)\right)\,, \tag{2.109}$$

with

$$b_\alpha(a) := \frac{\left({a_1}^{1+\alpha} - {a_2}^{1+\alpha}\right)}{(a_1 - a_2)} = (1+\alpha) \int_0^1 \left((1-t)a_2 + t a_1\right)^\alpha \,\mathrm{d}t\,. \tag{2.110}$$

By (2.86) and (2.105) we get that for $a \in \mathcal{B}_h$ and $a_1 \neq a_2$

$$b_\alpha(a) \sim 1 + \alpha\,, \tag{2.111}$$

but the rightmost term in (2.110) shows that the singularity in $a_1 = a_2$ is removable and one directly verifies that (2.109), and therefore (2.111), hold also for $a_1 = a_2 (= 1)$.

A number of considerations are in order:

1. By recalling the convexity arguments used in § 2.3.1, we directly have that the constraint in $\mathcal{B}_h$ can be rewritten as $a_2 = \widetilde{a}_{2,h}(a_1)$, with $\widetilde{a}_2(\cdot)$ a decreasing convex function (this is the curve appearing in Fig. 2.3, up to the affine change of variable we performed). It will be more practical at this stage to write rather $a_1 = \widetilde{a}_{1,h}(a_2)$, and

$\widetilde{a}_{1,h} : [0,1] \to [1,\infty)$ is also convex and decreasing with $\widetilde{a}_{1,h}(0) = 1 - \overline{\lambda}_1(h)/\mathrm{G}(h) \sim (1+\alpha)^{1/\alpha}$ and $\widetilde{a}_{1,h}(1) = 1$ where we have used (2.111). The choice in favor of $\widetilde{a}_{1,h}(\cdot)$ over $\widetilde{a}_{2,h}(\cdot)$ is because we prefer having a h dependence in the image rather than in the domain.

2. Since $1 - \Psi_h(\cdot, \mathrm{G})$ is concave, also $b_\alpha(\cdot)$ is (this can also be verified directly) and the constraint in $\mathcal{B}_h$, namely (2.111), becomes in the limit $b_\alpha(a) = 1+\alpha$, with $a_1 \in [1, (1+\alpha)^{1/\alpha}]$ and $a_2 \in [0,1]$. Such a constraint can be expressed (like in § 2.3.1) as $a_1 = \widetilde{a}_1(a_2)$, where $\widetilde{a}_1 : [0,1] \to [1,(1+\alpha)^{1/\alpha}]$ is a convex decreasing function. This function is smooth (in fact, analytic, by the Implicit Function Theorem [66]) in the interior of the domain and one directly computes $\widetilde{a}_1'(0^+) = -1/\alpha$ and $\widetilde{a}_1'(1^-) = -1$ by expanding (2.110) to second order, which actually coincide with the limits as $h \searrow 0$ respectively of $\widetilde{a}_{1,h}'(0^+)$ (this is precisely (2.95)) and $\widetilde{a}_{1,h}'(1^-) = -1$ (in fact: $\widetilde{a}_{1,h}'(1^-) = -1$ for every $h > 0$ by symmetry).

3. But (2.111), that is the convergence of the constraint function for $h \searrow 0$, implies $\lim_h \widetilde{a}_{1,h}(a_2) = \widetilde{a}_1(a_2)$ for every $a_2 \in [0,1]$, as well as the convergence of the derivative for $a_2 \in (0,1)$, because we are dealing with a sequence of convex functions and because the limit is differentiable. Note that we have already pointed out that there is convergence of the derivatives also at the endpoints and that, by analyticity, $\widetilde{a}_1(\cdot)$ is strictly convex. This allows to conclude that the unique minimizer $\hat{a}(h)$ for (2.103) converges to the unique minimizer of $\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$, with

$$\mathcal{B} := \left\{ a : 1 \le a_1 \le (1+\alpha)^{1/\alpha},\, 0 \le a_2 \le 1\,,\, b_\alpha(a) = 1+\alpha \right\}. \tag{2.112}$$

The limit problem, like the approaching ones, can of course be rewritten in a one-dimensional form using $\widetilde{a}_1(\cdot)$ and $\widetilde{a}_{1,h}(\cdot)$.

At this point we can treat separately the non Cramér case, in which $a_2 = 0$ (this happens when $\gamma > 1/\alpha$, but also when $\gamma = 1/\alpha$ by a limit procedure), so we have $a_1 = (1+\alpha)^{1/\alpha}$ and $\widetilde{\mathrm{F}}_\gamma(h) \sim (1+\alpha)^{1/\alpha}\mathrm{G}(h)$.

If $\gamma < 1/\alpha$ instead, $\widetilde{\mathrm{F}}_\gamma(h) \sim \min_{a\in\mathcal{B}}(a_1 + \gamma a_2)\mathrm{G}(h)$, so $c_{\alpha,\gamma} = \frac{1}{2}\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$ and (2.100) is proven. Since $a_1(\cdot)$ is strictly convex and decreases from $a_1(0) = (1+\alpha)^{1/\alpha}$ to $a_1(1) = 1$, one can easily see (just evaluate $a_1 + \gamma a_2$ at $a = (1,1)$ and $a = ((1+\alpha)^{1/\alpha}, 0)$) that $c_{\alpha,\gamma} \in (1, \frac{1}{2}((1+\alpha)^{1/\alpha} \wedge (1+\gamma)))$.

Remark 2.23. *For $\alpha \in (0,1)$ we can use the Lagrange multiplier method to solve the limit optimization problem $\min_{a\in\mathcal{B}}(a_1 + \gamma a_2)$. With s as multiplier we have*

$$\begin{cases} s\,(1+\alpha)\,(1 - {a_1}^{\alpha}) = 1\,, \\ s\,(1+\alpha)\,({a_2}^{\alpha} - 1) = \gamma\,, \end{cases} \tag{2.113}$$

which implies that

$$a_2{}^{\alpha} = 1 - \gamma(a_1{}^{\alpha} - 1)\,. \tag{2.114}$$

In particular, and consistently with what precedes, if $\gamma > 1/(a_1{}^{\alpha} - 1) \geq 1/\alpha$ *no solution is found and we are out of the Cramér regime, and* $(a_1, a_2) = ((1+\alpha)^{1/\alpha}, 0)$. *If* $\gamma < 1/\alpha$ *instead a solution is found and in fact we are in the Cramér regime.*

For the case $\alpha = 1$, from Lemma 2.4, we have $\gamma_c(0) = 1$, which implies that for every $\gamma > 1$, we are outside of the Cramér regime for h small and $a = \overline{a} = (\overline{a}_1, 0)$. Observe that

$$1 - \Psi_h(\overline{a}, \mathrm{G}) = \sum_{t \geq 2} \frac{L(t)}{t^3} \left((t-1) - \frac{\exp(-\overline{a}_1 \mathrm{G}(h)(t-1)) - 1}{1 - \exp(\overline{a}_1 \mathrm{G}(h))} \right), \tag{2.115}$$

and follow the same procedure adopted for the case $\alpha = 1$ in Lemma 2.21: split the sum in (2.115) into t larger than $\varepsilon/(\overline{a}_1 \mathrm{G})$, for an arbitrary $\varepsilon > 0$ (which yields $O(1)$ contribution) and to t smaller than $\varepsilon/(\overline{a}_1 \mathrm{G})$ to obtain

$$1 - \Psi_h(\overline{a}, \mathrm{G}) \overset{h \searrow 0}{\sim} \frac{1}{2} \overline{a}_1 \mathrm{G}(h) \widetilde{L}\left(1/(\overline{a}_1 \mathrm{G}(h))\right)\,. \tag{2.116}$$

We know that the left-hand side is equivalent to h from (2.105), and using the property of slowly varying functions (2.208) ($\widetilde{L}\left(1/(\overline{a}_1 \mathrm{G}(h))\right) \sim \widetilde{L}\left(1/(\mathrm{G}(h))\right)$), then by (2.89) we get $\overline{a}_1 \sim 2$. Therefore by (2.103), we obtain $\widetilde{\mathrm{F}}_\gamma(h) \sim 2\mathrm{G}(h)$.

For the case $\alpha = 0$, we do not strive for the sharp prefactor, since we did not go for the sharp behavior of $\mathrm{G}(h)$, cf. Lemma 2.21. We just use (2.103) and observe that $\min_{a \in \mathcal{B}_h}(a_1 + \gamma a_2)$ is bounded (because the supremum is over a set that is bounded uniformly in h, with h in a right neighborhood of zero). Hence in this case $\widetilde{\mathrm{F}}_\gamma(h) = O(h^{1/\varepsilon})$ for every $\varepsilon > 0$. The proof is therefore complete. □

2.3.4 The bio-physics model and other examples

In this section we make $\gamma_c(\cdot)$ explicit for some choices of $K(\cdot)$, starting with [72].

The bio-physics model

We refer to Section 2.1.2. We have seen that the geometric constant s is irrelevant then we take $s = 1$ and $B(l) = 1/l^c$. Recall that $E_b > 0$ is the binding energy and $E_l > 0$ is the loop initiation cost. Set

$$K_h(n) := \frac{c_K}{n^c} \exp\left(h(E_b - E_l \mathbf{1}_{n>2}) - n\mathrm{G}(h)\right)\,, \tag{2.117}$$

with $c_K = \sum_{n,m \geq 1} 1/(n+m)^c = \sum_{n \geq 2}(n-1)/n^c$ is the normalization constant and $\mathrm{G}(h)$ is the only solution to $\sum_{n,m} K_h(n+m) = 1$.

In Figure 2.4, a plot of $\gamma_c(\cdot)$ for values of E_B, E_l, c and γ chosen in [72], shows that the system exhibits a unique transition at $h_{c,\gamma} \simeq 1.676$ (the vertical dashed line). For $\gamma = 1.15$, observe that the system is in the Cramér regime for $h < h_{c,\gamma}$ and outside of the Cramér regime if $h \geq h_{c,\gamma}$. In [72] the transition at $h_{c,\gamma}$ is described as between a phase with microscopic free strands at the end of the polymer and macroscopic free strands.

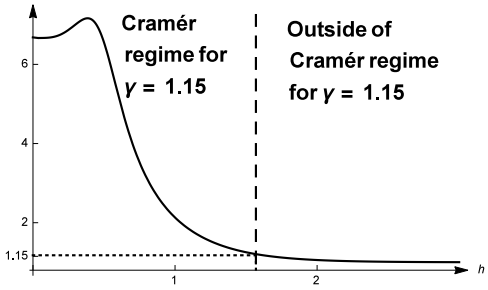

FIGURE 2.4. A representation of the function $\gamma_c(\cdot)$ for the distribution $K_h(\cdot)$ with $c = 2.15$, $E_b = 6$ and $E_l = 3$. The horizontal dashed line corresponds to the value of $\gamma = 1.15$ and the vertical one to the critical point $h_{c,\gamma} \simeq 1.676$. Computing $\gamma_c(h)$ is immediate from (2.23) once $\mathrm{G}(h)$ and $\overline{\lambda}_1(h)$ are known: we have first obtained $\mathrm{G}(h)$ by solving numerically $\sum_{n,m} \widetilde{K}_h(n,m) = 1$ and then we have solved numerically $q_h(\overline{\lambda}_1, \mathrm{G}) = 1$ for $\overline{\lambda}_1(h)$.

A more explicitly solvable class: basic case

We exploit the inter-arrival law used in [49] in the one-dimensional case. Recall that Euler's Gamma function $\Gamma(z) := \int_0^\infty t^{z-1} \exp(-t)\, \mathrm{d}t$ defines an analytic function on $\{z \in \mathbb{C} : Re(z) > 0\}$ and that for $z \in \mathbb{C} \setminus \{0, -1, -2, ...\}$, it verifies $\Gamma(z+1) = z\Gamma(z)$ (therefore $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$). Recall also that the Taylor coefficients of the function $(1-z)^\alpha$ for $|z| < 1$ and $\alpha \in \mathbb{R} \setminus \{0, 1, 2, ...\}$, is known exactly

$$\sum_{n \geq 0} \frac{\Gamma(n-\alpha)}{n!} z^n = \Gamma(-\alpha)(1-z)^\alpha\,, \tag{2.118}$$

and asymptotically from Stirling's formula we have $\Gamma(n-\alpha)/n! \overset{n\to\infty}{\sim} 1/n^{1+\alpha}$. Note that the first terms of the series in (2.118) have alternating signs.

First, let us suppose that $\alpha \in (0,1)$ and set

$$K_1(n) := \frac{\Gamma(n-\alpha-1)}{\Gamma(-\alpha-1)n!}\,, \quad \text{for } n \geq 2 \tag{2.119}$$

and from (2.118) we have that $\sum_{n\geq 2}(n-1)K_1(n) = 1$ and by (2.91) and the fact that $\sum_{n,m} nK(n+m)z^n = z\,\alpha\,(1-z)^{\alpha-1}$ and $\sum_{n,m} mK(n+m)z^n = z(1-z)^{\alpha-1}$ we obtain $\gamma_c(h) = 1/\alpha$. This implies that there is no transition in the localized phase and for every $h > 0$, the free energy is either – recall (2.70) – $\mathrm{N}(h)$ (if $\gamma \geq \gamma_c(h)$: out of the Cramér regime) or $\hat{c}_\gamma(h)$ (if $\gamma < \gamma_c(h)$, the Cramér regime, recall (2.69)).

A more explicitly solvable class: some generalizations of the basic case

More general explicit cases can be built by modifying $K_1(\cdot)$ on a finite number of sites. Let us choose now $K_2(2) = K_1(2)$, $K_2(3) = \kappa$ and normalize the rest

$$K_2(n) = c_\kappa K_1(n), \quad \text{for} \quad n \geq 4 \tag{2.120}$$

with $c_\kappa = (1 - K_1(2) - 2\kappa) / (1 - K_1(2) - 2K_1(3))$ and $\kappa \in [0, 1)$. With this choice of the inter-arrival $K(\cdot)$ there are two transitions in the localized phase if and only if $\gamma > 2$, see Figure 2.5: the system here is in the Cramér regime only for intermediate values of h.

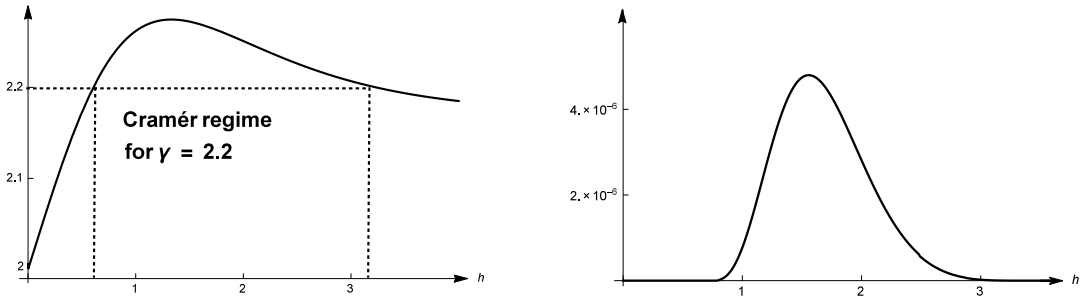

FIGURE 2.5. On the left the function $\gamma_c(\cdot)$ for the inter-arrival distribution $K_2(\cdot)$, with $\alpha = 0.5$ and $\kappa = 0.01$. The horizontal dashed line corresponds to the value of $\gamma = 2.2$. There are two values of h (vertical dashed lines) such that $\gamma_c(h) = \gamma$ and therefore there are two transitions. On the right we plot the difference $\mathrm{N}(\cdot) - \widetilde{\mathrm{F}}_\gamma(\cdot)$, that is $\mathrm{N}(\cdot) - \hat{c}_\gamma(\cdot)$ (recall (2.69) and (2.70)), which makes clear the presence of the two transitions. The resolution of the graph does not allow to appreciate the positivity of such a difference for example at $h = 3$ where it is about 6×10^{-9}.

We then present an example with $\alpha \in (1, 2)$: Proposition 2.22 shows that in this case for h small the system is outside the Cramér regime. We take $K_3(2) = \varrho$ and $K_3(n) = c_\varrho K_1(n)$ for $n \geq 3$ with $c_\varrho = (1-\varrho)/\left(\sum_{n\geq 3}(n-1)K_1(n)\right)$. A look at Figure 2.6 shows that, if $\gamma < 2.27$, the system is outside of the Cramér regime when h is below a critical value $h_{c,\gamma}$. For larger values of γ the system is outside of the Cramér regime for every $h > 0$.

2.4 The free model

In this section we use the notation $\gamma_N := M/N$ and $\gamma'_N := M'/N'$. Recall that we assume $M \geq N$, but $M' \in \{0, \dots, M\}$ may be smaller than $N' \in \{0, \dots, N\}$. We use the short-cut of saying that γ is in the Cramer region if $(1, \gamma) \in E_h$, i.e. if $(1, \gamma)$ is in the Cramer region. Figure 2.7 and its caption sum up properties of $D_h(1, \cdot)$ and of $\widetilde{\mathrm{F}}.(h)$ that will come handy in the remainder.

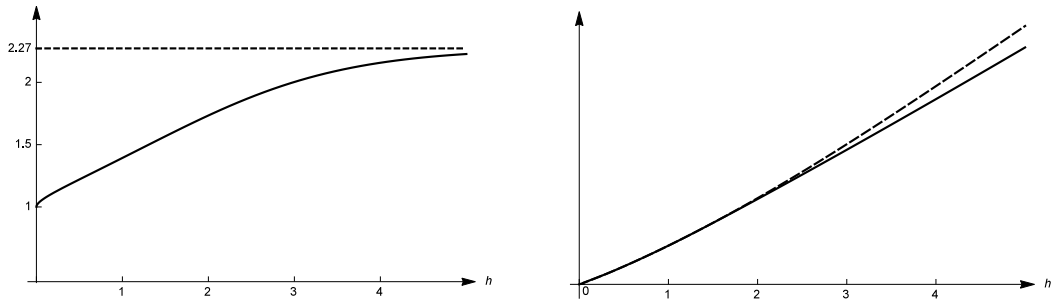

FIGURE 2.6. For the inter-arrival distribution $K_3(\cdot)$, the function $\gamma_c(\cdot)$ (on the left) is stricly increasing (for $\alpha = 1.5, \varrho = 0.02$), this implies that if $\gamma_0 := \lim_{h\to\infty} \gamma_c(h) \le \gamma$ (dashed line represent the value of γ_0 computed using (2.25)), there is no transition in the system. On the right, we present the free energy $\widetilde{\mathrm{F}}_\gamma(h)$ and $\mathrm{N}(h)$ (dashed line, recall (2.70)) for $\gamma = 1.5$.

2.4.1 A preliminary result and free energy equivalence

We start with a preliminary important lemma on the constrained free energy: a minimal part of its strength will be used to prove, just below (Proposition 2.25), that this free energy coincides with the one of the free model. The full strength of this lemma is used in § 2.4.2.

Lemma 2.24. *There exists $a(h) > 0$ such that for $\gamma'_N \ge \gamma_N$ we have*

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \ge a(h)\left(1 - \frac{\gamma_N}{\gamma'_N}\right) + \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{M'}{M}\right). \tag{2.121}$$

For $\gamma'_N \le \gamma_N$ we have

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \ge \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)\big|_{\gamma=\gamma_N}(\gamma_N - \gamma'_N) + \widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{N'}{N}\right), \tag{2.122}$$

where $\partial_\gamma \widetilde{\mathrm{F}}_\gamma(h) \ge 0$ and for every $\varepsilon \in (0, \gamma_c(h))$ we have $\inf_{\gamma\le\gamma_c(h)-\varepsilon} \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h) > 0$.

In particular for every ε as above and every $L > 0$ there exists $a_{\varepsilon,L}(h) > 0$ such that

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \ge a_{\varepsilon,L}(h)\left(\left|\gamma_N - \gamma'_N\right| + \mathbf{1}_{\gamma'_N\ge\gamma_N}\gamma_N\left(1 - \frac{M'}{M}\right) + \mathbf{1}_{\gamma'_N<\gamma_N}\left(1 - \frac{N'}{N}\right)\right), \tag{2.123}$$

for $\gamma_N \le \gamma_c(h) - \varepsilon$ and $\gamma'_N \in [1/L, L]$.

Proof. If $\gamma'_N \ge \gamma_N$ we write

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) = \left(\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\right) + \frac{\gamma_N}{\gamma'_N}\widetilde{\mathrm{F}}_{\gamma'_N}(h)\left(1 - \frac{M'}{M}\right), \tag{2.124}$$

and it suffices to focus on the first term between parentheses in the right-hand side and by using $\gamma = \gamma_N$ and $\gamma' = \gamma'_N$ to keep expressions short, since $\gamma \mapsto \widetilde{\mathrm{F}}_\gamma(h)$ is concave we have $\widetilde{\mathrm{F}}_{\gamma'}(h) \le \widetilde{\mathrm{F}}_\gamma(h) + \partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)(\gamma' - \gamma)$ so

$$\widetilde{\mathrm{F}}_\gamma(h) - \frac{\gamma}{\gamma'}\widetilde{\mathrm{F}}_{\gamma'}(h) \ge \left(1 - \frac{\gamma}{\gamma'}\right)\left(\widetilde{\mathrm{F}}_\gamma(h) - \gamma\partial_\gamma \widetilde{\mathrm{F}}_\gamma(h)\right), \tag{2.125}$$

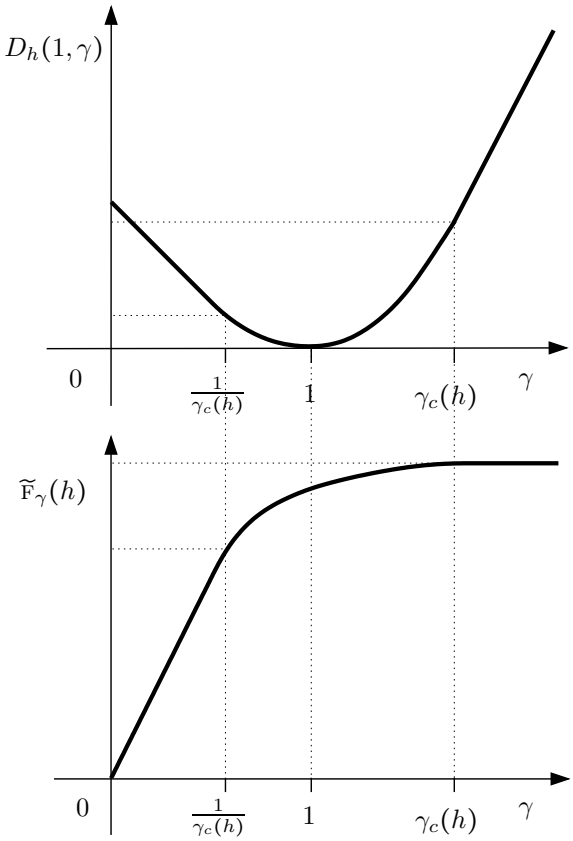

FIGURE 2.7. We plot the Large Deviations functional $D_h(1,\gamma)$ and the free energy $\widetilde{\mathrm{F}}_\gamma(h)$, that coincides with $\mathrm{F}_\gamma(h)$, as functions of γ for a given $h > 0$. Relevant features are the convexity of the first and concavity of the second, which become strict in the Cramer region or interval $(1/\gamma_c(h), \gamma_c(h))$. Both functions are real analytic except at the boundary of the Cramer interval and they are affine functions outside of this interval. This and a number of other properties can be extracted from the variational expression (2.43) for $D_h(1,\gamma)$, the analysis in §2.3 and formula (2.59) for $\mathrm{F}_\gamma(h)$. In particular we have that $\partial_\gamma D_h(1,\gamma)$ is equal to $\mathrm{G}(h)$ for $\gamma > \gamma_c(h)$ – see (2.23) for an explicit expression of $\gamma_c(h)$ – and to $-\mathrm{G}(h)/\gamma_c(h) + D_h(1,\gamma_c(h))$ for $\gamma < 1/\gamma_c(h)$. Note also that the fact that the lower endpoint of Cramer interval is $1/\gamma_c(h)$ follows from the symmetry $D_h(v_1,v_2) = D_h(v_2,v_1)$ and the scaling behavior $D_h(cv) = cD_h(v)$ for $c > 0$ and $v = (v_1,v_2)$, for every v_1 and v_2 positive. The free energy becomes constant above $\gamma_c(h)$ and it is equal to $\mathrm{G}(h) + |\overline{\lambda}_1(h)|$ (recall (2.22)).

and the right-hand side is non-negative because $\gamma \mapsto \widetilde{\mathrm{F}}_\gamma(h)$ is concave and $\widetilde{\mathrm{F}}_0(h) = 0$, so $\widetilde{\mathrm{F}}_\gamma(h) - \gamma\partial_\gamma\widetilde{\mathrm{F}}_\gamma(h)$ is non decreasing in γ and it is even increasing if $\gamma \in (1/\gamma_c(h), \gamma_c(h))$. This implies in particular that

$$a(h) := \inf_{\gamma\geq 1} \left(\widetilde{\mathrm{F}}_\gamma(h) - \gamma\partial_\gamma\widetilde{\mathrm{F}}_\gamma(h)\right) = \widetilde{\mathrm{F}}_1(h) - \partial_\gamma\widetilde{\mathrm{F}}_\gamma(h)\big|_{\gamma=1} > 0\,, \tag{2.126}$$

and (2.121) is proven.

Let us turn then to $\gamma'_N \leq \gamma_N$ and (2.122). This time we write

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) = \left(\widetilde{\mathrm{F}}_{\gamma_N}(h) - \widetilde{\mathrm{F}}_{\gamma'_N}(h)\right) + \left(1 - \frac{N'}{N}\right)\widetilde{\mathrm{F}}_{\gamma'_N}(h)\,, \tag{2.127}$$

so it suffices to bound from below the first term in parentheses in the right-hand side as claimed in (2.122), but this is a direct consequence of concavity in γ of the free energy.

The validity of (2.123) follows from (2.121) and (2.122) by elementary considerations. □

We are now ready to show that the free and constrained models have the same free energy:

Proposition 2.25. *For every $h \in \mathbb{R}$ we have that the limit that defines $\mathrm{F}_\gamma(h)$, i.e. (2.12), exists and*

$$\mathrm{F}_\gamma(h) = \widetilde{\mathrm{F}}_\gamma(h)\,. \tag{2.128}$$

Proof. Recall that we treat a result for $N \to \infty$ and $M \sim \gamma N$. First of all observe that, since $Z^f_{N,M,h} \geq Z^c_{N,M,h}$ we just need to worry about the upper bound. For the upper bound we use Proposition 2.16 and, more precisely the uniform estimate (2.60). As a matter of fact the supremum with respect to γ in (2.60) can be extended to $\gamma \in [1/L, L]$: the choice to restrict to $\gamma \geq 1$ in that proof was just to conform to the convention chosen from the beginning (but too restrictive for this proof as we will see). But the extension to $\gamma \in [1/L, L]$ can also be obtained from the result for $\gamma \in [1, L]$ by exploiting the symmetry of the expressions when exchanging M and N. Therefore we have that for every $\delta > 0$ and every $L > 1$ there exists ℓ_0 such that

$$Z^c_{N',M',h} \leq \exp\left(N'\left(\widetilde{\mathrm{F}}_{\gamma'_N}(h) + \frac{\delta}{2}\right)\right), \tag{2.129}$$

for every N' and M' such that both M' and $N' \geq \ell_0$ and $\gamma'_N = M'/N' \in [1/L, L]$. At the same time we require that

$$\frac{2\gamma h}{L} \leq \widetilde{\mathrm{F}}_\gamma(h)\,, \tag{2.130}$$

where γ is the one that appears in the final statement, so $\gamma \geq 1$ without loss of generality, and for (2.130) to hold is just a matter of choosing L sufficiently large and this, in turn, just affects the choice of ℓ_0. But then thanks to Lemma 2.24

$$\widetilde{\mathrm{F}}_{\gamma_N}(h) - \frac{N'}{N}\widetilde{\mathrm{F}}_{\gamma'_N}(h) \geq 0\,, \tag{2.131}$$

so

$$Z^c_{N',M',h} \le \exp\left(N\left(\widetilde{\mathrm{F}}_{\gamma_N}(h)+\frac{\delta}{2}\right)\right) \le \exp\left(N\left(\widetilde{\mathrm{F}}_{\gamma}(h)+\delta\right)\right), \tag{2.132}$$

still for M' and $N' \ge \ell_0$ and $\gamma'_N = M'/N' \in [1/L, L]$: we have used the continuity of $\gamma \mapsto F_\gamma(h)$, see the caption of Fig. 2.7, and we have chosen N sufficiently large. Therefore by using $K_f(n) \le n^C$ (for some $C > 0$) we have

$$Z^f_{N,M,h} \le N^{2C} \sum_{N',M'} Z^c_{N',M',h} \le (1+\gamma)N^{2C+2}\exp\left(N\left(\widetilde{\mathrm{F}}_{\gamma_N}(h)+\delta\right)\right) + N^{2C}\ell_0^2 \max_{N',M'\le\ell_0} Z^c_{N',M',h} + N^{2C} \sum_{\substack{N',M':\, N'\vee M'\ge\ell_0 \\ N'\le N, M'\le M, \gamma'_N\notin[1/L,L]}} Z^c_{N',M',h}\,. \tag{2.133}$$

Since the maximum in the second line is just a constant we are left with controlling the last sum. About this last sum, which we control without exploiting the condition $N' \vee M' \ge \ell_0$, we just remark that it can be split into the two cases $\gamma'_N < 1/L$ and $\gamma'_N > L$. The two terms are very similar and the second is larger than the first when $\gamma > 1$ (and N large) and if $\gamma = 1$ the bound is the same for both terms. So we focus on the second one and remark that the rough bound $Z^c_{N',M',h} \le \exp(hN')$ and the observation that $\gamma_N N \ge M' = \gamma'_N N' > LN'$, hence $N' \le 2\gamma N/L$ for N large, yield $Z^c_{N',M',h} \le \exp(2\gamma hN/L)$. Therefore, by (2.130), this term is negligible with respect to the first term in the right-hand side of (2.133) and we are done. □

2.4.2 Sharp estimates on $Z^f_{N,M,h}$

The case $h > 0$: Proof of (1) in Theorem 2.6

Recall the definition of $Z^f_{N,M,h}$ in (2.10) and that we work with $M \sim \gamma N$, γ in the Cramer region. For every $a \in (0,1)$

$$\sum_{i=\lfloor aN\rfloor}^{N}\sum_{j=0}^{M} K_f(i)K_f(j)Z^c_{N-i,M-j,h} = \sum_{i=\lfloor aN\rfloor}^{N}\sum_{j=0}^{M} K_f(i-\lfloor aN\rfloor)\frac{K_f(i)}{K_f(i-\lfloor aN\rfloor)}K_f(j)Z^c_{N-i,M-j,h}\,, \tag{2.134}$$

and since $K_f(\cdot)$ is positive and it has an asymptotic power law behavior there exists $C > 0$ such that (for N sufficiently large)

$$\sum_{i=\lfloor aN\rfloor}^{N}\sum_{j=0}^{M} K_f(i)K_f(j)Z^c_{N-i,M-j,h} \le N^C \sum_{i=0}^{N-\lfloor aN\rfloor}\sum_{j=0}^{M} K_f(i)K_f(j)Z^c_{N-\lfloor aN\rfloor-i,M-j,h} = N^C Z^f_{N-\lfloor aN\rfloor,M,h}\,. \tag{2.135}$$

Note that the leftmost term in (2.134) and (2.135) is $Z^f_{N,M,h}$ if $a = 0$. On the other hand by (2.60) and Proposition 2.25

$$\log\left(N^C Z^f_{N-\lfloor aN\rfloor,M,h}\right) \overset{N\to\infty}{\sim} N(1-a)\mathrm{F}_{\gamma_N/(1-a)}(h)\,, \tag{2.136}$$

and we observe that

$$\mathrm{F}_{\gamma_N}(h) > (1-a)\mathrm{F}_{\gamma_N/(1-a)}\,, \tag{2.137}$$

which follows from (2.121) by choosing $M' = M$ so $N'/N = (\gamma_N/\gamma'_N)$ so that one obtains $\mathrm{F}_\gamma(h) - (\gamma/\gamma')\mathrm{F}_{\gamma'}(h) > 0$ for every $\gamma' > \gamma \geq 1$, and this inequality becomes (2.137) if we choose $\gamma' = \gamma_N/(1-a)$.

At this point we observe that, since $K_f(0) = 1$, we have $Z^f_{N,M,h} \geq Z^c_{N,M,h}$ and $\log Z^c_{N,M,h} \sim N\mathrm{F}_{\gamma_N}(h)$ (by (2.60)) and therefore, by (2.135)–(2.137) we see that for every $a \in (0,1)$ there exists $q > 0$ such that

$$Z^f_{N,M,h} = (1 + O(\exp(-qN))) \sum_{i=0}^{\lfloor aN \rfloor} \sum_{j=0}^{M} K_f(i) K_f(j) Z^c_{N-i,M-j,h}\,. \tag{2.138}$$

A parallel, somewhat easier, argument can be put at work when we restrict the summation in the definition of $Z^f_{N,M,h}$ in (2.10) to $j \geq \lfloor aN \rfloor$. We have to use again Lemma 2.24: (2.121) for $N' = N$ simply becomes the fact that $\gamma \mapsto \mathrm{F}_\gamma(h)$ is (strictly) increasing for $\gamma < \gamma_c(h)$ and this allows to conclude that for every $a \in (0,1)$ there exists $q > 0$ such that

$$Z^f_{N,M,h} = (1 + O(\exp(-qN))) \sum_{i=0}^{\lfloor aN \rfloor} \sum_{j=0}^{\lfloor aN \rfloor} K_f(i) K_f(j) Z^c_{N-i,M-j,h}\,. \tag{2.139}$$

With (2.139) we now want to show that we can restrict the sum in (2.10) to a small (since we can choose $a > 0$ small) macroscopic square. We want now to show that we can restrict almost to a microscopic square: a microscopic square would be a square of size that does not diverge with N. The result we shall now prove is

$$Z^f_{N,M,h} = \left(1 + O\left(\exp(-(\log N)^{3/2})\right)\right) \sum_{i=0}^{\ell_N} \sum_{j=0}^{\ell_N} K_f(i) K_f(j) Z^c_{N-i,M-j,h}\,, \tag{2.140}$$

where

$$\ell_N := \lfloor (\log N)^2 \rfloor\,. \tag{2.141}$$

For this choose a small so that $(M-j)/(N-i)$ is in the Cramer region for all values of i and j in the summation in (2.139). We can then apply (2.49) and, more precisely, the following consequence of (2.49): for every N sufficiently large

$$Z^c_{N',M',h} \leq C \exp\left(N' \mathrm{F}_{\gamma'_N}(h)\,,\right)\,, \tag{2.142}$$

where $C > 0$ and $N - \lfloor aN \rfloor \leq N' \leq N$, $M - \lfloor aN \rfloor \leq M' \leq M = \gamma_N N$. For this we exploit (2.123) of Lemma 2.24: since i and j are (macroscopically) small we have that there exists $c = c(h,a) > 0$ such that

$$N\mathrm{F}_{\gamma_N}(h) - N'\mathrm{F}_{\gamma'_N}(h) \geq c\left(N\left|\gamma'_N - \gamma_N\right| + \mathbf{1}_{\gamma'_N \geq \gamma_N}\left(M - M'\right) + \mathbf{1}_{\gamma'_N < \gamma_N}\left(N - N'\right)\right)\,, \tag{2.143}$$

where $N' = N - i$, $M' = M - j$, N is sufficiently large and both i and j in their range of summation. But now we recall that we aim at (2.140) and therefore if $i < \ell_N$ then $j \geq \ell_N$

and the same is true if we exchange i and j. Therefore, omitting the constant c, the right-hand side of (2.143) is equal to $\gamma'_N(N-N') \geq (N-N')$ for $\gamma'_N \geq \gamma_N$ and since $N-N' = i$ and either $i \geq \ell_N$, or $i < \ell_N$ and then $j \geq \ell_N$, we get that $N-N' \geq (M-M')/\gamma_N \geq \ell_N/(2\gamma)$. For $\gamma'_N \leq \gamma_N$ instead the right-hand side of (2.143) is equal to $M-N\gamma'_N+N-N'$ which, on the one hand, is bounded from below by $M-N\gamma_N+N-N' = N-N'$. On the other hand it is equal to $(M-M')-(N-N')(\gamma'_N-1)$ which is bounded from below by $(M-M')-(N-N')(\gamma_N-1)$. For $\gamma'_N \leq \gamma_N$ we have $M-M' \geq \ell_N$ since if $M-M' = j < \ell_N$ then $N-N' = i \geq \ell_N$ and we get that $\gamma'_N = (M-j)/(N-i) \geq (\gamma_N - \ell_N/N)/(1-\ell_N/N)$ which is strictly larger than γ_N for N large. So either $N-N' \leq \frac{1}{2(\gamma_N-1)}\ell_N$, so $(M-M')-(N-N')(\gamma_N-1) \geq \frac{1}{2}\ell_N$, or $N-N' > \frac{1}{2(\gamma_N-1)}\ell_N$. Hence the right-hand side of (2.143) is bounded from below for $\gamma'_N \leq \gamma_N$ by

$$\frac{c}{2}\ell_N \min\left(\frac{1}{\gamma_N-1}, 1\right). \tag{2.144}$$

Recalling the lower bound found for $\gamma'_N \geq \gamma_N$, we see that if we set $c_\gamma := \frac{1}{2}\min(1/\gamma, 1)$ we have

$$N\mathrm{F}_{\gamma_N}(h) - N'\mathrm{F}_{\gamma'_N}(h) \geq c_{\gamma_N}\ell_N, \tag{2.145}$$

for $(N-N', M-M') \in ([0,aN]^2 \setminus [0,\ell_N)^2) \cap \mathbb{Z}^2$. But then (2.140) becomes evident from (2.142), (2.145) and the fact the summation is on less than $a^2N^2 = O(N^2)$ sites: so the total contribution by summing over the sites in small macroscopic square minus the almost microscopic square is $O(N^2\exp(N\mathrm{F}_{\gamma_N}(h) - c_{\gamma_N}\ell_N))$. On the other hand $Z^f_{N,M,h} \geq Z^c_{N,M,h} \geq AN^{-1/2}\exp(N\mathrm{F}_{\gamma_N}(h))$ for some $A > 0$ and N large, cf. (2.49), and (2.140) is proven.

The question of the sharp estimates on Z^f is then reduced to find the leading behavior of

$$\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)Z^c_{N-i,M-j,h}. \tag{2.146}$$

Observe that $(N-i, M-j)/|(N-i, M-j)|$ is close to $(1,\gamma)/|(1,\gamma)|$ and hence it is in a compact subset J of E_h for every $i, j \in [0, \ell_N]$. So we have by (2.49)

$$\begin{aligned}
&\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)Z^c_{N-i,M-j,h} \overset{N\to\infty}{\sim} \frac{A(\gamma)\exp\left(\mathrm{F}_{\gamma_N}(h)N\right)}{\sqrt{N}} \\
&\sum_{i=0}^{\ell_N}\sum_{j=0}^{\ell_N} K_f(i)K_f(j)\exp\left(-N\left(\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{(M-j)/(N-i)}(h) + \frac{i}{N}\mathrm{F}_{(M-j)/(N-i)}(h)\right)\right).
\end{aligned} \tag{2.147}$$

A Taylor expansion yields that there exists $\widetilde{\gamma} \in J$ such that

$$\begin{aligned}
&\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{(M-j)/(N-i)}(h) = \\
&\qquad \left(\gamma_N - \frac{M-j}{N-i}\right)\partial_\gamma \mathrm{F}_\gamma(h)\big|_{\gamma=(M-j)/(N-i)} + \frac{1}{2}\left(\gamma_N - \frac{M-j}{N-i}\right)^2 \partial^2_\gamma \mathrm{F}_\gamma(h)\big|_{\gamma=\widetilde{\gamma}},
\end{aligned} \tag{2.148}$$

and since

$$\gamma_N - \frac{M-j}{N-i} = \frac{j}{N} - \gamma_N\frac{i}{N} + O\left(\frac{\ell_N^2}{N^2}\right) = O\left(\frac{\ell_N}{N}\right), \tag{2.149}$$

the second term in the right-hand side of (2.148) is $O\left(\ell_N^2/N^2\right)$ and finally

$$\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{(M-j)/(N-i)}(h) + \frac{i}{N}\mathrm{F}_{(M-j)/(N-i)}(h) = \frac{j\partial_\gamma \mathrm{F}_\gamma(h) + i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)}{N} + \frac{i}{N}O\left(\gamma_N - \gamma\right) + O\left(\frac{\ell_N^2}{N^2}\right). \quad (2.150)$$

Recall now that $\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)$ is positive for $\gamma < \gamma_c(h)$, so that (2.150) implies that the double sum in the right-hand side of (2.147) converges to

$$C_{\gamma,h} := \left(\sum_{i=0}^{\infty} K_f(i)\exp\left(-i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)\right)\right)\left(\sum_{j=0}^{\infty} K_f(j)\exp\left(-j\partial_\gamma \mathrm{F}_\gamma(h)\right)\right), \quad (2.151)$$

and

$$Z^f_{N,M,h} \sim A(\gamma)C_{\gamma,h}\frac{\exp\left(\mathrm{F}_{\gamma_N}(h)N\right)}{\sqrt{N}}. \quad (2.152)$$

This completes the proof of (1) in Theorem 2.6 with $c_{\gamma,h} = A(\gamma)C_{\gamma,h}$. □

The case $h < 0$: Proof of (2) in Theorem 2.6

In the delocalized phase, $\widetilde{\tau}_h$ is terminating and $Z^c_{N,M,h} = \mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right)$. Recall that τ is recurrent, i.e. $\sum_{n,m}\mathrm{K}(n,m) = 1$, then

$$\widetilde{K}_h(\infty) = 1 - \sum_{(n,m)\in\mathbb{N}^2}\widetilde{K}_h(n,m) = 1 - \exp(h) > 0. \quad (2.153)$$

For the renewal function we write

$$\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right) = \sum_{j=0}^{\infty}\widetilde{K}_h^{j*}(N,M) = \sum_{j=0}^{\infty}\exp(jh)\mathrm{K}^{j*}(N,M), \quad (2.154)$$

where $\widetilde{K}_h^{j*}(\cdot,\cdot)$ is the j-fold convolution of $\widetilde{K}_h(\cdot,\cdot)$.

For the rest of the section, c_i, $i = 1,2,\ldots$ denote positive constants. We have the following estimate:

Lemma 2.26. *Let $\alpha > 0$, there exists $c > 0$ such that for every $j \in \mathbb{N}$ and for every (N,M)*

$$\mathrm{K}^{j*}(N,M) \le j^c\,\mathrm{K}(N,M). \quad (2.155)$$

Moreover,

$$\lim_{N,M\to\infty}\frac{\mathrm{K}^{j*}(N,M)}{\mathrm{K}(N,M)} = j. \quad (2.156)$$

Proof. It is clear that (2.155) holds for $j = 1$. Assume that it is true for $j < 2s$ and we want to show it for $j = 2s$. Observe that

$$\mathrm{K}^{2s*}(N,M) \le 2\sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor}\mathrm{K}^{s*}(n,m)\mathrm{K}^{s*}(N-n,M-m) + \left(\sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=1}^{\lfloor M/2\rfloor}\right)\mathrm{K}^{s*}(n,m)\mathrm{K}^{s*}(N-n,M-m) := Q_1 + (Q_2 + Q_3). \quad (2.157)$$

First we have by induction hypothesis

$$Q_1 \le 2s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor} \mathrm{K}^{s*}(n,m)\mathrm{K}(N-n,M-m)\,, \tag{2.158}$$

and from (2.208), we see that there exists $c_1 > 0$ such that $L(ux) \le c_1 L(x)$ for every $u \in [1/2,1]$ and $x \ge 1$, therefore $\mathrm{K}(N-n,M-m) \le c_1 2^{2+\alpha}\mathrm{K}(N,M)$ and

$$Q_1 \le c_1 2^{3+\alpha-c}(2s)^c \mathrm{K}(N,M) \sum_{n,m} \mathrm{K}^{s*}(n,m)\,. \tag{2.159}$$

Now observe that $Q_2(N,M) = Q_3(M,N)$, so in view of the bound we are after it suffices to consider Q_2. Applying (2.155) to the first term in the sum in Q_2, there exists $c_2 > 0$ such that

$$Q_2 \le s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n,m)\mathrm{K}^{s*}(N-n,M-m) \le c_2 2^{-c}(2s)^c K(M)\,. \tag{2.160}$$

On the other hand, applying (2.155) to the second term in the sum in Q_2, we get

$$Q_2 \le s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}^{s*}(n,m)\mathrm{K}(N-n,M-m) \le c_2 2^{-c}(2s)^c K(N)\,. \tag{2.161}$$

It suffices then to prove that there exists $c_3 > 0$ such that $K(N) \wedge K(M) \le c_3 \mathrm{K}(N,M)$. By elementary arguments we see that this follows if we can show that for every $x,y \ge 0$ and for every slowly varying function $L(\cdot)$, there exists $c_{L(\cdot)} > 0$ such that

$$xL(x) \vee yL(y) \ge c_{L(\cdot)}(x+y)L(x+y)\,. \tag{2.162}$$

By symmetry it suffices to consider the case $y \ge x$ and in this case it suffices to show that $yL(y) \ge c_{L(\cdot)}(x+y)L(x+y)$. Since $(x+y) \in [y,2y]$ we can apply (2.208) to see that there exists $c'_{L(\cdot)} > 0$ such that $L(y) \ge c'_{L(\cdot)} L(x+y)$ for every $y \ge x$ and this implies the desired inequality, and therefore (2.162), with $c_{L(\cdot)} = c'_{L(\cdot)}/2$.

Therefore $Q_2 + Q_3 \le c_4 2^{1-c}(2s)^c \mathrm{K}(N,M)$ and

$$\mathrm{K}^{2s*}(N,M) \le (c_1 2^{3+\alpha-c} + c_4 2^{1-c})(2s)^c \mathrm{K}(N,M)\,, \tag{2.163}$$

and if $c = 1 + \log_2(c_1 2^{2+\alpha} + c_4)$, we obtain (2.155) for $j = 2s$. The procedure can be repeated for $j = 2s+1$ with minor changes. Therefore (2.155) is proven.

For what concerns (2.156), observe that it holds for $j = 1$. Assume that it is still valid up to $j = s$ and write

$$\frac{\mathrm{K}^{(s+1)*}(N,M)}{\mathrm{K}(N,M)} = \sum_{n=1}^{N-1}\sum_{m=1}^{M-1} \frac{\mathrm{K}^{s*}(n,m)\mathrm{K}(N-n,M-m)}{\mathrm{K}(N,M)}\,. \tag{2.164}$$

Split the double sum in (2.164) to four terms:

$$S_1 = \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor} \mathrm{K}^{s*}(n,m)\frac{\mathrm{K}(N-n,M-m)}{\mathrm{K}(N,M)} \tag{2.165}$$

$$S_2 = \sum_{n=1}^{\lfloor N/2\rfloor-1}\sum_{m=1}^{\lfloor M/2\rfloor-1} \mathrm{K}(n,m)\frac{\mathrm{K}^{s*}(N-n,M-m)}{\mathrm{K}(N,M)} \tag{2.166}$$

$$S_3 = \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor-1} \mathrm{K}^{s*}(n,M-m)\frac{\mathrm{K}(N-n,m)}{\mathrm{K}(N,M)} \tag{2.167}$$

$$S_4 = \sum_{n=1}^{\lfloor N/2\rfloor-1}\sum_{m=1}^{\lfloor M/2\rfloor} \mathrm{K}^{s*}(N-n,m)\frac{\mathrm{K}(n,M-m)}{\mathrm{K}(N,M)} \tag{2.168}$$

For any fixed n and m and as $N, M \to \infty$, the two ratios in S_1 and S_2 converge respectively to 1 and to s (recall that $\mathrm{K}(n,m) = K(n+m)$) and the two ratios are uniformly bounded (from the uniform convergence property of the slowly varying functions (2.208) for the ratio in S_1 and from (2.155) for the ratio in S_2). Then by (DOM), we obtain $S_1 + S_2 \to 1 + s$ as $N, M \to \infty$.

Since S_3 and S_4 are essentially the same quantity when we exchange M and N, we just focus on S_3. If we first assume that $M \geq N$ (hence $M + N \in [M, 2M]$), from (2.155) and (2.208) we obtain that

$$\begin{aligned} S_3 \leq s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor-1} \mathrm{K}(n,M-m)\frac{\mathrm{K}(N-n,m)}{\mathrm{K}(N,M)} \leq \\ c_5 s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor-1} \frac{L(M+n)}{(M+n)^{2+\alpha}}\frac{L(N+m)}{(N+m)^{2+\alpha}}\frac{(N+M)^{2+\alpha}}{L(N+M)} \leq \\ c_6 s^c \sum_{n=1}^{\lfloor N/2\rfloor}\sum_{m=1}^{\lfloor M/2\rfloor-1} \frac{L(N+m)}{(N+m)^{2+\alpha}} \leq \frac{c_6 s^c}{2} N \sum_{n>N}\frac{L(n)}{n^{2+\alpha}} \leq c_7 L(N) N^{-\alpha}\,. \end{aligned} \tag{2.169}$$

By repeating the argument for $N \geq M$ we obtain that $S_3 = O(L(M)M^{-\alpha})$ in this case. Therefore, since $\alpha > 0$, $\lim_{N,M\to\infty} S_3 = 0$ by the basic properties of slowly varying functions and an elementary argument. □

To prove part (2) in Theorem 2.6, we need to know the sharp estimates in the constrained case for $h < 0$:

Proposition 2.27. *If $h < 0$, then there exists $c_h > 0$ such that for every (N, M)*

$$Z^c_{N,M,h} \leq c_h \mathrm{K}(N,M) \tag{2.170}$$

where $c_h = \sum_{j=0}^{\infty} j^c \exp(jh)$ and c is as in Lemma 2.26. Moreover

$$Z^c_{N,M,h} \overset{N,M\to\infty}{\sim} \frac{\exp(h)}{(1-\exp(h))^2}\mathrm{K}(N,M)\,. \tag{2.171}$$

Proof. From (2.154), we have

$$\frac{Z^c_{N,M,h}}{\mathrm{K}(N,M)} = \frac{\mathbf{P}\left((N,M) \in \widetilde{\tau}_h\right)}{\mathrm{K}(N,M)} = \sum_{j=0}^{\infty} \exp(jh)\frac{\mathrm{K}^{j*}(N,M)}{\mathrm{K}(N,M)}\,, \tag{2.172}$$

and using (2.155), for fixed j, we see that the ratio is bounded above by j^c, therefore we obtain (2.170).

For (2.171), the ratio in (2.172) converges to j as $N, M \to \infty$ from (2.156) and is bounded from (2.155), then by (DOM) we get

$$\lim_{N,M\to\infty} \frac{Z^c_{N,M,h}}{\mathrm{K}(N,M)} = \sum_{j=0}^{\infty} j \exp(jh) = \frac{\exp(h)}{(1-\exp(h))^2}\,, \tag{2.173}$$

□

We are now ready to prove the sharp estimate of $Z^f_{N,M,h}$:

Proposition 2.28. *Suppose that $M \sim \gamma N$. For $h < 0$, as $N \to \infty$*

- *If $\overline{\alpha} < 1 + \alpha/2$, we have*

$$Z^f_{N,M,h} \sim \frac{K_f(N)K_f(M)}{1-\exp(h)}\,. \tag{2.174}$$

- *If $\overline{\alpha} > 1 + \alpha/2$, we have*

$$Z^f_{N,M,h} \sim \frac{\exp(h)\big(\sum_{n\geq 0} K_f(n)\big)^2}{(1-\exp(h))^2}\mathrm{K}(N,M)\,. \tag{2.175}$$

Proof. Let us write

$$\frac{Z^f_{N,M,h}}{K_f(N)K_f(M)} = \sum_{n=0}^{N}\sum_{m=0}^{M} Z^c_{n,m,h}\frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}\,. \tag{2.176}$$

We split the last sum into

$$T_1 + T_2 + T_3 + T_4 = \left(\sum_{n=0}^{\lfloor N/2\rfloor}\sum_{m=0}^{\lfloor M/2\rfloor} + \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=0}^{\lfloor N/2\rfloor}\sum_{m=\lfloor M/2\rfloor+1}^{M} + \sum_{n=\lfloor N/2\rfloor+1}^{N}\sum_{m=0}^{\lfloor M/2\rfloor}\right) Z^c_{n,m,h}\frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)}\,. \tag{2.177}$$

For T_1, for fixed n and m, the ratio in (2.177) converges to 1 and by (2.208), this ratio is bounded. Then by (DOM), Fubini-Tonelli Theorem and the fact that $\mathrm{K}(\cdot,\cdot)$ is a discrete probability density we obtain (recall (2.17) and in this case $\mathrm{G}(h) = 0$)

$$\lim_{N,M\to\infty} T_1 = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} Z^c_{n,m,h} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\widetilde{\mathrm{K}}^{j*}_h(n,m) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\exp(jh)\mathrm{K}^{j*}(n,m) = \frac{1}{1-\exp(h)}\,. \tag{2.178}$$

For T_2, if $\overline{\alpha} \in [1, 1+\alpha/2)$ and $\sum_{n\geq 0} K_f(n) < \infty$, using (2.171), we obtain (recall that $M \sim \gamma N$)

$$T_2 \leq c_8 \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n,m) \frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)} = O\left(N^{2\overline{\alpha}-\alpha-2} \frac{L(N)}{(\overline{L}(N))^2}\right), \tag{2.179}$$

so $T_2 \to 0$ as $N \to \infty$.

If $\overline{\alpha} \leq 1$ and $\sum_{n\geq 0} K_f(n) = \infty$, using once again (2.171) with (2.209), we get

$$\begin{aligned} T_2 \leq c_8 \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=\lfloor M/2\rfloor+1}^{M} \mathrm{K}(n,m) \frac{K_f(N-n)K_f(M-m)}{K_f(N)K_f(M)} \leq \\ c_9 \mathrm{K}(N,M) \frac{\sum_{n=\lfloor N/2\rfloor+1}^{N} K_f(N-n) \sum_{m=\lfloor M/2\rfloor+1}^{M} K_f(M-m)}{K_f(N)K_f(M)} = O\left(N^{-\alpha} L(N)\right), \end{aligned} \tag{2.180}$$

and $T_2 \to 0$ as $N \to \infty$ in this case too.

Let us look at T_3 (the argument for T_4 is identical). We have

$$\begin{aligned} T_3 = \sum_{n=0}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} Z^c_{n,m,h} \frac{K_f(N-n)}{K_f(N)} \frac{K_f(M-m)}{K_f(M)} \leq \\ c_{10} \sum_{n=0}^{\lfloor N/2\rfloor} K(M+n) \sum_{m=\lfloor M/2\rfloor+1}^{M} \frac{K_f(M-m)}{K_f(M)} \leq c_{11} \frac{L(M)}{M^{1+\alpha}} \frac{\sum_{m=0}^{M} K_f(m)}{K_f(M)}. \end{aligned} \tag{2.181}$$

If $\overline{\alpha} \in [1, 1+\alpha/2)$ and $\sum_m K_f(m) < \infty$ then $T_3 = O(N^{\overline{\alpha}-\alpha-1} L(N)/\overline{L}(N))$ and it tends to zero because $\alpha > 0$ implies $1+\alpha > \overline{\alpha}$. If instead $\overline{\alpha} \leq 1$ and $\sum_m K_f(m) = \infty$ the last ratio in the rightmost term in (2.181) is $O(M) = O(N)$. Hence $T_3 = O(L(N)N^{-\alpha})$ and T_3 tends to zero also in this case. The proof of (2.174) is therefore complete.

Now for $\overline{\alpha} > 1 + \alpha/2$ (which implies that $\overline{\alpha} > 1$ since $\alpha > 0$), we write

$$\frac{Z^f_{N,M,h}}{\mathrm{K}(N,M)} = \sum_{n=0}^{N} \sum_{m=0}^{M} Z^c_{n,m,h} \frac{K_f(N-n)K_f(M-m)}{\mathrm{K}(N,M)}. \tag{2.182}$$

We split the last sum to $U_1 + U_2 + U_3 + U_4$ as in (2.177). Since $\sum_{i\geq 0}\sum_{j\geq 0} Z^c_{i,j,h} = 1/(1-\exp(h)) < \infty$ (see (2.178)) we have

$$U_1 = O\left(\frac{K_f(N)K_f(M)}{\mathrm{K}(N,M)}\right) = O\left(N^{2+\alpha-2\overline{\alpha}} \frac{\left(\overline{L}(N)\right)^2}{L(N)}\right) = o(1). \tag{2.183}$$

Also the terms U_3 and U_4 give a vanishing contribution. Let us see it for U_3 (the computation is identical for U_4):

$$\begin{aligned} U_3 = \sum_{n=0}^{\lfloor N/2\rfloor} \sum_{m=\lfloor M/2\rfloor+1}^{M} \frac{Z^c_{n,m,h}}{\mathrm{K}(N,M)} K_f(N-n)K_f(M-m) \leq \\ c_{12} \sum_{n=0}^{\lfloor N/2\rfloor} K_f(N-n) \sum_{m=\lfloor M/2\rfloor+1}^{M} K_f(M-m) \leq c_{13} \sum_{n\geq N/2} K_f(n) = O(\overline{L}(N) N^{1-\overline{\alpha}}). \end{aligned} \tag{2.184}$$

The relevant contribution comes from U_2:

$$\begin{aligned} U_2 &= \sum_{n=\lfloor N/2\rfloor+1}^{N} \sum_{m=\lfloor M/2\rfloor+1}^{M} \frac{Z^c_{n,m,h}}{\mathrm{K}(N,M)} K_f(N-n)K_f(M-m) \\ &= \sum_{n=0}^{N-\lfloor N/2\rfloor-1} \sum_{m=0}^{M-\lfloor M/2\rfloor-1} \frac{Z^c_{N-n,M-m,h}}{\mathrm{K}(N,M)} K_f(n)K_f(m)\,. \end{aligned} \tag{2.185}$$

But the ratio in the last term is bounded, cf. (2.171), and in fact (2.171) tells us also that for every m and n

$$\lim_{N,M\to\infty} \frac{Z^c_{N-n,M-m,h}}{\mathrm{K}(N,M)} = \frac{\exp(h)}{(1-\exp(h))^2}\,, \tag{2.186}$$

which, by applying (DOM), implies

$$\lim_{N,M\to\infty} U_2 = \frac{\exp(h)}{(1-\exp(h))^2}\left(\sum_{n\geq 0} K_f(n)\right)^2, \tag{2.187}$$

and completes the proof of (2.175) and, in turn, the proof of Proposition 2.27. □

2.4.3 Path properties: proof of Theorem 2.7

In this section, we suppose that $M \sim \gamma N$ and $\alpha > 0$.

Proof of (1) in Theorem 2.7. We first consider the case $h < 0$. Recall that $(\mathcal{F}_1, \mathcal{F}_2)$ is the last renewal epoch in $[0,N]\times[0,M]$. If $\overline{\alpha} < 1+\alpha/2$, for fixed i and j (so we can assume $i < N$ and $j < M$) we have

$$\begin{aligned} \mathbf{P}^f_{N,M,h}\left((\mathcal{F}_1,\mathcal{F}_2) = (i,j)\right) &= \frac{Z^c_{i,j,h} K_f(N-i)K_f(M-j)}{Z^f_{N,M,h}} \\ &\overset{N\to\infty}{\sim} (1-\exp(h)) Z^c_{i,j,h} \frac{K_f(N-i)K_f(M-j)}{K_f(N)K_f(M)}\,, \end{aligned} \tag{2.188}$$

where the estimation follows from (2.174). Since i and j are $O(1)$ and by (2.208) the ratio in the rightmost term in (2.188) converges to one. Hence it suffices to prove that $(1-\exp(h))\sum_{i,j} Z^c_{i,j,h} = 1$, but this is done in (2.178). We then recall that $Z_{i,j,h} = \mathbf{P}\left((i,j) \in \widetilde{\tau}_h\right)$ and (2.33) is proven.

Now recall that $(\mathcal{L}_1, \mathcal{L}_2) := (N-\mathcal{F}_1, M-\mathcal{F}_2)$. If $\overline{\alpha} > 1+\alpha/2$, for fixed i and j by using (2.171) and (2.175) we see that

$$\begin{aligned} \mathbf{P}^f_{N,M,h}\left((\mathcal{L}_1,\mathcal{L}_2) = (i,j)\right) &= \frac{Z^c_{N-i,M-j,h} K_f(i)K_f(j)}{Z^f_{N,M,h}} \\ &\overset{N\to\infty}{\sim} \frac{\mathrm{K}(N-i,M-j)}{\mathrm{K}(N,M)} \frac{K_f(i)K_f(j)}{\left(\sum_{n\geq 0} K_f(n)\right)^2}\,. \end{aligned} \tag{2.189}$$

The proof of (2.34) is therefore complete by observing that by (2.208) the first ratio in (2.189) converges to one.

We are then left (for $h < 0$) with (2.35). Here we prove more: consider $(\mathcal{E}_1, \mathcal{E}_2) := \max\{\tau \cap [0, \lfloor N/2 \rfloor] \times [0, \lfloor M/2 \rfloor]\}$ under $\mathbf{P}^f_{N,M,h}$ for $\overline{\alpha} > 1 + \alpha/2$. If i and j are fixed, by (2.175) we have

$$\mathbf{P}^f_{N,M,h}\left((\mathcal{E}_1, \mathcal{E}_2) = (i,j)\right) = \exp(h) \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} \frac{Z^c_{i,j,h} \mathrm{K}(s-i, t-j) Z^f_{N-s,M-t,h}}{Z^f_{N,M,h}}$$
$$\underset{N \to \infty}{\sim} \frac{(1-\exp(h))^2 Z^c_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2} \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} Z^f_{N-s,M-t,h} \frac{\mathrm{K}(s-i, t-j)}{\mathrm{K}(N,M)}. \quad (2.190)$$

By making the change of variable $(s,t) \to (N-s, M-t)$ we see that for i and j of $O(1)$, the very last ratio in (2.190) converges to one and the same ratio is bounded by (2.208) in all the range of the sum. Hence, by the (DOM), the expression in (2.190) converges to

$$\frac{(1-\exp(h))^2 Z^c_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2} \sum_{s,t \geq 0} Z^f_{s,t,h}, \quad (2.191)$$

and observe that

$$\sum_{s,t \geq 0} Z^f_{s,t,h} = \sum_{s,t \geq 0} \sum_{n=0}^{s} \sum_{m=0}^{t} Z^c_{n,m,h} K_f(s-n) K_f(t-m)$$
$$= \sum_{n,m \geq 0} Z^c_{n,m,h} \sum_{s \geq n} \sum_{m \geq t} K_f(s-n) K_f(t-m) = \frac{\left(\sum_{n \geq 0} K_f(n)\right)^2}{1-\exp(h)}, \quad (2.192)$$

where the last equality follows from (2.178). Therefore the law of $(\mathcal{E}_1, \mathcal{E}_2)$ under $\mathbf{P}^f_{N,M,h}$ converges for $N \to \infty$ to the probability distribution that assigns to (i,j) probability $(1-\exp(h)) Z^c_{i,j,h}$ (which is correctly normalized, like (2.33), by (2.178)).

Let $(\mathcal{C}_1, \mathcal{C}_2) := \min\{\tau \cap [\lfloor N/2 \rfloor, N] \times [\lfloor M/2 \rfloor, M]\}$ be the first renewal epoch in $[\lfloor N/2 \rfloor, N] \times [\lfloor M/2 \rfloor, M]$ and set $(\mathcal{H}_1, \mathcal{H}_2) := (N - \mathcal{C}_1, M - \mathcal{C}_2)$. For fixed i and j by (2.175) we have

$$\mathbf{P}^f_{N,M,h}\left((\mathcal{H}_1, \mathcal{H}_2) = (i,j)\right) = \exp(h) \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} \frac{Z^f_{i,j,h} \mathrm{K}(s-i, t-j) Z^c_{N-s,M-t,h}}{Z^f_{N,M,h}}$$
$$\underset{N \to \infty}{\sim} \frac{(1-\exp(h))^2 Z^f_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2} \sum_{s \geq \lfloor N/2 \rfloor} \sum_{t \geq \lfloor M/2 \rfloor} Z^c_{N-s,M-t,h} \frac{\mathrm{K}(s-i, t-j)}{\mathrm{K}(N,M)}. \quad (2.193)$$

The second ratio converges to 1 (same argument as above) and therefore (2.193) converges to

$$\frac{(1-\exp(h)) Z^f_{i,j,h}}{\left(\sum_{n \geq 0} K_f(n)\right)^2}, \quad (2.194)$$

and from (2.192), we see that this expression adds up ($i, j \geq 0$) to one. The law of $(\mathcal{H}_1, \mathcal{H}_2)$ converges as $N \to \infty$ to the probability distribution that assigns to (i,j) probability (2.194). We conclude that the contacts are either close to $(0,0)$ or to the last renewal epoch: therefore we get (2.35).

□

Proof of (2) in Theorem 2.7. In this case $h > 0$ and $\gamma \in (1/\gamma_c(h), \gamma_c(h))$: the positivity of $\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)$ and $\partial_\gamma \mathrm{F}_\gamma(h)$ is a direct consequence of the strict concavity of $\mathrm{F}_\cdot(h)$ in the Cramér region (see caption of Figure 2.7). Then choose (i,j) with non-negative integer entries. By (2.49) and (2.152) we have

$$\mathbf{P}^f_{N,M,h}\left((\mathcal{L}_1,\mathcal{L}_2) = (i,j)\right) = \frac{Z^c_{N-i,M-j,h} K_f(i) K_f(j)}{Z^f_{N,M,h}} \underset{N\to\infty}{\sim} \frac{A\left((M-j)/(N-i)\right)\sqrt{N}}{C_{\gamma,h} A(\gamma)\sqrt{N-i}} \exp\left(\mathrm{F}_{(M-j)/(N-i)}(h)(N-i) - \mathrm{F}_{M/N}(h) N\right) K_f(i) K_f(j)\,. \tag{2.195}$$

Observe now that the ratio in the rightmost term of (2.195) converges as $N \to \infty$ to $1/C_{\gamma,h}$ (defined in (2.151)) and using (2.150) and the fact that $(M-j)/(N-i)$ is close to γ, (2.195) converges to

$$\frac{1}{C_{\gamma,h}} \exp\left(-j\partial_\gamma \mathrm{F}_\gamma(h) - i\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right)\right) K_f(i) K_f(j)\,, \tag{2.196}$$

where $C_{\gamma,h}$ is defined in (2.151). Hence (2.36) is proven.

To complete the proof of Theorem 2.7, we need to show that for $h > 0$ and if $M \sim \gamma N$ such that $\gamma \in (1/\gamma_c(h), \gamma_c(h))$, $\mathbf{P}^f_{N,M,h}$ converges, for $N \to \infty$, to the law of a bivariate renewal with the inter-arrival probability given in (2.37). For this, fix a $k \in \mathbb{N}$ and for every $\{(i_n, j_n)\}_{n=0,1,\dots,k}$ with $(i_0, j_0) = (0,0)$, $i_1 < i_2 < \cdots < i_k < N$ and $j_1 < j_2 < \ldots < j_k < M$, we have

$$\mathbf{P}^f_{N,M,h}\left(\tau_n = (i_n, j_n) \text{ for } n = 1, 2, \dots, k\right) = \left(\prod_{n=1}^{k} K(i_n - i_{n-1} + j_n - j_{n-1}) \exp(h)\right) \frac{Z^f_{N-i_k,M-j_k,h}}{Z^f_{N,M,h}}\,. \tag{2.197}$$

Since $\gamma_k = (M - j_k)/(N - i_k)$ is close to γ (in the Cramér region) as $N \to \infty$ for fixed i_k and j_k, by (2.152), we see that the ratio in (2.197) is equal to

$$(1 + o(1)) \exp\left(-N\left(\mathrm{F}_{\gamma_N}(h) - \mathrm{F}_{\gamma_k}(h) + \frac{i_k}{N}\mathrm{F}_{\gamma_k}(h)\right)\right)\,. \tag{2.198}$$

Following the same procedure used for (2.150), we see that the exponent converges to

$$-i_k\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right) - j_k \partial_\gamma \mathrm{F}_\gamma(h)\,, \tag{2.199}$$

and therefore the left-hand side of (2.197) converges to

$$\prod_{n=1}^{k} K(i_n - i_{n-1} + j_n - j_{n-1}) \exp\left(h - (i_n - i_{n-1})\left(\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h)\right) - (j_n - j_{n-1})\partial_\gamma \mathrm{F}_\gamma(h)\right)\,. \tag{2.200}$$

We are left with proving that (2.37) is a probability distribution. Recall from (2.61) and (2.62) that

$$D_h(1,\gamma) = \max_{(\lambda_1,\lambda_2)\in B_h} (\lambda_1 + \gamma\lambda_2) = \hat{\lambda}_1(\gamma) + \gamma\hat{\lambda}_2(\gamma)\,, \tag{2.201}$$

with of course $(\hat{\lambda}_1(\gamma), \hat{\lambda}_2(\gamma)) \in B_h$ (keep in mind that they depend also on h) i.e.

$$\sum_{n,m} K(n+m) \exp\big(h - n(\mathrm{G}(h) - \hat{\lambda}_1(\gamma)) - m(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\big) = 1\,, \tag{2.202}$$

therefore $\mathrm{F}_\gamma(h) = (\mathrm{G}(h) - \hat{\lambda}_1(\gamma)) + \gamma(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))$. Replace $\mathrm{G}(h) - \hat{\lambda}_1(\gamma)$ in (2.202) by $\mathrm{F}_\gamma(h) - \gamma(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))$, we obtain

$$\sum_{n,m} K(n+m) \exp\big(h - n\mathrm{F}_\gamma(h) - (m-\gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\big) = 1\,, \tag{2.203}$$

and recall that $\hat{c}_\gamma(h) = \mathrm{F}_\gamma(h)$ in the Cramér regime, then by (2.75) one obtains

$$\sum_{n,m} (m-\gamma n) K(n+m) \exp\big(h - n\mathrm{F}_\gamma(h) - (m-\gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\big) = 0\,. \tag{2.204}$$

Differentiating (2.203) with respect to γ and using (2.204), we obtain that

$$\big(\partial_\gamma \mathrm{F}_\gamma(h) - (\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\big) \sum_{n,m} nK(n+m) \exp\big(-n\mathrm{F}_\gamma(h) - (m-\gamma n)(\mathrm{G}(h) - \hat{\lambda}_2(\gamma))\big)\,, \tag{2.205}$$

is zero, and this implies directly that

$$\partial_\gamma \mathrm{F}_\gamma(h) = \mathrm{G}(h) - \hat{\lambda}_2(\gamma)\,, \tag{2.206}$$

$$\mathrm{F}_\gamma(h) - \gamma\partial_\gamma \mathrm{F}_\gamma(h) = \mathrm{G}(h) - \hat{\lambda}_1(\gamma)\,. \tag{2.207}$$

Therefore from (2.202), we get that (2.37) is a probability distribution. □

Appendix

2.A Slowly and regularly varying functions

A function $L : [0,\infty) \to (0,\infty)$ is a *slowly varying* function at ∞ if it is measurable and if $\lim_{x\to\infty} \frac{L(ux)}{L(x)} = 1$ for every $u > 0$. The function $x \mapsto L(1/x)$ is slowly varying at zero if $L(\cdot)$ is slowly varying at ∞. It can be shown that this convergence holds uniformly in u [14, Theorem 1.2.1]: for every $0 < c_1 < c_2 < \infty$

$$\lim_{x\to\infty} \sup_{u\in[c_1,c_2]} \left| \frac{L(ux)}{L(x)} - 1 \right| = 0\,. \tag{2.208}$$

A function of the form $x \mapsto x^a L(x)$, $a \in \mathbb{R}$, is said to be *regularly varying (at ∞) of exponent a* with an analogous definition for regularly varying at zero. Examples of slowly varying

functions include logarithmic functions (of course the trivial example is the constant) like $a(\log(x))^b$ as $x \to \infty$ with $a > 0$ and $b \in \mathbb{R}$. We refer to [14] for the full theory of slowly and regularly varying functions: we just recall some basic important facts. First of all, both $L(x)$ and $1/L(x)$ are $o(x^\varepsilon)$ for every $\varepsilon > 0$, which directly implies that if $f(\cdot)$ is regularly varying with exponent a and $g(\cdot)$ is regularly varying with exponent $b < a$, then $g(x) = o(f(x))$.

We will often use that for $\beta > 0$ [14, Sec. 1.5.6]

$$\sum_{n \geq N} \frac{L(n)}{n^{1+\beta}} \underset{\sim}{^{N\to\infty}} \frac{L(N)}{\beta N^\beta} \quad \text{and} \quad \sum_{n=1}^{N} \frac{L(n)}{n^{1-\beta}} \underset{\sim}{^{N\to\infty}} \frac{L(N)}{\beta N^{-\beta}}, \tag{2.209}$$

which can be proven by Riemann sum approximation. We will often use Riemann sum approximations involving regularly varying functions also beyond (2.209) and the central tool to control these approximations are the so called *Potter bounds* [14, Th. 1.5.6].

Another important issue is about asymptotic invertibility of regular functions: a regular function of exponent $a > 0$ (respectively $a < 0$) is asymptotically equivalent to an increasing (respectively decreasing) function. Moreover the inverse of a monotonic regularly varying function of exponent $a \neq 0$ is a regularly varying function of exponent $1/a$. In different terms, if $f(\cdot)$ is regularly varying of exponent $a \neq 0$, then there exists $g(\cdot)$ regularly varying of exponent $1/a$ such that $f(g(x)) \sim g(f(x)) \sim x$ [14, Sec. 1.5.7]. Occasionally we use other properties of slowly varying functions and we refer directly to [14].

Chapter 3

The generalized Poland-Scheraga model: disordered case

Abstract

We study the disordered version of the generalized Poland-Scheraga model. We focus on the influence of disorder on the denaturation transition: we want to determine whether the presence of randomness modifies the critical properties of the system with respect to the homogeneous case. We prove that the disorder is irrelevant if $\alpha < 1$ and we show that for $\alpha > 1$, the quenched and annealed critical points differ (basing on coarse graining techniques and fractional moment method), proving the presence of a relevant disorder regime. This chapter is work in collaboration with G. Giacomin and Q. Berger.

3.1 Introduction of the model and results

The analysis of the DNA denaturation phenomenon – that is the unbinding at high temperature of two strands of DNA – lead to the proposal of a very elementary model, the Poland-Scheraga (PS) model [78], that turns out to be relevant not only at a conceptual and qualitative level [42, 48], but also at a quantitative level [17, 18]. This model can naturally embody the inhomogeneous character of the DNA polymer, which is a monomer sequence of four different types (A,T, G and C). The binding energy for A-T pairs is different from the binding energy for G-C pairs. The quantitative analysis is then based on finite length chains with a given sequences of pairs, but in order to analyse general properties of inhomogeneous chains bio-physicists focused on the cases in which the base sequence is the realization of a sequence of random variables, that is often referred to as disorder in statistical mechanics. The PS model is limited to the case in which the two strands are of equal length and the n^{th} base of one strand can only bind with the n^{th} base of the other strand and does not allow *mismatches* or, more generally, *asymmetric loops* as explained in Figure 1.5 dans l'introduction. A less elementary model, the generalized Poland-Scheraga model (gPS) [51] allows asymmetric loops and different length strands can be considered too, see Figure 1.6.

A remarkable feature of the non disordered PS model (this corresponds to the case in which all the bases are the same: for example a strand AAA... and a second strand TTT...) is its solvable character. Notably, one can show that the model has a *denaturation* transition in the limite of infinite strand length, identify the critical point (for example, the critical temperature) and the critical behavior, that is the nature of the singularity of the free energy at the critical value. Somewhat surprisingly, also the gPS model is exactly solvable, in spite of the fact that it is considerably more complex than the PS model. This has been pointed out first in [46, 47, 72] and a mathematical treatment can be found in [51]. We stress from now that the higher complexity level of the gPS model is however reflected in a richer behavior. Notably, in the gPS model there are other phase transitions beyond the denaturation transition. Another relevant remark is that PS and gPS models contain a parameter – the *loop exponent* – that, in a mathematical or theoretical physics perspective, can be chosen arbitrarily and the critical behavior depends on this parameter. In fact in this class of models the critical exponent depends on this parameter and arbitrary critical exponents can be observed by tuning the loop exponent.

Stepping to the disordered model is not (at all) straightforward. One way to attack the problem is by looking at it as a stability issue: is the transition (and we will focus on the denaturation one) still present in the model if we introduce disorder, for example a small amount? And, if it does, what is the new critical value and is the critical behavior the same as without disorder? We refer to [48, Ch. 5] for an outline on this general very

important issue in statistical mechanics and on the renormalization group ideas that lead to the so called *Harris criterion* of disorder irrelevance. We speak of disorder relevance when the disorder, irrespectively of its strength, makes the critical behavior of the model different from the one of the non disordered model. Disorder is instead irrelevant if the two critical behaviors coincide for a small disorder strength. In the relevant (resp. irrelevant) case one can argue that applying a coarse graining procedure makes the disorder stronger (resp. weaker). Harris' idea is that disorder (ir)relevance, at least if disorder is not too strong, can be read out of the critical exponent in the non disordered model.

Harris criterion says that, if ν denotes the correlation length exponent of the non-disordered system and d the dimension, $\nu > 2/d$ implies disorder irrelevance, at least if the disorder is not too strong. One also expects disorder relevance if $\nu < 2/d$. The case $\nu = 2/d$ is dubbed marginal and deciding whether disorder is relevant or not is usually a delicate issue, even leaving aside mathematical rigor. The PS and gPS models, with their wide spectra of critical behaviors, therefore become an ideal framework for testing the validity of the physical predictions. In fact, the mathematical activity on the PS model has been very successful. Results include:

- Very complete understanding of the PS model when disorder is irrelevant [1, 50, 69, 88];
- Precise estimates on the disorder induced shift of the critical point (with respect to the annealed model) in the relevant disorder case [3, 35] and a proof of the fact that disorder does change the critical exponent (without determining the new one: this is an open problem also in the physical literature, see [37] and references therein) [31, 56];
- Determination of whether or not there is a disordered induced critical point shift, and precise estimates of this shift: this issue was controversial in the physical literature [10, 53]. It has also been shown that in absence of critical point shift also the critical exponent is unchanged by the noise. Showing that disorder does change the critical behavior when there is a critical point shift at marginality is an open issue, and determining the critical behavior in presence of disorder does not appear to be easier than attacking the same issue in the relevant case [37].

Our aim is to analyze the disordered gPS model and to understand the effect of disorder on the denaturation transition for this generalized model.

3.1.1 The generalized Poland-Scheraga model

Let us introduce the model: set $\tau = \{\tau_n\}_{n=0,1,\dots} = \{(\tau_n^{(1)}, \tau_n^{(2)})\}_{n=0,1,\dots}$ to be a bivariate renewal process with inter-arrival law $\mathrm{K}(\cdot,\cdot)$, i.e $\tau_0 = (0,0)$ and $\{\tau_{n+1} - \tau_n\}_{n=0,1,\dots}$ are identically distributed two dimensional random vectors. We assume that $\mathbf{P}(\tau_1 = (n,m)) = \mathrm{K}(n,m) = K(n+m)$ where

$$K(n) := \frac{L(n)}{n^{2+\alpha}}\,, \tag{3.1}$$

for some $\alpha \geq 0$ and slowly varying function $L(\cdot)$ and let $\mu := \mathbf{E}[\tau_1^{(1)}] = \mathbf{E}[\tau_1^{(2)}] \in (1,\infty]$. Without loss of generality, we assume that this process is persistent, i.e $\sum_{n,m} K(n+m) = 1$. We denote by $\mathbf{P}$ the law of τ.

Let $\omega := \{\omega_{n,m}\}_{n,m\in\mathbb{N}}$ be a sequence of IID centered random variables taking values in $\mathbb{R}$ with law $\mathbb{P}$. We assume that the variables $\omega_{n,m}$ are centered, have unit variance and exponential moments of all order, and set for $\beta \in \mathbb{R}$

$$Q(\beta) := \mathbb{E}[\exp(\beta\omega)] < \infty\,. \tag{3.2}$$

The choice of this disorder is discussed in detail in Section 3.1.3.

Given $\beta > 0$, $h \in \mathbb{R}$ (the pinning parameter) and $N, M \in \mathbb{N}$, we define $\mathbf{P}_{N,M}^{\beta,h,\omega}$ a measure whose Radon-Nikodym derivative w.r.t. $\mathbf{P}$ is given by

$$\frac{\mathrm{d}\mathbf{P}_{N,M,\omega}^{\beta,h}}{\mathrm{d}\mathbf{P}}(\tau) := \frac{1}{Z_{N,M,\omega}^{\beta,h}} \exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\omega_{n,m}+h)\mathbf{1}_{(n,m)\in\tau}\right)\mathbf{1}_{(N,M)\in\tau}\,, \tag{3.3}$$

where $Z_{N,M,\omega}^{\beta,h}$ is the constrained partition function (the normalization constant)

$$Z_{N,M,\omega}^{\beta,h} := \mathbf{E}\left[\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\omega_{n,m}+h)\mathbf{1}_{(n,m)\in\tau}\right)\mathbf{1}_{(N,M)\in\tau}\right]. \tag{3.4}$$

Note that the presence of $\mathbf{1}_{(N,M)\in\tau}$ in the right-hand side means that we are considering trajectories that are pinned at the endpoint of the system (at a technical level it is more practical to work with the system pinned at the endpoint, in particular for the proof of Theorem 3.1).

We also define the *free* partition function, where the endpoints are free

$$Z_{(N,M),\omega}^{f,\beta,h} = \sum_{n=0}^{N}\sum_{m=0}^{M} Z_{(n,m),\omega}^{\beta,h} \sum_{i=N-n+1}^{\infty}\sum_{j=M-m+1}^{\infty} K(i+j)\,, \tag{3.5}$$

that can be compared to the *constrained* partition function (3.4), see Lemma 3.8.

For notational convenience, we will sometimes suppress the β, h from the partition function.

One then defines the *quenched* free energy of the system. We prove the following theorem in Section 3.2.

Theorem 3.1. *For all $\gamma > 0$, $h \in \mathbb{R}$, $\beta \geq 0$ and every choice of $\{M(N)\}_{N=1,2,\dots}$ such that $\lim_N M(N)/N = \gamma$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \log Z^{\beta,h}_{N,M(N),\omega} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z^{\beta,h}_{N,M(N),\omega} =: \mathrm{F}_\gamma(\beta,h)\,, \tag{3.6}$$

where the first limit exists $\mathbb{P}(\mathrm{d}\omega)$-almost surely and in $L^1(\mathbb{P})$. Moreover the same result holds for the free model, namely:

$$\mathrm{F}_\gamma(\beta,h) = \lim_{N\to\infty} \frac{1}{N} \log Z^{f,\beta,h}_{N,M(N),\omega} \qquad \mathbb{P}(\mathrm{d}\omega)\text{-a.s. and in } L^1(\mathbb{P})\,. \tag{3.7}$$

The function $(\beta,h) \mapsto \mathrm{F}_\gamma(\beta,h)$ is convex, $h \mapsto \mathrm{F}_\gamma(\beta,h)$ and $\beta \mapsto \mathrm{F}_\gamma(\beta,h)$ are nondecreasing, and $\gamma \mapsto \mathrm{F}_\gamma(\beta,h)$ is nondecreasing and continuous.

The homogeneous model corresponds to the case $\beta = 0$ and there is no dependence in ω, so we will drop the β and ω dependence in the partition function that will be simply denoted by $Z^h_{N,M}$. The homogeneous model is *exactly solvable* and sharp estimates of $\mathrm{F}_\gamma(0,h)$ near criticality are given in [51].

Theorem 3.2. *For every $\gamma \geq 1$, when $\beta = 0$ we have $h_c(0) = 0$. Moreover there exists a slowly varying function $L_{\alpha,\gamma}(\cdot)$ such that as $h \searrow 0$ one has*

$$\mathrm{F}_\gamma(0,h) \sim L_{\alpha,\gamma}(h)\, h^{1/\min(1,\alpha)}\,. \tag{3.8}$$

Moreover, if $\sum_n n^2 K(n) < \infty$ (that is if $\mu < +\infty$), then $L_{\alpha,\gamma}(h)^{-1} = c^{-1} := \frac{1}{2}\sum_{n\geq 2} n(n-1)K(n)$.

In fact, the disordered system also presents this transition and we define the critical point

$$h_c(\beta) := \sup\{h : \mathrm{F}_\gamma(\beta,h) = 0\} = \min\{h : \mathrm{F}_\gamma(\beta,h) > 0\}\,. \tag{3.9}$$

On the other hand, we define the annealed free energy as

$$\mathrm{F}^a_\gamma(\beta,h) := \lim_{\substack{N\to\infty:\\ \frac{M(N)}{N}\to\gamma}} \frac{1}{N} \log \mathbb{E} Z_{N,M(N),\omega} = \mathrm{F}_\gamma(0, h + \log Q(\beta))\,. \tag{3.10}$$

This link with the homogeneous model and the fact that $h_c(0) = 0$ allows immediately to identify the annealed critical point:

$$h^a_c(\beta) := \min\{h : \mathrm{F}^a_\gamma(\beta,h) > 0\} = -\log Q(\beta)\,. \tag{3.11}$$

Now observe that by Jensen's inequality, we have that $\mathbb{E}\log Z_{N,M,\omega} \leq \log \mathbb{E} Z_{N,M,\omega}$ and then $\mathrm{F}^q_\gamma(\beta,h) \leq \mathrm{F}^a_\gamma(\beta,h)$. Moreover, since $\beta \mapsto \mathrm{F}_\gamma(\beta,h)$ is non-decreasing (this follows from convexity and the fact that the $\mathbb{E}\partial_\beta \log Z_{N,M,\omega}$ is zero at $\beta = 0$), we have that $\mathrm{F}_\gamma(0,h) \leq \mathrm{F}^q_\gamma(\beta,h)$. Therefore for every β we have

$$h^a_c(\beta) \leq h_c(\beta) \leq h_c(0)\,. \tag{3.12}$$

One can show, by adapting the argument of proof of [50, Th. 5.2], that the second inequality is strict for every $\beta \neq 0$. The first inequality may or may not be strict and this is an important issue which is directly linked to disorder relevance and irrelevance.

The Harris criterion predicts that disorder is irrelevant if $\nu > 2/d$. Here, Theorem 3.2 suggests that $\nu = 1/\min(1,\alpha)$, if we admit that the correlation length of the non-disordered system can be given by the reciprocal of the free energy, as it is the case for the PS model, see [49]. Here, the model is 2-dimensional, (contrary to the PS model which is 1-dimensional), and hence disorder should be irrelevant when $\nu > 1$, that is when $\alpha < 1$.

And in fact our first result states that the first inequality in (3.12) is an equality if $\alpha < 1$ and β is not too large. For the same values of β we can also show that the critical behavior is the same as for the $\beta = 0$ case (disorder irrelevance). We will moreover see that the inequality is strict for $\alpha > 1$. We interpret this critical point shift, with a certain abuse, as disorder relevance: see however the discussion in Section 3.1.3, in particular Conjecture 3.6 regarding the change in the critical behavior. We therefore prove that disorder is irrelevant if $\alpha < 1$, and relevant (only in terms of critical points) if $\alpha > 1$, confirming Harris' prediction.

3.1.2 Relevance and irrelevance of disorder

We start by showing disorder irrelevance for $\alpha < 1$.

Theorem 3.3. *Let us define $\nu := \tau \cap \tau'$, where τ and τ' are two independent copies of τ: ν is another bivariate renewal process. Assume that ν is terminating ($\alpha \in (0,1)$ is a sufficient condition, while $\alpha \in (0,1]$ is necessary, see Proposition 3.18). Then there exists $\beta_1 > 0$ (see (3.47)), such that for every $\beta \in (0,\beta_1)$ we have $h_c(\beta) = h_c^a(\beta)$, and moreover*

$$\lim_{h \searrow h_c(\beta)} \frac{\log \mathrm{F}_\gamma(\beta,h)}{\log(h - h_c(\beta))} = \frac{1}{\alpha}\,. \tag{3.13}$$

Hence, the order of the phase transition is unchanged when ν is terminating, at least when β is small enough. We prove Theorem 3.3 in Section 3.3.

We are also able to obtain a more precise result regarding the critical behavior of $\mathrm{F}_\gamma(\beta,h)$ when the disorder distribution is Gaussian, via a replica-coupling method detailed in Appendix 3.B.

Theorem 3.4. *Assume that $\omega_1 \sim \mathcal{N}(0,1)$. If ν is terminating, for every $\epsilon > 0$, there exists $\beta_0(\epsilon) > 0$ and $\Delta_0(\epsilon) > 0$ such that, for every $\beta \le \beta_0(\epsilon)$ and $0 < \Delta < \Delta_0(\epsilon)$, one has*

$$(1-\epsilon)\mathrm{F}_\gamma(0,\Delta) \le \mathrm{F}_\gamma(\beta, h_c^a(\beta)+\Delta) \le \mathrm{F}_\gamma(0,\Delta)\,. \tag{3.14}$$

On the other hand, when $\alpha > 1$, we are able to show that the quenched and annealed critical points differ, and give a lower bound on the critical point shift.

Theorem 3.5. *For $\alpha > 1$ we have $h_c(\beta) > h_c^a(\beta)$ for every $\beta > 0$. Moreover, we have that for every $\varepsilon > 0$, there exists $\beta_\varepsilon > 0$ such that for any $\beta \leqslant \beta_\varepsilon$*

$$h_c(\beta) - h_c^a(\beta) \geq \Delta_\beta^\varepsilon := \begin{cases} \beta^{\frac{2\alpha}{\alpha-1}+\varepsilon} & \text{if } \alpha \in (1,2]\,, \\ \beta^4 |\log \beta|^{-(1+\varepsilon)} & \text{if } \alpha > 2\,. \end{cases} \tag{3.15}$$

Moreover

$$h_c(\beta) - h_c^a(\beta) \leq \widetilde{L}(1/\beta)\beta^{\frac{2\alpha}{\alpha-1}\vee 4}\,, \tag{3.16}$$

where $\widetilde{L}(\cdot)$ is a slowly varying function.

We add that $\beta \mapsto h_c(\beta) - h_c^a(\beta)$ is a non decreasing function of β: this result can be proven by the exact same procedure as the one used to prove Proposition 6.1 in [54]. It is to be interpreted that disorder relevance is *non-decreasing* in β.

3.1.3 On the results, perspectives and related work

On the main theorems

A two replica computation plays a central role in the proof of Theorem 3.3 and in the proof of (3.16) of Theorem 3.3. And when dealing with two replica computation the intersection renewal ν emerges naturally, like in the PS model. In the PS context we now know that disorder is irrelevant (for small values of β) if and only if the intersection renewal is terminating. For the gPS model our results go in the same direction, but it is not sharp in the marginal case $\alpha = 1$ where we do not cover all the cases. The difficulty is already at the renewal process level: a necessary and sufficient condition for persistence of ν when $\alpha = 1$ – it depends therefore on the slowly varying function $L(\cdot)$ – remains to be found, and seems out of reach without further understanding of bivariate renewal processes, see Remark 3.19.

The proof of (3.15) is based on coarse graining techniques and fractional moment method: we have chosen to adapt the method proposed in [35] and the difficulties in this generalization come from dealing with the richness of a multidimensional path with respect to the one dimensional structure of the PS models. A keyword for these difficulties is *off-diagonal estimates*. It can certainly be improved in the direction of getting rid of the ε in the exponent for $\alpha \in (1,2]$ and of the logarithmic term in the case $\alpha > 1$ by using more sophisticated coarse graining techniques (see [50, Ch. 6] and references therein). And one could probably shoot also for sharp estimates, like in [10], but the estimates are technically rather demanding already to obtain (3.15): we have chosen to stick to these simplified non-optimal (but almost optimal) bounds because sharper results would have required a substantially heavier argument of proof. Such a sharp treatment would make sense if we were to treat also the (marginal) case $\alpha = 1$ in which, as we have explained,

we still have some more basic difficulties at the level of the intersection of bivariate renewals. We mention that, in analogy with the PS model, we expect that the necessary and sufficient condition for a critical point shift is the persistence of the intersection renewal $\nu = \tau \cap \tau'$. Indeed, the techniques developed in [10, 11] should transfer to this model, but suffer from the lack of estimates for bivariate renewals in the $\alpha = 1$ case.

Discussion on the presence of a *smoothing* phenomenon

Of course a fully satisfactory result on disorder relevance would include showing that the critical exponent is modified by the disorder. We do not have such a result yet, but we make one observation and formulate a conjecture.

The observation is that Theorem 3.3 may appear at first surprising in view of the smoothing inequality [31, 56] for PS models that ensures that the free energy exponent cannot be smaller than two in presence of disorder: for the gPS model we go down to one, since in (3.13) we can choose α arbitrarily close to 1. The reason of the difference is that the PS model is 1-dimensional whereas the gPS model is 2-dimensional: Harris criterion tells that disorder should be irrelevant if $\nu > 2$ for the PS model, and $\nu > 1$ for the gPS model. In the gPS model, the irrelevant disorder regime therefore holds even if $\nu (= 1/\min(1, \alpha))$ is arbitrarily close to 1, and one should therefore not hope for a general smoothing inequality valid whatever α is.

It is however worthwhile attempting to sketch the argument in [56], in the simplified set-up of Gaussian charges [48, Ch. 5, Sec. 4]. This is useful both to understand were the argument fails and because we can realize that a suitable generalization of the argument naturally leads to a conjecture that we state just below.

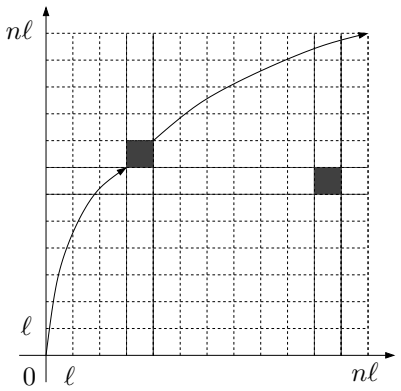

FIGURE 3.1. A schematic view of the coarse graining procedure proposed for a smoothing inequality. The environment is considered in independent blocks of linear size ℓ, called ℓ-boxes. An ℓ-box is good (shadowed blocks in the figure) if the pinned partition function in this block is larger than expected, in the sense that the partition function grows at an exponential rate that is larger than the free energy of the system. The good ℓ-boxes will therefore be rare, but we can choose n such that in a system of linear size $n\ell$ with positive probability there will be at least one ℓ-box. A lower bound on the partition function follows by the limitation to trajectories that visit only a given good ℓ-box (say, the one with largest vertical coordinate).

The argument [56] is based on introducing a coarse graining scale $\ell \in \mathbb{N}$ and considering

the environment in terms of ℓ-boxes, see Figure 3.1. We argue for the case $\gamma = 1$ and we consider the system at criticality, that is $h = h_c(\beta)$:

1. A good ℓ-box is a box for which the pinned partition function in this box – i.e. we consider the model that is pinned at the south-west corner as well as the north-east corner of the ℓ-box – is larger than $\exp\left(\frac{1}{2}\ell\, \mathrm{F}_1(\beta, h+\delta)\right)$, with $\delta > 0$. For $\ell \to \infty$ this is a rare event. The probability of such an event can be estimated from below by shifting the environment of δ/β, that is we replace $\omega_{i,j}$ with $\omega_{i,j} + \delta/\beta$, and by performing a relative entropy estimate [50, Ch. 5]. This shows that the probability of such a rare event is at least $\exp(-\delta^2\ell^2/(2\beta^2))$: note the ℓ^2 term, with respect to ℓ in PS case [56].

2. We then make a lower bound on the partition function of the system by neglecting the renewal trajectories that visit ℓ-boxes that are not good and permit only trajectories that enter ℓ-boxes passing through the south-west corner and exiting by passing through the north-east corner.

The trajectories are therefore alternated jumps to a good box, *visit* of the box, and then a new jump to another good box. Jumps are long because good boxes are rare. The analysis in [56] is ultimately reduced to see what happens in one *jump and visit*: by exploiting super-additivity one can even just choose $N = n\ell$ such that there is (say, with probability at least 1/2), at least one good box in the system (like it is done in [9]). We therefore see that we need $n^2 \exp(-\delta^2\ell^2/(2\beta^2)) \approx 1$, so that $n \approx \exp(-\delta^2\ell^2/\beta^2)$ and, with this level of precision, jumping to enter such a box costs $K(n\ell) = (n\ell)^{-(2+\alpha)}$ (let us consider the case in which $L(\cdot)$ is trivial, that is asymptotic to a constant, but the computation goes through in the same way also in the general case). In the box there will be a contribution $\exp\left(\frac{1}{2}\ell\, \mathrm{F}_1(\beta, h+\delta)\right)$ so the net contribution to the logarithm of the partition function divided by the size $n\ell$ of the system is

$$\frac{1}{n\ell}\left(\log K\left(\exp(-\delta^2\ell^2/\beta^2)\ell\right) + \frac{\ell}{2}(\mathrm{F}_1(\beta, h+\delta))\right) \geq \frac{1}{n}\left(-\frac{c}{\beta^2}\delta^2\ell + \frac{1}{2}\mathrm{F}_1(\beta, h+\delta)\right), \tag{3.17}$$

with c a positive constant that we have left implicit (it depends on more accurate computations, and can be in principle just reduced to $2+\alpha$).

Now let us choose $h = h_c(\beta)$. So the argument we just outlined goes in the direction of saying that

$$0 = \mathrm{F}_1(\beta, h_c(\beta)) \geq \frac{1}{n}\left(-\frac{c}{\beta^2}\delta^2\ell + \mathrm{F}_1(\beta, h_c(\beta)+\delta)\right), \tag{3.18}$$

so that

$$\mathrm{F}_1(\beta, h_c(\beta)+\delta) \leq \frac{c}{\beta^2}\delta^2\ell\,. \tag{3.19}$$

At this stage choosing ℓ arbitrarily large is of no help. The steps we have performed up to now require $\ell\delta \gg 1$ (say, larger than a large constant, so that the good boxes we have

chosen are really sparse). On the other hand we need to have chosen the size of the boxes so that $Z^{\beta,h}_{\ell,\ell,\omega} \geq \exp(\ell(\mathrm{F}_1(\beta, h_c(\beta)+\delta))/2)$. This is a delicate issue, but it definitely appears that for this to hold, $\ell \mathrm{F}_1(\beta, h_c(\beta)+\delta)$ needs to be sufficiently large (say, larger than a suitable constant): see for example the discussion on the notion of correlation length given in [50, Ch. 2] and references therein, notably [57], where the correlation length is identified by the reciprocal of the free energy. But if ℓ is (a constant times) $1/\mathrm{F}_1(\beta, h_c(\beta)+\delta)$ then from (3.19) we obtain

$$\mathrm{F}_1(\beta, h_c(\beta)+\delta) \leq C\delta\,, \tag{3.20}$$

for some $C > 0$. But such a bound is trivial: it holds with $C = 1$ just because the contact density cannot exceed one! On the other hand, as we have already pointed out, we could not have hoped for a better bound valid for any $\alpha > 0$.

In spite of the fact that it leads to a trivial result, we insist that the argument we have just outlined can be made rigorous: the delicate step is the last one, where one has to use arguments developed in [57]. It can be therefore taken as a starting point to push things farther. Indeed, it appears useless to modify the environment in the whole ℓ-box, at least if $\alpha > 1$. In fact if $\alpha > 1$ one can show that if $q > 1/\max(\alpha, 2)$

$$\lim_{N\to\infty} \mathbf{P}\left(\tau \cap [0,N]^2 \subset \{(i,j) \in \mathbb{Z}^2 : |i-j| \leq N^q\right) = 1\,. \tag{3.21}$$

We can then consider modifying only the environment that is close to the diagonal, that is in a subset of the ℓ-box with $|i-j| \leqslant \ell^q$. This would improve the lower bound on the probability of a good ℓ-box to $\exp(-cst.\,\delta^2\ell^{q+1})$, and (3.19) would become

$$\mathrm{F}_1(\beta, h_c(\beta)+\delta) \leqslant \frac{c}{\beta^2}\delta^2\ell^q.$$

Taking ℓ a constant times $1/\mathrm{F}_1(\beta, h)$ as in the argument leading (3.20) and q arbitrarily close to $1/\min(\alpha, 2)$ supports the following:

Conjecture 3.6. *For every $\alpha > 0$ and every $\beta > 0$*

$$\limsup_{h\searrow h_c(\beta)} \frac{\log \mathrm{F}_1(\beta, h)}{\log(h - h_c(\beta))} \geq \begin{cases} \frac{2\alpha}{\alpha+1} & \text{for } \alpha \in (1,2)\,, \\ \frac{4}{3} & \text{for } \alpha \geq 2\,. \end{cases} \tag{3.22}$$

An important modeling issue: the choice of the disorder

There is no doubt that the first disorder that comes to mind when thinking of DNA modeling is not the one we have used. One would rather choose $\omega_{i,j} = f(\omega_i, \omega_j)$ for a suitable choice of a function f and a sequence $\{\omega_j\}_{j=1,2,\ldots}$ of random variables (let us say IID for simplicity, but if we want to stick to DNA problems very closely it appears that some sort of strongly correlated sequence may be more appropriate [74]). For example, we could choose ω_j taking only two values e_{AT} and e_{GC} and then make a choice for f that

reflects the fact that AT bounds are weaker than GC bounds, and that all other possible bounds are even weaker. Even restricting to $\{\omega_j\}_{j=1,2,\dots}$ that is IID this model is highly non trivial (gPS model with this type of disorder has been considered at a numerical level in [46, 47], see also [41, 83] for related work). But one could also choose to consider the binding of two sequences that are not complementary (the case considered in [72] goes in this direction, even if only heuristics and numerics are presented): choose for example two independent sequences $\{\omega_j^{(1)}\}_{j=1,2,\dots}$ and $\{\omega_j^{(2)}\}_{j=1,2,\dots}$ and use $\omega_{i,j} = f(\omega_i^{(1)}, \omega_j^{(2)})$. This is somewhat closer to what we are using, but it is still very difficult to deal with (the problem is in any case due to correlations in the disorder, which can be dealt with in some cases, see e.g. [8, 12]. Our choice is in a sense a toy choice, but we stress that is conceptually similar to the simplification made for example in [24] in the RNA context. Moreover it recovers importance once we leave somewhat the DNA context and focus rather on moving toward understanding mathematically Harris' theory of disorder (ir)relevance.

We point out also that this disordered version of the gPS model gives really a bridge between pinning model and directed polymers in random environment [32, 68], in particular, to the long range directed polymer [32, 94]. Moreover a different class of two strand polymer problems is the case treated in [13, 15, 16].

Open questions and perspectives

Several natural issues remain open: let us list some of them.

1. Proving a smoothing inequality, showing thus disorder relevance in the original sense of Harris, for $\alpha > 1$ (see Conjecture 3.6).

2. What is the effect of disorder on the other phase transitions? Here we have addressed only the denaturation transition, but in [51] other transitions are shown to exist. Do they withstand the introduction of disorder? If so, does the corresponding critical behavior differ from the homogeneous case? This is the question (quickly) addressed [72] where a rather bold conjecture is set forth.

3. We have dealt only with free energy estimates, but, like for the standard PS model, obtaining precise estimates on the gPS process is not a straightforward issue, see [50, Ch.8] and references therein.

4. Dealing with the marginal case ($\alpha = 1$) is open, in particular because of the incomplete understanding of the bivariate $\alpha = 1$ renewal, as pointed out before.

Organization of the rest of the work

The issues of existence and self-averaging of the free energy, hence the proof of Theorem 3.1, are treated in Section 3.2. In Section 3.3 we prove Theorem 3.3, as well as the upper bound

(3.16) of Theorem 3.5. The rest of the Theorem 3.5 is proven in Section 3.4. We collect in Appendix 3.A a number of statements and proofs about bivariate renewals. Finally, in Appendix 3.B one can find the proof of Theorem 3.4.

3.2 Free Energy: existence and properties

In this section we often assume $\gamma \in \mathbb{Q}$: in this case we write it as $\gamma = p/q$ with p and q relatively prime positive integer numbers.

Proposition 3.7. *For every $\gamma > 0$ and every $\{M(N)\}_{N=1,2,\dots}$ such that $\lim_{N\to\infty} M(N)/N = \gamma$ we have that*

$$\lim_{N\to\infty} \frac{1}{N} \log Z^{\beta,h}_{N,M(N),\omega} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z^{\beta,h}_{N,M(N),\omega} =: \mathrm{F}_\gamma(\beta,h)\,, \tag{3.23}$$

where the first limit is meant $\mathbb{P}(\mathrm{d}\omega)$-a.s. and in $L^1(\mathbb{P})$. $\mathrm{F}_\gamma(\cdot,\cdot)$ is convex and $\mathrm{F}_\gamma(\beta,\cdot)$ is non decreasing, as well as $\mathrm{F}_\gamma(\cdot,h)$ is non decreasing on the positive semi-axis and non increasing in the negative one. Moreover if $\gamma = \frac{p}{q} \in \mathbb{Q}$

$$\mathrm{F}_\gamma(\beta,h) = \sup_{N:\, \frac{N}{q}\in\mathbb{N}} \frac{1}{N} \mathbb{E} \log Z^{\beta,h}_{N,\gamma N,\omega}\,. \tag{3.24}$$

Finally we have the bound: for every $\gamma_2 \geq \gamma_1 > 0$

$$\mathrm{F}_{\gamma_1}(\beta,h) \leq \mathrm{F}_{\gamma_2}(\beta,h) \leq \frac{\gamma_2}{\gamma_1} \mathrm{F}_{\gamma_1}(\beta,h)\,, \tag{3.25}$$

which implies that $\gamma \mapsto \mathrm{F}_\gamma(\beta,h)$ is locally Lipschitz (hence continuous).

Proof. The proof is divided into several steps:

1. We first show that for $\gamma \in \mathbb{Q}$ the sequence made up with $\log Z_{N,\gamma N,\omega}$ with $N/q \in \mathbb{N}$ is super-additive in an ergodic sense, which implies the existence of the free energy limit (3.23) along this subsequence.

2. The restriction $\gamma N \in \mathbb{N}$ is then removed by a direct estimate, for what concerns the existence of the free energy limit, still with $\gamma \in \mathbb{Q}$.

3. We then prove a comparison estimate between $Z_{N,\gamma_1 N,\omega}$ and $Z_{N,\gamma_2 N,\omega}$ and use it to establish the existence of the free energy limit for $Z_{N,\gamma N,\omega}$, every $\gamma > 0$.

4. The same comparison estimate yields also (3.25) and the fact that one can take the limit along an arbitrary sequence satisfying $M \sim \gamma N$, for $N \to \infty$.

5. We prove the convexity and monotonicity statements.

Step 1. With $\gamma = p/q$ set $\mathcal{Z}_j(\omega) := Z_{jq,jp,\omega}$. Then one directly sees that

$$\mathcal{Z}_{j_1+j_2}(\omega) \geq \mathcal{Z}_{j_1}(\omega)\mathcal{Z}_{j_2}(\Theta_{j_1q,j_1p}\omega)\,, \tag{3.26}$$

where $(\Theta_{q,p}\omega)_{n,m} = \omega_{q+n,p+m}$. Since ω is an IID sequence of L^1 random variables, it is straightforward to see that $\log \mathcal{Z}_j \in L^1(\mathbb{P})$ and that $\sup_n \frac{1}{n}\mathbb{E}|\log \mathcal{Z}_j| < \infty$, so that we see that $\{-\log \mathcal{Z}_j(\omega)\}_{j=1,2,\dots}$ satisfies the hypotheses of Kingman Sub-Additive Theorem (see for example [48, Sec. A.7]). Hence $\{\frac{1}{j}\log \mathcal{Z}_j(\omega)\}_{j=1,2,\dots}$ converges $\mathbb{P}(\,\mathrm{d}\omega)$-a.s. and in $L^1(\mathbb{P})$. Moreover (3.26) directly tells us that $\{\mathbb{E}\log \mathcal{Z}_j(\omega)\}_{j=1,2,\dots}$ is super-additive, so that $\lim_{j\to\infty}\frac{1}{j}\mathbb{E}\log \mathcal{Z}_j(\omega) = \sup_{j\in\mathbb{N}}\frac{1}{j}\mathbb{E}\log \mathcal{Z}_j(\omega)$. This establishes (3.24), and also (3.23) but only for $M(N) = \gamma N$ with $\gamma \in \mathbb{Q}$ and along the subsequence satisfying $\gamma N \in \mathbb{N}$.

Step 2. The restriction to $\gamma N \in \mathbb{N}$ can be removed by observing that we can write $N = jq+r$, with $r \in \{0, 1, \dots, q-1\}$ ($\gamma = p/q$), and for $r \neq 0$

$$\begin{aligned} Z_{N,\lfloor\gamma N\rfloor,\omega} &\geq Z_{jq,jp,\omega}\exp(\beta\omega_{N,\lfloor\gamma N\rfloor}+h)K\left(r+\left\lfloor jp+\frac{p}{q}r\right\rfloor - jp\right) \\ &\geq c(p,q)\exp(\beta\omega_{N,\lfloor\gamma N\rfloor}+h)Z_{jq,jp,\omega}\,, \end{aligned} \tag{3.27}$$

where $c(p,q) > 0$.

In the same way

$$\begin{aligned} Z_{(j+1)q,(j+1)p,\omega} &\geq Z_{N,\lfloor\gamma N\rfloor,\omega}\exp(\beta\omega_{(j+1)q,(j+1)p}+h)K\left(q-r+(j+1)p-\left\lfloor jp+\frac{p}{q}r\right\rfloor\right) \\ &\geq c(p,q)\exp(\beta\omega_{(j+1)q,(j+1)p}+h)Z_{N,\lfloor\gamma N\rfloor,\omega}\,, \end{aligned} \tag{3.28}$$

possibly redefining $c(p,q) > 0$. From (3.27) and (3.28) one easily removes the restriction to $\gamma N \in \mathbb{N}$ and establishes (3.23) for $\gamma \in \mathbb{Q}$.

Step 3. We now establish (3.23) for $M = \lfloor\gamma N\rfloor$ for an arbitrary $\gamma > 0$, by proving the announced comparison bounds, upper and lower.

The upper bound is more general: if $M_2 > M_1$ and if there exists $c > 0$ such that $M_2 \leq cN$ we see that

$$\begin{aligned} Z_{N,M_1,\omega} &= \sum_{\substack{n\in\{0,\dots,N-1\}\\ m\in\{0,\dots,M_1-1\}}} Z_{n,m,\omega}K(N-n+M_1-m)\exp\left(\beta\omega_{N,M_1}+h\right) \\ &\leq c_K N^{c_K}\exp\left(\beta(\omega_{N,M_1}-\omega_{N,M_2})\right) \\ &\qquad \sum_{\substack{n\in\{0,\dots,N-1\}\\ m\in\{0,\dots,M_1-1\}}} Z_{n,m,\omega}K(N-n+M_2-m)\exp\left(\beta\omega_{N,M_2}+h\right) \\ &\leq c_K N^{c_K}\exp\left(\beta(\omega_{N,M_1}-\omega_{N,M_2})\right)Z_{N,M_2,\omega}\,, \end{aligned} \tag{3.29}$$

where in the first inequality we have used that $K(\cdot)$ is regularly varying and that $M_2 \leq cN$ to see that there exists $c_K > 0$ such that

$$\frac{K(N-n+M_1-m)}{K(N-n+M_2-m)} \leq c_K N^{c_K}\,, \tag{3.30}$$

for every N. For the second inequality we have relaxed the constrained $m < M_1$ to $m < M_2$.

On the other hand, we prove a comparison lower bound only for M of the form $\lfloor \gamma N \rfloor$. Let us choose $\gamma_2 > \gamma_1 > 0$. Note that for

$$N' := \left\lfloor \frac{\gamma_1}{\gamma_2} N \right\rfloor - \left\lceil \frac{2}{\gamma_2} \right\rceil , \tag{3.31}$$

we have $\lfloor \gamma_2 N' \rfloor + 1 \le \lfloor \gamma_1 N \rfloor$ so that

$$\begin{aligned} Z_{N,\lfloor \gamma_1 N \rfloor,\omega} &\ge K \left(N - N' + \lfloor \gamma_1 N \rfloor - \lfloor \gamma_2 N' \rfloor \right) \exp\left(\beta \omega_{N,\lfloor \gamma_1 N \rfloor} + h \right) Z_{N',\lfloor \gamma_2 N' \rfloor,\omega} \\ &\ge (c_K N^{c_K})^{-1} \exp\left(\beta \omega_{N,\lfloor \gamma_1 N \rfloor} + h \right) Z_{N',\lfloor \gamma_2 N' \rfloor,\omega} , \end{aligned} \tag{3.32}$$

possibly changing the value of $c_K > 0$.

We now choose $0 < \gamma_1 < \gamma_2 \in \mathbb{Q}$ and (3.29) implies that $\mathbb{P}(\,\mathrm{d}\omega)$-a.s.

$$\limsup_{N \to \infty} \frac{1}{N} \log Z_{N,\lfloor \gamma_1 N \rfloor,\omega} \le \mathrm{F}_{\gamma_2}(\beta, h) , \tag{3.33}$$

and (3.32) implies that $\mathbb{P}(\,\mathrm{d}\omega)$-a.s.

$$\liminf_{N \to \infty} \frac{1}{N} \log Z_{N,\lfloor \gamma_1 N \rfloor,\omega} \ge \lim_{N \to \infty} \frac{1}{N'} \log Z_{N',\lfloor \gamma_2 N' \rfloor,\omega} = \frac{\gamma_1}{\gamma_2} \mathrm{F}_{\gamma_2}(\beta, h) , \tag{3.34}$$

and the proof of (3.23) is achieved in the $\mathbb{P}(\,\mathrm{d}\omega)$-a.s. sense for $M(N) = \lfloor \gamma_1 N \rfloor$, by choosing a sequence of values for γ_2 converging to γ_1, defining thus $F_{\gamma_1}(\beta, h)$ also by this limit procedure. Note that a byproduct is that (3.25) holds, hence $\mathrm{F}_\gamma(\beta, h)$ is non decreasing and (locally) Lipschitz continuous. To upgrade (3.23) to the $L^1(\mathbb{P})$ sense one simply applies the expectation $\mathbb{E}[\,\cdot\,]$ to (3.32) and (3.29) so that one obtains $\lim_{N\to\infty}(1/N)\mathbb{E} \log Z_{N,\lfloor \gamma N \rfloor,\omega} = F_\gamma(\beta, h)$ for every $\gamma > 0$, and the first limit in (3.23) holds in the $L^1(\mathbb{P})$ sense by Schéffé's Lemma.

Step 4. The generalization to a sequence $M(N) \sim \gamma N$ is just made by observing that given arbitrary $\gamma_1 < \gamma_2$ with $\gamma \in (\gamma_1, \gamma_2)$ for N_0 sufficiently large we have $\lfloor \gamma_1 N \rfloor < M(N) < \lfloor \gamma_2 N \rfloor$ for every $N \ge N_0$. At this point we can apply the comparison bounds like in the previous step and conclude by an approximation procedure.

Step 5. The function $(\beta, h) \mapsto \mathrm{F}_\gamma(\beta, h)$ is convex because it is the limit of a sequence of convex functions. Monotonicity in h for β fixed is also evident from the finite N expression. The fact that $\beta \mapsto \mathrm{F}_\gamma(\beta, h)$ is non increasing for $\beta \le 0$ and non decreasing for $\beta \ge 0$ follows from convexity and the fact that $\partial_\beta \mathbb{E} \log Z^{\beta,h}_{N,M,\omega} = 0$ (by direct computation, since the ω variables are centered), so $\partial_\beta \mathrm{F}_\gamma(\beta, h)|_{\beta=0} = 0$. This completes the proof of Proposition 3.7. □

We compare now the constrained and the free partition function:

Lemma 3.8. *There exists C such that for every $N, M \in \mathbb{N}$ and $\alpha_+ > \alpha$*

$$\begin{aligned} &Z^c_{N,M,\omega} \le Z^f_{N,M,\omega} \le \\ &Z^c_{N,M,\omega} \left(1 + C(N+M)^{3+\alpha_+} \exp(-\beta \omega_{N,M}) \sup_{\substack{1 \le n \le N \\ 1 \le m \le M}} \{ \exp(\beta \omega_{n,M}), \exp(\beta \omega_{N,m}) \} \right) . \end{aligned} \tag{3.35}$$

Proof. For $N, M \geq 1$, we write

$$Z^f_{N,M,\omega} = Z^c_{N,M,\omega} + \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} Z^c_{n,m,\omega} \sum_{i=N-n+1}^{\infty}\sum_{j=M-m+1}^{\infty} K(i+j) + \sum_{n=0}^{N-1} Z^c_{n,M,\omega} \sum_{i=N-n+1}^{\infty}\sum_{j\geq 1} K(i+j) + \sum_{m=0}^{N-1} Z^c_{N,m,\omega} \sum_{i\geq 1}\sum_{j=M-m+1}^{\infty} K(i+j)\,, \quad (3.36)$$

from which we see that the lower bound holds and observe that

$$\begin{aligned}
&\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} Z^c_{n,m,\omega} \sum_{i=N-n+1}^{\infty}\sum_{j=M-m+1}^{\infty} K(i+j) \\
&\leq C_1 (N+M)^{2+\alpha_+} \exp(-\beta\omega_{N,M} - h) \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} Z^c_{n,m,\omega} K(M+N-n-m) \exp(\beta\omega_{N,M} + h) \\
&\leq C_1 (N+M)^{2+\alpha_+} Z^c_{N,M,\omega} \exp(-\beta\omega_{N,M} - h)\,,
\end{aligned} \quad (3.37)$$

for any $\alpha_+ > \alpha$.

Using the fact that if $n < N$, there exists c_2 such that

$$Z^c_{n,M,\omega} \leq c_2 N^{2+\alpha_+} Z^c_{N,M,\omega} \exp\left(\beta\omega_{n,M} - \beta\omega_{N,M}\right)\,, \quad (3.38)$$

then we obtain

$$\begin{aligned}
\sum_{n=0}^{N-1} Z^c_{n,M,\omega} &\leq c_2 N^{2+\alpha_+} \exp(-\beta\omega_{N,M})\, Z^c_{N,M,\omega} \left(\sum_{n=0}^{N-1} \exp(\beta\omega_{n,M})\right) \\
&\leq C_2 N^{3+\alpha_+} \sup_{1\leq n\leq N} \{\exp(\beta\omega_{n,M})\} \exp(-\beta\omega_{N,M})\, Z^c_{N,M,\omega}\,.
\end{aligned} \quad (3.39)$$

The same for the last term in (3.36), there exists $C_3 > 0$

$$\sum_{m=0}^{M-1} Z^c_{N,m,\omega} \leq C_3 M^{3+\alpha_+} \sup_{1\leq m\leq M} \{\exp(\beta\omega_{N,m})\} \exp(-\beta\omega_{N,M})\, Z^c_{N,M,\omega}\,. \quad (3.40)$$

The proof is therefore complete. □

The proof of (3.7) follows immediately from Lemma 3.8.

This notation is used later: for positive integers verifying $a_1 < a_2$ and $b_1 < b_2$, we define the partition function of the system on $[a_1, a_2] \times [b_1, b_2]$ by

$$Z_{(a_1,b_1),(a_2,b_2),\omega} := \mathbf{E}\left[\exp\left(\sum_{n=a_1+1}^{a_2}\sum_{m=b_1+1}^{b_2} (\beta\omega_{n,m} + h)\mathbf{1}_{(n,m)\in\tau}\right)\mathbf{1}_{(a_2,b_2)\in\tau} \,\middle|\, (a_1,b_1) \in \tau\right], \quad (3.41)$$

with the convention that $Z_{(a_1,b_1),(a_1,b_1),\omega} = 1$ and $Z_{(a_1,b_1),(a_1,b_2),\omega} = Z_{(a_1,b_1),(a_2,b_1),\omega} = 0$.

3.3 Upper bound on the critical point shift

The arguments in this section follow the line of proof of H. Lacoin in [69], and is mainly based on a second moment computation. We start with some preliminary results.

Proposition 3.9. *If* $\{Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega}\}_N$ *is* uniformly integrable *there exists* $\zeta > 0$ *such that for every sequence of events* $\{A_N\}_{N=1,2,\dots}$ *satisfying* $\lim_N \mathbf{P}(A_N) = 0$ *there exists* $N_0 \in \mathbb{N}$ *such that*

$$\inf_{N\geq N_0} \mathbb{P}\left(\mathbf{P}^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega}(A_N) \leq \frac{1}{2} \text{ and } Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega} > \frac{1}{2}\right) \geq \zeta\,. \tag{3.42}$$

Proof. We set $h = h_c^a(\beta)$, and use $Z^f_{N,M,\omega}$ instead of $Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega}$ for notational convenience. It is sufficient to prove that there exists $\zeta > 0$ such that

$$\inf_N \mathbb{P}\left(Z^f_{N,M,\omega} > \frac{1}{2}\right) \geq 2\zeta\,. \tag{3.43}$$

and

$$\lim_{N\to\infty} \mathbb{P}\left(\mathbf{P}^f_{N,M,\omega}(A_N) > \frac{1}{2} \text{ and } Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega} > \frac{1}{2}\right) = 0\,. \tag{3.44}$$

Since $\mathbb{E}[Z^f_{N,M,\omega}] = 1$, and because $\{Z^f_{N,M,\omega}\}_N$ is uniformly integrable, then (3.43) follows immediately from [50, Lemma 4.6].

For (3.44) instead we observe that the Fubini-Tonelli Theorem imples

$$\begin{aligned}\mathbb{E}\left[Z^f_{N,M,\omega}\mathbf{P}^f_{N,M,\omega}(A_N)\right] &= \mathbb{E}\mathbf{E}\left[\exp\left(\sum_{n=1}^N\sum_{m=1}^M(\beta\omega_{n,m} - \log Q(\beta))\delta_{n,m}\right)\mathbf{1}_{A_N}\right]\\ &= \mathbf{P}(A_N)\,,\end{aligned} \tag{3.45}$$

with $\delta_{n,m} = \mathbf{1}_{(n,m)\in\tau}$, so $\lim_N \mathbb{E}\left[Z^f_{N,M,\omega}\mathbf{P}^f_{N,M,\omega}(A_N)\right] = 0$ and (3.44) follows because

$$\begin{aligned}\mathbb{P}\left(Z^f_{N,M,\omega} > \frac{1}{2} \text{ and } \mathbf{P}^f_{N,M,\omega}(A_N) > \frac{1}{2}\right) &= \mathbb{P}\left(Z^f_{N,M,\omega}\mathbf{1}_{\left\{\mathbf{P}^f_{N,M,\omega}(A_N)>\frac{1}{2}\right\}} > \frac{1}{2}\right)\\ &\leq 2\mathbb{E}\left[Z^f_{N,M,\omega}\mathbf{1}_{\left\{\mathbf{P}^f_{N,M,\omega}(A_N)>\frac{1}{2}\right\}}\right] \leq 4\mathbb{E}\left[Z^f_{N,M,\omega}\mathbf{P}^f_{N,M,\omega}(A_N)\right].\end{aligned} \tag{3.46}$$

□

We now prove that $\{Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega}\}_N$ is uniformly integrable (and this holds for an arbitrary choice of $M(N)$) provided that the intersection renewal $\nu = \tau \cap \tau'$ is transient, where τ and τ' are two independent copies of τ, and β is small enough. Let us point out that, since ν is a transient renewal then the total number $|\nu|$ of renewal points (except the origin), that is $|\nu| = \sum_{(n,m)\in\mathbb{N}^2}\widetilde{\delta}_{n,m}$ with $\widetilde{\delta}_{n,m} = \mathbf{1}_{(n,m)\in\nu}$, is a geometric random variable of parameter $\mathbf{P}^{\otimes 2}(\nu_1 < \infty)$, where $\nu_1 < \infty$ simply means that both components of ν_1 are finite. This in particular implies that $\mathbf{P}^{\otimes 2}(\nu_1 < \infty) = 1/\mathbf{E}^{\otimes 2}[|\nu|]$. Moreover it is straightforward to see that $\mathbf{E}^{\otimes 2}[|\nu|] = \sum_{n,m}\mathbf{P}((n,m)\in\tau)^2$.

Lemma 3.10. *If $\nu := \tau \cap \tau'$ is transient, then defining*

$$0 < \beta_1 := \sup\left\{ \beta : \log Q(2\beta) - 2\log Q(\beta) < -\log \mathbf{P}^{\otimes 2}(\nu_1 < \infty)\right\}, \tag{3.47}$$

we have that for every $\beta \in (0, \beta_1)$ the sequence $\{Z^{f,\beta,h_c^a(\beta)}_{N,M(N),\omega}\}_N$ is bounded in $L^2(\mathbb{P})$, and therefore uniformly integrable.

Proof. We write $M = M(N)$ and use the Fubini-Tonelli Theorem to compute the second moment

$$\begin{aligned}\mathbb{E}\left[\left(Z^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\right)^2\right] &= \mathbf{E}^{\otimes 2}\left[\mathbb{E}\left[\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\omega_{n,m} - \log Q(\beta))(\delta_{n,m} + \delta'_{n,m})\right)\right]\right] \\ &= \mathbf{E}^{\otimes 2}\left[\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}(\log Q(2\beta) - 2\log Q(\beta))\widetilde{\delta}_{n,m}\right)\right].\end{aligned} \tag{3.48}$$

The sequence $\{Z^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\}_N$ is bounded in $L^2(\mathbb{P})$ if

$$\mathbf{E}^{\otimes 2}\left[\exp\left(\sum_{n,m=1}^{\infty}(\log Q(2\beta) - 2\log Q(\beta))\widetilde{\delta}_{n,m}\right)\right] < \infty. \tag{3.49}$$

Since $|\nu|$ is a geometric random variable of parameter $\mathbf{P}^{\otimes 2}(\nu_1 < \infty)$, (3.49) holds if

$$\log Q(2\beta) - 2\log Q(\beta) < -\log \mathbf{P}^{\otimes 2}(\nu_1 < \infty). \tag{3.50}$$

□

We now have all the ingredients to complete the proof of Theorem 3.3.

Proof of Theorem 3.3. In view of what we want to prove and of Proposition 3.7, notably the explicit continuity estimate (3.25), it suffices to establish the result for $\gamma \in \mathbb{Q}$ and for $M = \lfloor \gamma N \rfloor$, which we shall assume till the end of the proof, even if this explicit choice is used in full only at the very end.

Because of Lemma 3.10, we have that the sequence $\{Z^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\}_N$ is uniformly integrable for $\beta < \beta_1$. Now for all $0 < \eta < \alpha$ we set

$$A_N := \{|\tau \cap ((0,N] \times (0,M])| \le N^\eta\}. \tag{3.51}$$

From Lemma 3.21, we have that $\lim_N \mathbf{P}(A_N) = 0$. Observe also that

$$\begin{aligned}Z^{f,\beta,h_c^a(\beta)+h}_{N,M,\omega} &= Z^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\mathbf{E}^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\left[\exp\left(h|\tau \cap ((0,N] \times (0,M])|\right)\right] \\ &\ge Z^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\mathbf{P}^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\left(A_N^c\right)\exp\left(hN^\eta\right).\end{aligned} \tag{3.52}$$

Let us call E_N the event whose probability is estimated from below in (3.42). Then on E_N, whose probability is at least $\zeta > 0$, we have

$$Z^{f,\beta,h_c^a(\beta)+h}_{N,M,\omega} \ge \frac{1}{2}\left(1 - \mathbf{P}^{f,\beta,h_c^a(\beta)}_{N,M,\omega}\left(A_N\right)\right)\exp\left(hN^\eta\right) \ge \frac{1}{4}\exp\left(hN^\eta\right). \tag{3.53}$$

Therefore we obtain

$$\mathbb{P}\left(Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)+h} \geq \frac{1}{4}\exp\left(hN^\eta\right)\right) \geq \mathbb{P}\left(E_N\right) \geq \zeta\,. \tag{3.54}$$

Our aim is to prove that $\mathrm{F}_\gamma(\beta, h + h_c^a(\beta) > 0$ or more precisely give a lower bound for $\mathrm{F}_\gamma(\beta, h + h_c^a(\beta)$. We aim at using (3.24), this is why we have chosen $\gamma \in \mathbb{Q}$ and now we choose also N such that $\gamma N \in \mathbb{N}$, so $N = jq$, $j \in \mathbb{N}$ ($\gamma = p/q$). Since the first part of the proof exploits the free partition function, and not the constrained one for which (3.24) holds, we use Lemma 3.8 that guarantees that

$$\begin{aligned}\log Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)+h} &\leq \log Z_{N,M,\omega}^{c,\beta,h_c^a(\beta)+h} \\ &\quad + c_1\Big(1 + \log(N+M) + \beta|\omega_{N,M}| + \beta \sup_{\substack{1\leq n\leq N\\ 1\leq m\leq M}} \{|\omega_{n,M}|, |\omega_{N,m}|\}\Big)\,.\end{aligned} \tag{3.55}$$

By a direct estimate we see that there exists $c_2 > 1$ such that

$$\mathbb{P}\left(\beta|\omega_{N,M}| + \beta \sup_{\substack{1\leq n\leq N\\ 1\leq m\leq M}} \{|\omega_{n,M}|, |\omega_{N,m}|\} \geqslant c_2\log(NM)\right) \leq \frac{\zeta}{2}\,, \tag{3.56}$$

so that combining with (3.55), and recalling that $M \sim \gamma N$ (so that $\log(NM) = O(\log N)$), we get that there exists $c_3 > 0$ such that

$$\mathbb{P}\Big(\log Z_{N,M,\omega}^{c,\beta,h_c^a(\beta)+h} \leq \log Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)+h} - c_3\log(N)\Big) \leqslant \frac{\zeta}{2}\,.$$

Combining this with (3.54), we get that

$$\mathbb{P}\left(\log Z_{N,M,\omega}^{c,\beta,h_c^a(\beta)+h} \geq \frac{1}{2}hN^\eta - c_3\log N\right) \geq \frac{\zeta}{2}\,. \tag{3.57}$$

Now using the uniform bound $Z_{N,M,\omega}^{c,\beta,h_c^a(\beta)} \geq K(N+M)\exp((\beta\omega_{N,M} - \log Q(\beta))$, we arrive at

$$\begin{aligned}\mathbb{E}\log Z_{N,M,\omega}^{c,\beta,h_c^a(\beta)+h} &\geq \frac{\zeta}{4}hN^\eta - \frac{c_3\zeta}{2}\log N + \log K(N+M) - \beta\mathbb{E}[|\omega_{1,1}|] - \log Q(\beta) \\ &\geq c_4hN^\eta - c_5\log N\,,\end{aligned} \tag{3.58}$$

for suitably chosen $c_4, c_5 > 0$.

At this point the choice $\gamma = p/q$ and $M = \gamma N \in \mathbb{N}$ enters the game. By (3.24) we have

$$\mathrm{F}_\gamma(\beta, h_c^a(\beta) + h) \geqslant \sup_{N=jq:\, j=j_0,j_0+1,\dots} \left\{c_4hN^{\eta-1} - c_5N^{-1}\log N\right\}\,, \tag{3.59}$$

and the fact that j has to be chosen larger than a certain j_0 just reflects the fact that the estimates in this proof have been performed for a N larger than a suitable N_0. We now estimate from below the right-hand side in (3.59) by choosing $N = h^{-\frac{1+\varepsilon}{\eta}}$ (for some $\varepsilon > 0$ fixed): this means that we have chosen $h = (jq)^{-\eta/(1+\varepsilon)}$. With this choice

$$\mathrm{F}_\gamma(\beta, h_c^a(\beta) + h) \geqslant c_4h^{-\varepsilon}N^{-1} - c_5\frac{1+\varepsilon}{\eta}N^{-1}\log\frac{1}{h} \geqslant N^{-1} = h^{\frac{1+\varepsilon}{\eta}}\,, \tag{3.60}$$

and this is the estimate we were after since we can choose η arbitrarily close to α and ε close to 0, but we have established it only for $h = (jq)^{-\eta/(1+\varepsilon)}$, $j = j_0, j_0+1, \dots$. But $h \mapsto \mathrm{F}_\gamma(\beta, h_c^a(\beta)+h)$ is non decreasing so $\mathrm{F}_\gamma(\beta, h_c^a(\beta)+h) \geq h^{\frac{1+\varepsilon}{\eta}}$ for $h = h_j := (jq)^{-\eta/(1+\varepsilon)}$ implies $\mathrm{F}_\gamma(\beta, h_c^a(\beta)+h) \geq h^{\frac{1+\varepsilon}{\eta}}/2$ for every sufficiently small h, which can be verified by checking that $h_j^{\frac{1+\varepsilon}{\eta}}/2$ is smaller than $h_{j+1}^{\frac{1+\varepsilon}{\eta}}$, which is equivalent to $jq+q \leq 2jq$. This completes the proof of Theorem 3.3. □

The technique used to prove Theorem 3.3 could be adapted for $\alpha > 1$ to deduce the upper bound for the difference between quenched and annealed critical points.

Proposition 3.11. *Let $\alpha > 1$. There exists a slowly varying function $\widetilde{L}(\cdot)$ such that*

$$h_c^q(\beta) - h_c^a(\beta) \leq \widetilde{L}(1/\beta)\, \beta^{\frac{2\alpha}{\alpha-1} \vee 4}, \tag{3.61}$$

for $\beta \leq 1$.

Proof. As in the previous proof, it suffices to work with the case $\gamma = p/q \in \mathbb{Q}$ and $M = \lfloor \gamma N \rfloor$. We set

$$N_\beta := \max\left\{ N \in q\mathbb{N} : \mathbb{E}\left[\left(Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)} \right)^2 \right] \leq 2 \right\}. \tag{3.62}$$

Using Paley-Zygmund inequality, we therefore get that $\mathbb{P}(Z_{N,M,\omega}^f > 1/2) \geqslant 1/8$ for any $N \leqslant N_\beta$, and we can then adapt the proof of Proposition 3.9.

Let us take $A_N := \{|\tau \cap ((0,N] \times (0,M])| \leq N/2\mu\}$. Since $\lim_{N\to\infty} \mathbf{P}(A_N) = 0$, and $\mathbb{P}(Z_{N,M,\omega}^f > 1/2) > 1/8$ for $N \leqslant N_\beta$, we find that exactly as in the proof of Proposition 3.9 there exist $N_0 \in \mathbb{N}$ such that for every $N_0 \leq N \leq N_\beta$ we have

$$\mathbb{P}\left(\mathbf{P}_{N,M,\omega}^{f,\beta,h_c^a(\beta)}(A_N) \leq \frac{1}{2} \text{ and } Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)} > \frac{1}{2} \right) \geq \frac{1}{20}. \tag{3.63}$$

Following the proof of Theorem 3.3, provided $N_\beta \geqslant N_0$, and since $N_\beta \in q\mathbb{N}$, we get that

$$\mathrm{F}_\gamma(\beta, h) \geqslant \left\{ c_6(h - h_c^a(\beta)) - c_7 N_\beta^{-1} \log N_\beta \right\}. \tag{3.64}$$

We therefore observe that if $h - h_c^a(\beta) > c_7/c_6 N_\beta^{-1} \log N_\beta$ then $\mathrm{F}_\gamma(\beta, h) > 0$. Hence we get that

$$h_c^q(\beta) - h_c^a(\beta) \leq \frac{c_7}{c_6} . \frac{\log N_\beta}{N_\beta}. \tag{3.65}$$

It therefore boils down to estimating N_β, namely obtaining a lower bound. Recall from (3.48) that

$$\mathbb{E}\left[\left(Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)} \right)^2 \right] = \mathbf{E}^{\otimes 2}\left[\exp\left((\log Q(2\beta) - 2\log Q(\beta)) \mathcal{H}_{N,M}(\nu) \right) \right], \tag{3.66}$$

with $\mathcal{H}_{N,M}(\nu) = \sum_{n=1}^N \sum_{m=1}^M 1_{(n,m)\in\nu}$, and where $\nu = \tau \cap \tau'$ is the intersection renewal. Recall that for $\alpha > 1$ ν is recurrent, see Proposition 3.18.

Note that for $\beta \le 1$, there exists c_9 such that $\log Q(2\beta) - 2\log Q(\beta) \le c_9\,\beta^2$, and that $\mathcal{H}_{N,M} \leqslant \mathcal{H}_{M,M}$. The question is therefore reduced to estimating $\mathbf{E}^{\otimes 2}\left[\exp(t\mathcal{H}_{M,M}(\nu))\right]$, with $t = c_9\beta^2$. Actually we have

$$\begin{aligned}\mathbf{E}^{\otimes 2}\Big[\exp(t\mathcal{H}_{M,M}(\nu))\Big] &= 1 + \sum_{k=1}^{M}\Big(e^{tk} - e^{t(k-1)}\Big)\mathbf{P}^{\otimes 2}\left(\mathcal{H}_{M,M}(\nu) \ge k\right)\\ &\le 1 + (e^t - 1)\sum_{k=1}^{N} e^{tk}\mathbf{P}^{\otimes 2}\left(\mathcal{H}_{N,N}(\nu) \ge k\right).\end{aligned} \tag{3.67}$$

In order to obtain an upper bound, we use the following fact

$$\mathbf{P}^{\otimes 2}\left(\mathcal{H}_{M,M}(\nu) \ge k\right) = \mathbf{P}^{\otimes 2}\left(\nu_k \in (0,M]^2\right) \le \left(\mathbf{P}^{\otimes 2}\left(\nu_1 \in (0,M]^2\right)\right)^k. \tag{3.68}$$

Then we get

$$\mathbf{E}^{\otimes 2}\left[\exp(t\mathcal{H}_{M,M}(\nu))\right] \quad \le \quad 1 \quad + \quad 2\sum_{k=1}^{N}\exp\Big[k\Big(t + \log\mathbf{P}^{\otimes 2}\left(\nu_1 \in (0,M]^2\right)\Big)\Big]. \tag{3.69}$$

Let $\underline{\nu} := \nu^{(1)} + \nu^{(2)}$. An elementary observation is that $\mathbf{P}^{\otimes 2}\left(\nu_1 \notin (0,M]^2\right)$ is of the same order as $\mathbf{P}^{\otimes 2}\left(\underline{\nu}_1 > M\right)$: indeed, for every $M \in \mathbb{N}$ we have

$$\mathbf{P}^{\otimes 2}\left(\underline{\nu}_1 > 2M\right) \le \mathbf{P}^{\otimes 2}\left(\nu_1 \notin (0,M]^2\right) \le \mathbf{P}^{\otimes 2}\left(\underline{\nu}_1 > M\right). \tag{3.70}$$

Therefore, using that $\log x \leqslant x - 1$ for $x > 0$, we get that (see (3.156) for the definition of $U_{N,M}$)

$$\log\mathbf{P}^{\otimes 2}\left(\nu_1 \in (0,M]^2\right) \le -\mathbf{P}^{\otimes 2}\left(\underline{\nu}_1 > 2M\right) \leqslant -c_{11}/U_{M,M} \leqslant -c_{12}/U_{N,N}, \tag{3.71}$$

where we used Lemma 3.20 to estimate $\mathbf{P}^{\otimes 2}\left(\underline{\nu}_1 > 2M\right)$ (provided that M is large enough), and then the fact that $M \leqslant \gamma N$ and $U_{N,N}$ is regularly varying, see Proposition 3.18.

Thus we obtain

$$\mathbf{E}^{\otimes 2}\left[\exp(t\mathcal{H}_{N,N}(\nu))\right] \le 1 + (e^t - 1)\sum_{k=1}^{N}\exp\Big(k(t - c_{12}/U_{N,N})\Big). \tag{3.72}$$

We therefore choose N such that $kc_{12}/U_{N,N} \ge 3tk = 3c_9\beta^2 k$. By Proposition 3.18, for $\alpha > 1$, we can choose

$$N = \widetilde{\psi}(1/\beta)\beta^{-(\frac{2\alpha}{\alpha-1}\vee 4)}, \tag{3.73}$$

for some slowly varying function $\widetilde{\psi}(\cdot)$. For this choice of N, we therefore get that

$$\mathbb{E}\left[\left(Z_{N,M,\omega}^{f,\beta,h_c^a(\beta)}\right)^2\right] \le 1 + (e^{c_9\beta^2} - 1)\sum_{k=1}^{N}\exp(-2c_9\beta^2 k) \le 1 + \frac{e^{c_9\beta^2} - 1}{1 - e^{-2c_9\beta^2}}, \tag{3.74}$$

which is smaller than 2 provided that β is small enough. It therefore implies that there exists some $\beta_1 > 0$ such that

$$N_\beta \ge \widetilde{\psi}(1/\beta)\beta^{-(\frac{2\alpha}{\alpha-1}\vee 4)} \qquad \text{for } \beta \leqslant \beta_1. \tag{3.75}$$

The proof is therefore complete by putting (3.75) in (3.65). □

3.4 Lower bound on the critical point shift

From now on, $L_i(\cdot)$ will denote slowly varying functions and C_i positive constants for $i = 1, 2, \ldots$ Also, we sometimes treat certain large quantities as if they were integers, simply to avoid the integer-part notation; in all cases these can be treated as if the integer-part notation were in use.

Our proof is based on combining the fractional moment method and the change of measure, following the same strategy adopted in [35]. Let

$$z_{n,m} := \exp\left(\beta\omega_{n,m} + h\right) . \tag{3.76}$$

Choose $k \le N$ and M such that $M \sim \gamma N$ and decompose the partition function (3.4) as follows (recall the notation (3.41)), see Figure 3.1:

$$Z_{N,M,\omega} = Z^1_{N,M,\omega} + Z^2_{N,M,\omega} + Z^3_{N,M,\omega} , \tag{3.77}$$

with

$$Z^1_{N,M,\omega} = \sum_{n=k}^{N}\sum_{m=k}^{M} Z_{N-n,M-m,\omega} \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} K(n-i+m-j) z_{N-i,M-j} Z_{(N-i,M-j),(N,M),\omega} , \tag{3.78}$$

$$Z^2_{N,M,\omega} = \sum_{n=1}^{k-1}\sum_{m=k}^{M} Z_{N-n,M-m,\omega} \sum_{i=0}^{n-1}\sum_{j=0}^{k-1} K(n-i+m-j) z_{N-i,M-j} Z_{(N-i,M-j),(N,M),\omega} , \tag{3.79}$$

$$Z^3_{N,M,\omega} = \sum_{n=k}^{N}\sum_{m=1}^{k-1} Z_{N-n,M-m,\omega} \sum_{i=0}^{k-1}\sum_{j=0}^{m-1} K(n-i+m-j) z_{N-i,M-j} Z_{(N-i,M-j),(N,M),\omega} , \tag{3.80}$$

where $Z_{(N-i,M-j),(N,M),\omega}$ has the same law as $Z_{i,j,\omega}$. Moreover $Z_{N-n,M-m,\omega}$, $z_{N-i,M-j}$ and $Z_{(N-i,M-j),(N,M),\omega}$ are independent for $i < n$ and $j < m$.

Let $\delta \in (0,1)$ (that will be chosen close to 1 later in the proof), and define

$$\mathcal{A}_{N,M} := \mathbb{E}\left[(Z_{N,M,\omega})^\delta\right] \qquad \text{for every } N, M \in \mathbb{N}^2, \tag{3.81}$$

with $\mathcal{A}_{0,0} = 1$, and $\mathcal{A}_{i,0} = \mathcal{A}_{0,i} = 0$ for every $i \ge 1$. We apply the inequality $\left(\sum a_i\right)^\delta \le \sum {a_i}^\delta$ (which holds for any finite and countable connection of positive real numbers) to the decomposition (3.77) to get

$$\mathcal{A}_{N,M} \le \mathcal{A}^1_{N,M} + \mathcal{A}^2_{N,M} + \mathcal{A}^3_{N,M} , \tag{3.82}$$

where

$$\mathcal{A}^1_{N,M} \le \mathbb{E}[z^\delta_{1,1}] \sum_{n=k}^{N}\sum_{m=k}^{M} \mathcal{A}_{N-n,M-m} \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} K(n-i+m-j)^\delta \mathcal{A}_{i,j} . \tag{3.83}$$

$$\mathcal{A}^2_{N,M} \le \mathbb{E}[z^\delta_{1,1}] \sum_{n=1}^{k-1}\sum_{m=k}^{M} \mathcal{A}_{N-n,M-m} \sum_{i=0}^{n-1}\sum_{j=0}^{k-1} K(n-i+m-j)^\delta \mathcal{A}_{i,j} . \tag{3.84}$$

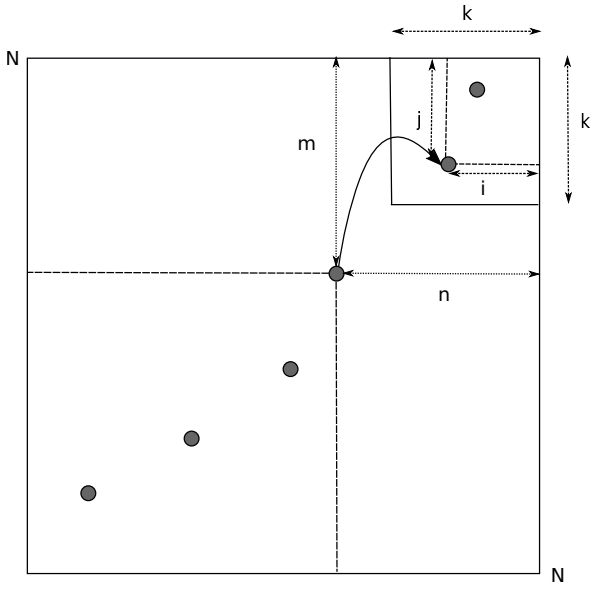

FIGURE 3.1. Fixing a value k, the partition function is decomposed by summing over the values of the last renewal epoch before $(N-k, N-k)$ and the first after $(N-k, N-k)$ to obtain $Z^1_{N,N,\omega}$ and the values of the last renewal epoch $(N-n, N-m)$ in $[N-k, N]\times[1, N-k]$ and the first in $[N-n, N]\times[N-k, N]$ to obtain $Z^2_{N,N,\omega}$ (same idea to obtain $Z^3_{N,N,\omega}$).

$$\mathcal{A}^3_{N,M} \le \mathbb{E}[z^\delta_{1,1}] \sum_{n=k}^{N}\sum_{m=1}^{k-1} \mathcal{A}_{N-n,M-m} \sum_{i=0}^{k-1}\sum_{j=0}^{m-1} K(n-i+m-j)^\delta \mathcal{A}_{i,j}\,. \tag{3.85}$$

The key idea of the proof is to see that

Proposition 3.12. *For fixed β and h, if there exist $k \in \mathbb{N}$ such that $\rho_1+\rho_2+\rho_3 \le 1$ with*

$$\rho_1+\rho_2+\rho_3 := \mathbb{E}[z^\delta_{1,1}]\left(\sum_{n=k}^{\infty}\sum_{m=k}^{\infty}\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} + \sum_{n=1}^{k-1}\sum_{m=k}^{\infty}\sum_{i=0}^{n-1}\sum_{j=0}^{k-1} + \sum_{n=k}^{\infty}\sum_{m=1}^{k-1}\sum_{i=0}^{k-1}\sum_{j=0}^{m-1}\right) K(n-i+m-j)^\delta \mathcal{A}_{i,j}\,, \tag{3.86}$$

then $\mathrm{F}_\gamma(\beta,h) = 0$.

Proof. Define $\overline{A} := \max\{\{\mathcal{A}_{i,j}, \mathcal{A}_{i,s}, \mathcal{A}_{t,j}\},\ 1 \le i,j \le k-1,\ k \le s \le M,\ k \le t \le N\}$. Note that by Jensen's inequality we have $\mathcal{A}_{i,j} \leqslant \mathbb{E}[Z_{i,j}]^\delta \leqslant \exp(\delta h \min\{i,j\})$, since there are at most $\min(i,j)$ renewals in the region $\{1,\dots,i\}\times\{1,\dots,j\}$. Hence we get that $\overline{A} \leqslant e^{hk}$. Then from (3.82) and the fact that $\rho_1+\rho_2+\rho_3 \le 1$, we deduce (by induction) that $\mathcal{A}_{N,M} \le \overline{A}$ for all N, M. Then by Jensen's inequality

$$\mathrm{F}^q_\gamma(\beta,h) = \lim_{\substack{N\to\infty \\ M/N\to\gamma}} \frac{1}{\delta N}\mathbb{E}\log\left(Z_{N,M,\omega}\right)^\delta \le \lim_{\substack{N\to\infty \\ M/N\to\gamma}} \frac{1}{\delta N}\log \mathcal{A}_{N,M} = 0\,. \tag{3.87}$$

□

Our aim is therefore to prove that for $h = h_c^a(\beta) + \Delta_\beta^\varepsilon$ (where Δ_β^ε is defined in Theorem 3.5) we have that $\mathrm{F}_1^q(\beta, h) = 0$ (provided that β is small enough), by showing that ρ_1, ρ_2, ρ_3 are smaller than $1/3$ for such h, for some $k = k_\beta$ wisely chosen. For the choice of k, we pick k proportional to the correlation length of the annealed system, that is $k \propto \mathrm{F}(0, \Delta_\beta^\varepsilon)^{-1}$, and in view of Theorem 3.2 (here $\alpha > 1$ or $\mu < +\infty$), we can take

$$k = k_\beta = \frac{1}{\Delta_\beta^\varepsilon} = \begin{cases} \beta^{-(1+\varepsilon)\frac{2\alpha}{\alpha-1}} & \text{if } \alpha \in (1,2]\,, \\ \beta^{-4}|\log\beta|^6 & \text{if } \alpha > 2. \end{cases} \tag{3.88}$$

Note that, in view of (3.86) and (3.1), we have

$$\rho_1 \le \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \frac{L_1(2k-i-j)}{(2k-i-j)^{(2+\alpha)\delta-2}} \mathcal{A}_{i,j}\,, \tag{3.89}$$

and

$$\rho_2 \le \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \frac{L_2(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \mathcal{A}_{i,j}\,, \tag{3.90}$$

the ρ_3 case being symmetric to ρ_2, we can therefore focus on ρ_1 and ρ_2.

3.4.1 Finite-volume fractional moment estimate

To estimate (3.89) and (3.90), we need a good control over the fractional moment $\mathcal{A}_{i,j}$ for any $i, j \leqslant k$, and we provide estimates in this section.

First of all, using Jensen's inequality, we have that $\mathcal{A}_{i,j} \le (\mathbb{E}Z_{i,j,\omega})^\delta$. Moreover, because $h = h_c^a(\beta) + \Delta_\beta^\varepsilon$, we get that for any $i, j \leqslant k$

$$\mathbb{E}Z_{i,j,\omega} = \mathbf{E}\left[\exp\left(\Delta_\beta^\varepsilon |\tau \cap \{1,...,i\} \times \{1,...,j\}|\right) \mathbf{1}_{(i,j)\in\tau}\right] \le e\,\mathbf{P}((i,j) \in \tau)\,,$$

since $|\tau \cap \{1,...,i\} \times \{1,...,j\}| \leqslant k$ and thanks to of our choice of $k = (\Delta_\beta^\varepsilon)^{-1}$. We therefore get that,

$$\mathcal{A}_{i,j} \leqslant e^\delta \mathbf{P}((i,j) \in \tau)^\delta, \tag{3.91}$$

and $\mathbf{P}((i,j) \in \tau)$ can be estimated thanks to Theorems 3.17.

However, this estimate is rather rough, especially when i, j is close to the diagonal (that is for example $i \leqslant j \leqslant i + a_i$ where $(a_n)_{n\ge 0}$ is defined in Appendix 3.A). We therefore prove the following proposition:

Proposition 3.13. *Let $h = h_c^a(\beta) + \Delta_\beta^\varepsilon$ and $k = (\Delta_\beta^\varepsilon)^{-1}$. Then, define also*

$$\ell_i := \begin{cases} i^{(1+\varepsilon^3)/\alpha} & \text{if } \alpha \in (1,2], \\ C\sqrt{i\log i} & \text{if } \alpha > 2, \end{cases} \tag{3.92}$$

so in any case $\ell_i \gg a_i$. There exists some k_0 such that, provided that $k \geqslant k_0$ then for all $\sqrt{k} \leqslant i \leqslant k$ and $i \leqslant j \leqslant i + \ell_i$ we have that

$$\mathcal{A}_{i,j} \leqslant L_{10}(i)\Big(i^{\delta(1-\frac{1}{\alpha\wedge 2})}(\ell_i)^{-\delta\alpha} + \frac{i^{-\delta(1+\alpha)}\ell_i^\delta}{\beta^{2\delta}} + i^{-\frac{\delta}{\alpha\wedge 2}} e^{-c(\beta^2 i/\ell_i)^{1/2}}\Big). \tag{3.93}$$

This result is the core of the proof, and is based on a change of measure argument. With this result in hand, we are able to show that ρ_1 and ρ_2 are small, in Section 3.4.2 for $\alpha > 2$ and in Section 3.4.3 for $\alpha \in (1,2]$.

In particular, we can apply this proposition in the different cases.

- If $\alpha > 2$, we get that uniformly for $k/2 \leqslant i \leqslant k$ and $i \leqslant j \leqslant i + C'\sqrt{k \log k}$

$$\begin{aligned}\mathcal{A}_{i,j} &\leqslant L_{11}(k) k^{\frac{\delta}{2}(1-\alpha)} + L_{12}(k) \frac{k^{-\delta(1+\alpha)} (k \log k)^{\delta/2}}{\beta^{2\delta}} + L_{13}(k) k^{-\delta/2} e^{-c(\beta^4 k/\log k)^{1/4}} \\ &\leqslant L_{11}(k) \Big(k^{\frac{\delta}{2}(1-\alpha)} + k^{-\delta\alpha} + k^{-\delta/2} e^{-c(\log k)^5/4} \Big) \leqslant L_{11}(k) k^{-\frac{\delta}{2}(\alpha-1)} ,\end{aligned} \tag{3.94}$$

where for the last inequality, we observe that the first term dominates. Here the choice (3.88) of $k = (\Delta_\beta^\varepsilon)^{-1} = \beta^{-4} |\log \beta|^6$ is crucial, to get that $\beta^4 k / \log k \geqslant c(\log k)^5$.

- if $\alpha \in (1,2]$, we use also the choice (3.88) of $k = (\Delta_\beta^\varepsilon)^{-1} = \beta^{-(1+\varepsilon)2\alpha/(\alpha-1)}$ to get that uniformly for $k^{1-\varepsilon^2} \leqslant i \leqslant k$, we have provided that ε is small enough

$$\beta^2 i / \ell_i = k^{-\frac{\alpha-1}{(1+\varepsilon)\alpha}} i^{\frac{\alpha-1-\varepsilon^3}{\alpha}} \geqslant k^{\frac{1}{(1+\varepsilon)\alpha}(\varepsilon(\alpha-1)+O(\varepsilon^2))} \geqslant k^{\varepsilon(\alpha-1)/2\alpha} . \tag{3.95}$$

Therefore, using also that $\beta^{-2} \leqslant k^{3/4} \leqslant i$ (if ε has been fixed small enough) we have that uniformly for $k^{1-\varepsilon^2} \leqslant i \leqslant k$ and $i \leqslant j \leqslant i + \ell_i$,

$$\begin{aligned}\mathcal{A}_{i,j} &\leqslant L_{10}(i) i^{-\delta(\frac{1}{\alpha}+\varepsilon^3)} + L_{10}(i) i^{-\delta\alpha} i^{(1+\varepsilon^3)\delta/\alpha} + L_{10}(i) i^{-\delta/\alpha} e^{-ck^{\varepsilon(\alpha-1)/4\alpha}} \\ &\leqslant L_{11}(k) k^{-\delta(1-\varepsilon^2)(\frac{1}{\alpha}+\varepsilon^3)} \leqslant L_{11}(k) k^{-\frac{\delta}{\alpha}(1+\varepsilon^2/2)}\end{aligned} \tag{3.96}$$

where again, for the second to last inequality, we observe that the first term dominates, and used that ε is small enough for the last inequality.

Proof of Proposition 3.13. The idea is to use a change of measure argument. We define a strip $J_{i,j}$ in which we will tilt the environment by some quantity λ (to be chosen wisely):

$$J_{i,j} := \Big\{ (n,m) \in [\![0,i]\!] \times [\![0,j]\!] \, ; \, |n-m| \leqslant 2\ell_i \Big\} , \tag{3.97}$$

and hence $\# J_{i,j} \leqslant 2i\ell_i$. The width $2\ell_i$ of the strip is chosen because of the scaling of the bivariate renewal: it is very unlikely that the renewal deviates from the diagonal by more than ℓ_i, see Theorem 3.17.

Now, for $\lambda \in \mathbb{R}$ and $i, j \in \mathbb{N}$, we define a new probability measure $\mathbb{P}_{i,j,\lambda}$, under which the $\omega_{n,m}$ are still independent variables, but tilted by λ in the strip $J_{i,j}$:

$$\frac{\mathrm{d}\mathbb{P}_{i,j,\lambda}}{\mathrm{d}\mathbb{P}}(\omega) = \frac{1}{Q(-\lambda)^{\# J_{i,j}}} \exp\left(-\lambda \sum_{(n,m)\in J_{i,j}} \omega_{n,m} \right) , \tag{3.98}$$

where $Q(\cdot)$ is defined in (3.2). Observe now that by Hölder inequality

$$\begin{aligned}\mathcal{A}_{i,j} &= \mathbb{E}_{i,j,\lambda}\left[(Z_{i,j,\omega})^\delta \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}_{i,j,\lambda}}(\omega) \right] \\ &\leqslant \mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega} \big]^\delta \left(\mathbb{E}_{i,j,\lambda}\left[\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}_{i,j,\lambda}}(\omega) \right)^{1/(1-\delta)} \right] \right)^{1-\delta} .\end{aligned} \tag{3.99}$$

The second term in the right-hand side of (3.99) is equal to

$$\left(\mathbb{E}_{i,j,\lambda}\left[\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}_{i,j,\lambda}}(\omega)\right)^{1/(1-\delta)}\right]\right)^{1-\delta} = \left(Q(-\lambda)^{\delta}Q(\lambda\delta/(1-\delta))^{1-\delta}\right)^{\#J_{i,j}}. \tag{3.100}$$

Observe that there exists $c_1 > 0$ such that $0 \le \log Q(x) \le c_1 x^2$ for $|x| \le 1$. Therefore for $|\lambda| \le \min(1, (1-\delta)/\delta)$ and by (3.99) and (3.100), we get

$$\mathcal{A}_{i,j} \le \mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big]^{\delta} \exp\left(c_1\left(\frac{\delta}{1-\delta}\right)\lambda^2 \# J_{i,j}\right), \tag{3.101}$$

and we therefore choose $\lambda := (i\ell_i)^{-1/2}$. Hence $\lambda^2 \# J_{i,j} \leqslant 2$, and

$$\mathcal{A}_{i,j} \leqslant e^{2c_1\delta/(1-\delta)}\mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big]^{\delta}, \tag{3.102}$$

so that we are left with estimating $\mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big]$ for $\lambda := (i\ell_i)^{-1/2}$.

Recall (3.11) and the definition (3.98) of $\mathbb{P}_{i,j,\lambda}$. Using that $\mathbb{E}_{i,j,\lambda}[e^{\beta\omega_{n,m}}]$ equals $Q(\beta) = e^{-h_c^a(\beta)}$ if $(n,m) \notin J_{i,j}$ and $Q(\beta-\lambda)/Q(-\lambda)$ if $(n,m) \in J_{i,j}$, we have that for every β, h, λ and (i,j)

$$\begin{aligned}\mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big] &= \mathbf{E}\left[e^{(h-h_c^a(\beta))|\tau\cap[\![0,i]\!]\times[\![0,j]\!]|}\left(\frac{Q(\beta-\lambda)}{Q(\beta)Q(-\lambda)}\right)^{|\tau\cap J_{i,j}|}\mathbf{1}_{(i,j)\in\tau}\right]\\ &\leqslant e\,\mathbf{E}\left[\left(\frac{Q(\beta-\lambda)}{Q(\beta)Q(-\lambda)}\right)^{|\tau\cap J_{i,j}|}\mathbf{1}_{(i,j)\in\tau}\right],\end{aligned} \tag{3.103}$$

where we used that $|\tau \cap [\![0,i]\!] \times [\![0,j]\!]| \leqslant k$ and $h - h_c^a(\beta) = \Delta_\beta^\varepsilon = k^{-1}$.

Now, observe that $\frac{Q(\beta-\lambda)}{Q(\beta)Q(-\lambda)} = 1 - \lambda\beta + o(\lambda^2+\beta^2)$ as $\lambda, \beta \downarrow 0$. Here, because of our choice (3.88) of $k = (\Delta_\beta^\varepsilon)^{-1}$, we have that $k \geqslant \beta^{-4}$. Since we are considering $i \geqslant \sqrt{k}$, and using that $\ell_i \geqslant \sqrt{i}$, we have that $\lambda := (i\ell_i)^{-1/2} \leqslant i^{-3/4} \leqslant k^{-3/8}$, and hence we have that $\lambda \leqslant \beta$. Therefore, there exists a constant $c_6 > 0$ such that provided that β is small enough (or k is large enough) we have

$$\frac{Q(\beta-\lambda)}{Q(\beta)Q(-\lambda)} \le \exp(-c_6\beta\lambda), \tag{3.104}$$

and we end up with

$$\begin{aligned}e^{-1}\mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big] &\leqslant \mathbf{E}\left[e^{-c_6\beta\lambda|\tau\cap J_{i,j}|}\mathbf{1}_{(i,j)\in\tau}\right]\\ &\leqslant \mathbf{P}\left(\exists s, \tau_s \notin J_{i,j}, (i,j)\in\tau\right) + \mathbf{E}\left[e^{-c_6\lambda\beta|\tau\cap\{1,\dots,i\}\times\{1,\dots,j\}|}\mathbf{1}_{(i,j)\in\tau}\right],\end{aligned} \tag{3.105}$$

where in the last term we dropped the indicator function that all renewals occur in the strip $J_{i,j}$. We now estimate these two terms separately.

Lemma 3.14. *There exists a slowly varying function L_{29} such that, for every $1 \leqslant i \leqslant j \leqslant i + \ell_i$ we have*

$$\mathbf{P}\left(\exists s, \tau_s \notin J_{i,j}, (i,j)\in\tau\right) \leqslant L_{29}(i)\, i^{1-1/\alpha\wedge 2}\,(\ell_i)^{-\alpha}. \tag{3.106}$$

Proof. Let us first observe that

$$\mathbf{P}\left(\exists s, \tau_s \notin J_{i,j}\,,\,(i,j)\in\tau\right) \le \sum_{s_0=1}^{i-1} \mathbf{P}\left(\tau_{s_0}\notin J_{i,j}\,,\,\tau_{s_0-1}\in J_{i,j}\,,\,(i,j)\in\tau\right)$$
$$= \sum_{(a,b)\in J_{i,j}} \sum_{\substack{(k,l);\\ (a+k,b+l)\notin J_{i,j}}} \mathbf{P}((a,b)\in\tau)K(k+l)\mathbf{P}\left((i-a-k,j-b-l)\in\tau\right)\,. \quad (3.107)$$

From Theorem 3.17, we see that the last term in the double sum of (3.107) is bounded above by $\frac{C}{a_i}$ and using some symmetry to consider $\{(a,b) : 1 \le a \le i/2,\, b-a \le \ell_i\}$. We get that (3.107) is bounded above by

$$\frac{C}{a_i}\sum_{a=1}^{i/2}\sum_{r=0}^{\ell_i}\mathbf{P}\left((a,a+r)\in\tau\right)\sum_{k=1}^{i/2-a}\sum_{l\ge\ell_i-a-r}K(k+l) \le \frac{C}{a_i}\mathbf{P}\left(\exists s\le i/2, \tau_s\notin\overline{J}_i\right)\,, \quad (3.108)$$

with

$$\overline{J}_i := \left\{(a,b)\,;\,|a-b|\leqslant\ell_i\right\}. \quad (3.109)$$

Let us now define

$$S_k = \tau_k^{(1)} - \tau_k^{(2)}\,, \quad (3.110)$$

Then we see that

$$\mathbf{P}\left(\exists k\le i/2, \tau_k\notin\overline{J}_i\right) \le \mathbf{P}\left(\max_{k\le i/2}|S_k|\ge\ell_i\right). \quad (3.111)$$

Observe that $\{S_k\}$ is a centred random walk in the domain of attraction of a stable law of index $\alpha>1$. From the Lemma in [81] for the case $\alpha\in(1,2]$ (and infinite variance) and [75, Corollary 1] or equation (12) in [19] for the case $\alpha>2$ we get that

$$\mathbf{P}\left(\max_{k\le i/2}|S_k|\ge\ell_i\right) \le i\,L_{28}(i)\,(\ell_i)^{-\alpha}\,. \quad (3.112)$$

Therefore by (3.108), (3.111) and (3.112), we obtain (3.106). □

Lemma 3.15. *Assume that $i\leqslant j$ and $\mu<+\infty$. There exists a constant $c_{14}>0$ such that, for any $u=u(i)\leqslant 1$ (we may take $u(i)\to 0$ as $i\to+\infty$), we have*

$$Z_{i,j}(-u(i)) := \mathbf{E}\left[e^{-u(i)\,|\tau\cap\{1,\dots,i\}\times\{1,\dots,j\}|}\mathbf{1}_{(i,j)\in\tau}\right] \leqslant c_{14}\frac{K(i+j)}{u(i)^2} + \mathbf{P}((i,j)\in\tau)e^{-c'iu(i)}\,.$$

In particular, we always have

$$Z_{i,j}(-u(i)) \leqslant \frac{L_{27}(i)i^{-(2+\alpha)}}{u(i)^2} + L_{28}(i)i^{-1/\alpha\wedge 2}e^{-ciu(i)}$$

Proof. The last inequality simply comes from the fact that for $i\leqslant j$ we have $K(i+j)\leqslant cL(i)i^{-(2+\alpha)}$ and Theorem 3.17 that gives that $\mathbf{P}((i,j)\in\tau)\leqslant c(a_i)^{-1}$ with $a_i = \psi(i)i^{1/\alpha\wedge 2}$.

We write

$$Z_{i,j}(-u_i) = \sum_{k=1}^{i} e^{-ku(i)}\mathbf{P}(\tau_k = (i,j)) = \left(\sum_{k=1}^{i/2\mu} + \sum_{k=i/2\mu}^{i}\right) e^{-ku(i)}\mathbf{P}(\tau_k = (i,j))\,.$$

For the first sum, we use Theorem 2.8 in [6] to get that for $k \leqslant i/2\mu$ we have $\mathbf{P}(\tau_k = (i,j)) \leqslant ckK(i+j)$, so

$$\sum_{k=1}^{i/2\mu} e^{-ku(i)}\mathbf{P}(\tau_k = (i,j)) \leqslant c\frac{K(i+j)}{u(i)^2}\sum_{k=1}^{+\infty} u(i)ku(i)e^{-ku(i)} \leqslant c\frac{K(i,j)}{u(i)^2}, \tag{3.113}$$

where for the last inequality we bounded the sum by a constant times $\int_{\mathbb{R}_+} xe^{-x}\,\mathrm{d}x$ (thanks to a Riemann-sum approximation along subsequences $u(i) \to 0$).

For the second sum we simply bound k by $i/2\mu$ to get that it is smaller than

$$e^{-iu(i)/2\mu}\sum_{k=i/2\mu}^{i}\mathbf{P}(\tau_k = (i,j)) \leq \mathbf{P}((i,j) \in \tau)e^{-c'iu(i)}\,. \tag{3.114}$$

Combining (3.113) and (3.114), we obtain Lemma 3.15. □

Using Lemma 3.14 and Lemma 3.15 in (3.105), and with $u(i) = c_6\lambda\beta = c_6\beta(i\ell_i)^{-1/2} \leqslant 1$,

$$\mathbb{E}_{i,j,\lambda}\big[Z_{i,j,\omega}\big] \leqslant L_{30}(i)i^{1-1/\alpha\wedge 2}(\ell_i)^{-\alpha} + L_{31}(i)\frac{i^{-(1+\alpha)}\ell_i}{\beta^2} + L_{32}(i)i^{-1/\alpha\wedge 2}e^{-c_7\beta i^{1/2}\ell_i^{-1/2}}\,. \tag{3.115}$$

Finally, one concludes thanks to (3.102), using that $(a+b+c)^\delta \leqslant a^\delta + b^\delta + c^\delta$ for $\delta \in (0,1)$. □

3.4.2 Proof of (3.15) in Theorem 3.5: the case $\alpha > 2$

Let $\delta < 1$ be sufficiently close to 1 to have

$$(2+\alpha)\delta > 4\,, \tag{3.116}$$

which implies that $\delta(\alpha - 1) > 1$.

We start by estimating ρ_1. Let R be a large constant and split the sum in (3.89) as

$$S_1 + S_2 := \left(\sum_{i,j=0}^{k-R-1} + \sum_{i,j=k-R}^{k-1}\right)\frac{L_1(2k-i-j)}{(2k-i-j)^{(2+\alpha)\delta-2}}\mathcal{A}_{i,j}\,, \tag{3.117}$$

and

$$S_3 + S_4 := \left(\sum_{i=0}^{k-R-1}\sum_{j=k-R}^{k-1} + \sum_{i=k-R}^{k-1}\sum_{j=0}^{k-R-1}\right)\frac{L_1(2k-i-j)}{(2k-i-j)^{(2+\alpha)\delta-2}}\mathcal{A}_{i,j}\,. \tag{3.118}$$

Using the fact that $A_{i,j} \leq e^\delta$ from (3.91), we get

$$S_1 \leq \frac{L_2(R)}{R^{(2+\alpha)\delta-4}}\,, \tag{3.119}$$

and the right-hand side of (3.119) can be made small by (3.116) and because R is large.

For S_2, there exists C_2 such that

$$S_2 \le C_2 \max_{k-R\le i,j<k} \mathcal{A}_{i,j}\,, \tag{3.120}$$

and from (3.91), combined with Theorem 3.17, there exists C_3 such that

$$\max_{k-R\le i,j<k} \mathcal{A}_{i,j} \le e^{\delta} \max_{k-R\le i,j<k} \mathbf{P}((i,j)\in\tau)^{\delta} \le \frac{C_3}{k^{\delta/2}}\,, \tag{3.121}$$

then S_2 is arbitrarily small for k large.

Since S_3 and S_4 are the same quantity, we just focus on S_3. Since $A_{i,j} \le e^{\delta}$ from (3.91), we obtain

$$S_3 \le \frac{L_3(R)}{R^{(2+\alpha)\delta-4}}\,, \tag{3.122}$$

which again can be made small in view of the condition (3.116) and because R is large. Hence ρ_1 can be made arbitrarily small by choosing R large and k large (*i.e.* β small).

Let us now look at ρ_2 in (3.90) (the argument for ρ_3 is identical). We split the sum to:

$$S_5 + S_6 = \left(\sum_{i=0}^{k-1}\sum_{j=0}^{i} + \sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\right)\mathcal{A}_{i,j}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}}\,. \tag{3.123}$$

Let us first study S_5:

$$\begin{aligned} S_5 &= \sum_{j=0}^{k-1}\sum_{i=j}^{k-1}\mathcal{A}_{i,j}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \\ &= \left(\sum_{j=0}^{k/2}\sum_{i=j}^{k-1} + \sum_{j=k/2+1}^{k-1}\sum_{i=j}^{k-1}\right)\mathcal{A}_{i,j}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} := S_{5a} + S_{5b}\,. \end{aligned} \tag{3.124}$$

Using (3.91) (in particular the fact $\mathcal{A}_{i,j} \le e^{\delta}$), we get

$$S_{5a} \le \frac{L_5(k)}{k^{(2+\alpha)\delta-4}}\,. \tag{3.125}$$

For S_{5b}, we use (3.91) and Theorem 3.17 to get

$$S_{5b} \le \sum_{j=k/2+1}^{k-1}\sum_{i=j}^{k-1}\frac{C}{j^{\delta/2}}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \le \frac{C_4}{k^{\delta/2}}\,. \tag{3.126}$$

Then S_5 can be made small for k large and from condition (3.116).

Now we split S_6 as

$$S_{6a} + S_{6b} = \left(\sum_{i=0}^{k/2} + \sum_{i=k/2+1}^{k-1}\right)\sum_{j=(i+\ell_i+1)\wedge(k-1)}^{k-1}\mathcal{A}_{i,j}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}}\,. \tag{3.127}$$

and

$$S_{6c} + S_{6d} = \left(\sum_{i=0}^{k/2} + \sum_{i=k/2+1}^{k-1}\right)\sum_{j=i+1}^{(i+\ell_i)\wedge(k-1)}\mathcal{A}_{i,j}\frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}}\,. \tag{3.128}$$

Using (3.91) and Theorem 3.17, we see that

$$\begin{aligned} S_{6a} &\leq \sum_{i=0}^{k/2} \sum_{j=1}^{3k/4} \frac{L_6(k-j)}{(k-j)^{(2+\alpha)\delta-2}} + \sum_{i=0}^{k/2} \sum_{j=3k/4}^{k-1} \frac{Ci^\delta L(j-i)^\delta}{(j-i)^{(1+\alpha)\delta}} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \\ &\leq \frac{L_7(k)}{k^{(2+\alpha)\delta-4}} + \frac{L_8(k)}{k^{\alpha\delta-1}}, \end{aligned} \tag{3.129}$$

and

$$\begin{aligned} S_{6b} &\leq \sum_{i=k/2+1}^{k-1} \sum_{j=(i+c\sqrt{k\log k})\wedge(k-1)}^{k-1} \frac{Ci^\delta L(j-i)^\delta}{(j-i)^{(1+\alpha)\delta}} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \\ &\leq Ck^{\alpha\delta/2} \sum_{j=k/2}^{k-1} \sum_{x \geqslant c\sqrt{k\log k}} \frac{L(x)^\delta}{x^{(1+\alpha)\delta}} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \leqslant L_9(k) \frac{k^\delta}{\sqrt{k}^{(1+\alpha)\delta-1}} = \frac{L_9(k)}{k^{\frac{1}{2}((\alpha-1)\delta-1)}}. \end{aligned}$$

By (3.91) (bounding the probability by 1), and since provided that k is large enough we have $i + \sqrt{i\log i} \leqslant 3k/4$ for $i \leqslant k/2$, we obtain

$$S_{6c} \leq C \sum_{i=0}^{k/2} c\sqrt{i\log i} \frac{L_4(k)}{k^{(2+\alpha)\delta-2}} \leq \frac{L_{10}(k)}{k^{(2+\alpha)\delta-7/2}}. \tag{3.130}$$

Therefore again by the condition (3.116), $S_{6a} + S_{6b} + S_{6c}$ is arbitrarily small for k large.

For the term S_{6d}, since there are at most $C_5\sqrt{k\log k}$ terms in the sum over j, we have

$$S_{6d} \leq C_5\sqrt{k\log k} \max_{\substack{k/2\leq i\leq k \\ i\leq j\leq i+\sqrt{i\log i}}} \mathcal{A}_{i,j}. \tag{3.131}$$

Then we use Proposition 3.13, and more precisely (3.94), to get that

$$S_{6d} \leq L_{13}(k) k^{-\frac{1}{2}(\delta(\alpha-1)-1)} \tag{3.132}$$

In view of the condition (3.116), S_{6d} can be made arbitrarily small for k large. This completes the proof of (3.15) in the case $\alpha > 2$.

3.4.3 Proof of (3.15) in Theorem 3.5: the case $\alpha \in (1, 2]$

Fix $\varepsilon > 0$ small and let $0 < \delta < 1$ such that

$$\delta\left[(2+\alpha) + (1-\varepsilon^2)/\alpha\right] > 4 - \varepsilon^2, \tag{3.133}$$

and

$$\delta(1-\varepsilon^2)\left[\alpha + \frac{\varepsilon}{2}\left(\frac{\alpha-1}{\alpha}\right)\right] > \frac{1}{\alpha}. \tag{3.134}$$

In particular, the condition (3.133) implies that

$$\delta(2+\alpha) > 3. \tag{3.135}$$

Let us start with showing that ρ_1 is small: we split the sum in (3.89) to

$$T_1 + T_2 := \left(\sum_{i,j=0}^{k^{1-\varepsilon^2}} + \sum_{i,j=k^{1-\varepsilon^2}+1}^{k-1} \right) \frac{L_1(2k-i-j)}{(2k-i-j)^{(2+\alpha)\delta-2}} A_{i,j} \,, \tag{3.136}$$

and

$$T_3 + T_4 := \left(\sum_{i=0}^{k^{1-\varepsilon^2}} \sum_{j=k^{1-\varepsilon^2}+1}^{k-1} + \sum_{i=k^{1-\varepsilon^2}+1}^{k-1} \sum_{j=0}^{k^{1-\varepsilon^2}} \right) \frac{L_1(2k-i-j)}{(2k-i-j)^{(2+\alpha)\delta-2}} A_{i,j} \,. \tag{3.137}$$

For $\alpha \leqslant 2$, we know that there exists a slowly varying function $\psi(\cdot)$ such that $a_i = \psi(i) i^{1/\alpha}$. For T_1, using (3.91) and Theorem 3.17, we get

$$T_1 \leq \frac{L_{15}(k)}{k^{(2+\alpha)\delta-2}} \sum_{i,j=0}^{k^{1-\varepsilon^2}} \frac{1}{(a_{\min(i,j)})^\delta} \leqslant \frac{L_{16}(k)}{k^{(1-\varepsilon^2)(\delta/\alpha-2)+(2+\alpha)\delta-2}} \,, \tag{3.138}$$

and from the condition (3.133), T_1 can be made small for k large.

For T_2, using (3.91) and Theorem 3.17, we have (since $(2+\alpha)\delta - 2 \in (1,2)$)

$$T_2 \leq L_8(k) k^{-(2+\alpha)\delta+4} \max_{k^{1-\varepsilon^2} \leq i,j \leq k} \mathcal{A}_{i,j} < \frac{L_9(k)}{k^{(1-\varepsilon^2)\delta/\alpha+(2+\alpha)\delta-4}} \,, \tag{3.139}$$

then T_2 is small for k large thanks to (3.133).

For T_3 (the argument for T_4 is identical), we use again (3.91) and Theorem 3.17 to get that

$$T_3 \leq \sum_{i=1}^{k^{1-\varepsilon^2}} \frac{1}{(a_i)^\delta} \frac{L_{16}(k-i)}{(k-i)^{(2+\alpha)\delta-3}} \leqslant \frac{L_{17}(k)}{k^{(2+\alpha)\delta-3+(1-\varepsilon^2)(\delta/\alpha-1)}} \,, \tag{3.140}$$

which can be made small by taking k large, thanks to (3.133). In the end, we get that ρ_1 is bounded above by a small constant for k large.

As far as ρ_2 is concerned, we split the right-hand side of (3.90) to

$$T_5 + T_6 = \sum_{i=0}^{k-1} \sum_{j=i+1}^{k-1} \mathcal{A}_{i,j} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} + \sum_{i=0}^{k-1} \sum_{j=0}^{i} \mathcal{A}_{i,j} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \,. \tag{3.141}$$

Recall the definition of ℓ_i in Proposition 3.13, and define $\bar{\ell}_i = i^{(1+\varepsilon^4)/\alpha} \ll \ell_i$. We split T_5 as

$$T_{5a} + T_{5b} = \left(\sum_{i=0}^{k/2} + \sum_{i=k/2+1}^{k-1} \right) \sum_{j=(i+\bar{\ell}_i)\wedge(k-1)}^{k-1} \mathcal{A}_{i,j} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \,. \tag{3.142}$$

and

$$T_{5c} + T_{5d} = \left(\sum_{i=0}^{k^{1-\varepsilon^2}} + \sum_{i=k^{1-\varepsilon^2}+1}^{k-1} \right) \sum_{j=i+1}^{(i+\bar{\ell}_i)\wedge(k-1)} \mathcal{A}_{i,j} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \,. \tag{3.143}$$

From (3.91) and Theorem 3.17, we get that

$$\begin{aligned} T_{5a} &\le \sum_{i=0}^{k/2} \left(\sum_{j=i+\bar\ell_i}^{3k/4} + \sum_{j=3k/4+1}^{k-1} \right) \frac{C i^\delta L(j-i)^\delta}{(j-i)^{(1+\alpha)\delta}} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \\ &\le \frac{L_{18}(k)}{k^{(2+\alpha)\delta-2}} \sum_{i=1}^{k/2} i^\delta \frac{L_{19}(i)}{i^{\frac{1+\varepsilon^4}{\alpha}((1+\alpha)\delta-1)}} + \frac{L_{20}(k)}{k^{\delta\alpha-1}} \\ &\leqslant \frac{L_{21}(k)}{k^{(1+\alpha)\delta-3+(1+\varepsilon^4)((1+\alpha)\delta-1)/\alpha}} + \frac{L_{20}(k)}{k^{\delta\alpha-1}}, \end{aligned} \tag{3.144}$$

and also

$$\begin{aligned} T_{5b} &\le \sum_{i=k/2+1}^{k-1} \sum_{j=(i+c\bar\ell_k)\wedge(k-1)}^{k-1} \frac{C i^\delta L(j-i)^\delta}{(j-i)^{(1+\alpha)\delta}} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \\ &\le C k^\delta \sum_{j=k/2}^{k-1} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \sum_{x \geqslant c\bar\ell_k} \frac{L_{22}(x)}{x^{(1+\alpha)\delta}} \\ &\le L_{23}(k) k^\delta (\bar\ell_k)^{1-(1+\alpha)\delta} = \frac{L_{23}(k)}{k^{(1+\varepsilon^4)((1+\alpha)\delta-1)/\alpha-\delta}}. \end{aligned} \tag{3.145}$$

Using (3.91) and Theorem 3.17, we get

$$\begin{aligned} T_{5c} &\le \sum_{i=0}^{k^{1-\varepsilon^2}} \sum_{j=i+1}^{(i+\bar\ell_i)\wedge(k-1)} \frac{C}{a_i^\delta} \frac{L_4(k-j)}{(k-j)^{(2+\alpha)\delta-2}} \le \frac{L_{23}(k)}{k^{(2+\alpha)\delta-2}} \sum_{i=0}^{k^{1-\varepsilon^2}} L(i) i^{(1+\varepsilon^4)/\alpha-\delta/\alpha} \\ &\leqslant \frac{L_{24}(k)}{k^{(2+\alpha)\delta-2}} k^{(1-\varepsilon^2)((1+\varepsilon^4-\delta)/\alpha+1)}. \end{aligned} \tag{3.146}$$

Then the condition (3.133) imposed on δ guarantees that T_{5a}, T_{5b} and T_{5c} tend to zero when k tends to infinity.

Finally, it remains to bound T_{5d}. As for (3.131), there are at most $\bar\ell_k$ terms in the sum over i, so that

$$T_{5d} \le C_{11}\, \bar\ell_k \max_{\substack{k^{1-\varepsilon^2} \le i < k-1 \\ i \le j \le i+\ell_i}} \mathcal{A}_{i,j}. \tag{3.147}$$

Then we use Proposition 3.13, and more precisely (3.96), to get that

$$T_{5d} \le L_{12}(k) k^{(1+\varepsilon^4)/\alpha} k^{-\delta(1+\varepsilon^2/2)/\alpha}, \tag{3.148}$$

we have that this term goes to 0 as $k \to \infty$ if $\delta > (1+\epsilon^4)/(1+\epsilon^2/2)$ which can be achieved by taking δ sufficiently close to 1.

For T_6, we have

$$T_6 \le \sum_{j=0}^{k-1} \sum_{i=j}^{k-1} \mathcal{A}_{i,j} \frac{L_4(k-i)}{(k-i)^{(2+\alpha)\delta-2}}. \tag{3.149}$$

By following the same procedure adopted for T_5, T_6 is bounded above by a small term when k is large. The proof of (3.15) in the case $\alpha \in (1,2]$ is therefore complete and, with it, also the proof of the lower bound part of Theorem 3.5.

Appendix

3.A Bivariate renewal theory, important estimates

We now present some results on the bivariate renewal process τ defined in Section 3.1.1, and in particular Proposition 3.18 which gives some conditions on the transience/recurrence of the intersection renewal $\nu = \tau \cap \tau'$.

First, let us stress that τ is in the domain of attraction of a $\min(\alpha, 2)$-stable distribution: we denote $(b_n)_{n \geqslant 1}$ be the recentering sequence and $(a_n)_{n \geq 0}$ the renormalizing sequence for τ_n, that is such that $\frac{1}{a_n}(\tau_n - (b_n, b_n))$ converges to a $\min(\alpha, 2)$ stable distribution, whose density is denoted $g_\alpha(\cdot, \cdot)$.

For b_n, we have $b_n = \mu n$ if $\mu < +\infty$, $b_n = n\mathbf{E}[X_1 \wedge n]$ if $\alpha = 1, \mu = +\infty$, and $b_n = 0$ if $\alpha \in (0, 1)$. For a_n, we have

- if $\alpha \in (0, 2]$ and $\mathbf{E}[X_1^2] = +\infty$, the asymptotic behavior of a_n as $n \to \infty$ is characterized by

$$\begin{aligned} L(a_n)(a_n)^{-\alpha} &\sim \frac{1}{n} \qquad \text{if } \alpha < 2 \\ \sigma(a_n)(a_n)^{-2} &\sim \frac{1}{n} \qquad \text{if } \alpha = 2 \end{aligned} \tag{3.150}$$

  where $\sigma(n) := \mathbf{E}[(X_1 \wedge n)^2] \asymp \sum_{k=1}^n L(k)k^{-1}$ verifies $\sigma(n)/L(n) \to +\infty$.

- if $\mathbf{E}[X_1^2] < +\infty$, then $a_n = \sqrt{n}$ (note that, up to a constant, this fits the definition (3.150)).

In any case, there exists some slowly varying function $\psi(\cdot)$ such that

$$a_n \overset{n\to\infty}{\sim} \psi(n) n^{\frac{1}{\min(\alpha,2)}}. \tag{3.151}$$

First of all, before further study, we need sharp renewal theorems, that can be found in [6]: it provides sharp asymptotics along the favorite scaling, and general upper bounds away from it. We collect the results giving estimates of the quantity $\mathbf{P}((n, n+k) \in \tau)$ for any $k \geqslant 0$ (this is enough by symmetry) in the two following theorems, dealing with the cases $\alpha \in (0, 1)$ and $\alpha \geqslant 1, \mu < +\infty$. The case $\alpha = 1, \mu = +\infty$ is much more technical and subtle, and it is not treated in [6].

Theorem 3.16. *If $\alpha \in (0, 1)$, and for $k/n \to t \in \mathbb{R}_+$, we have*

$$\mathbf{P}\left((n, n+k) \in \tau\right) \overset{n\to\infty}{\sim} C_\alpha(t) L(n)^{-1} n^{-(2-\alpha)}, \tag{3.152}$$

with $C_\alpha(t) := \alpha \int_0^{+\infty} x^{1-\alpha} g_\alpha(x, (1+t)x)\, \mathrm{d}x$. Moreover, there exits a constant C such that for any $k \geqslant n$

$$\mathbf{P}\left((n, n+k) \in \tau\right) \leq C L(n)^{-1} n^{-(2-\alpha)} \frac{L(k)}{L(n)} \left(\frac{k}{n}\right)^{-(1+\alpha)}. \tag{3.153}$$

Theorem 3.17. *If $\alpha \geqslant 1, \mu < +\infty$, for any $n \to \infty$ and k with $k/a_n \to t \in \mathbb{R}_+$, we have that*

$$\mathbf{P}\left((n, n+k) \in \tau\right) \overset{n\to\infty}{\sim} c_\alpha(t)\,\frac{1}{a_n}, \tag{3.154}$$

where $c_\alpha(t) = \mu_\alpha \int_{-\infty}^{+\infty} g_\alpha(x, x+\mu_\alpha t)$ with $\mu_\alpha := \mu^{1/\min(\alpha,2)}$. Moreover, there exists a constant $C > 0$ such that for any $k \geqslant a_n$,

$$\mathbf{P}\big((n, n+k) \in \tau\big) \leqslant CnL(k)k^{-(1+\alpha)} + \frac{C}{a_n} e^{-ck^2/n\sigma(k)} \mathbf{1}_{\{\alpha \geqslant 2\}},$$

and

$$\mathbf{P}\big((n, n+k) \in \tau\big) \leqslant \frac{C}{a_n}\left(\frac{L(k)}{L(a_n)}\Big(\frac{k}{a_n}\Big)^{-(1+\alpha)} + Ce^{-ck^2/n\sigma(k)} \mathbf{1}_{\{\alpha \geqslant 2\}}\right).$$

When $\alpha \geqslant 2$, then the second term is negligible as soon as $k \geqslant C\sqrt{n \log n}$. Note that we have some constant C' such that (provided $\alpha \geqslant 1$, $\mu < +\infty$) for any $k \geqslant a_n$

$$\mathbf{P}\big((n, n+k) \in \tau\big) \leqslant \frac{C'}{a_n} \times \Big(\frac{k}{a_n}\Big)^{-3/2}. \tag{3.155}$$

The two Theorems 3.16-3.17 enables us to estimate the mean overlap of two copies τ and τ' in the region $(0, N] \times (0, M]$. We define

$$U_{N,M} := \mathbf{E}\left[|\nu \cap ([0, N] \times [0, M])|\right] = \sum_{n=0}^{N} \sum_{m=0}^{M} \mathbf{P}((n, m) \in \tau)^2, \tag{3.156}$$

and for any $\lambda > 0$

$$\hat{U}(\lambda) := \sum_{n,m=0}^{+\infty} e^{-\lambda(n+m)} \mathbf{P}((n, m) \in \tau)^2. \tag{3.157}$$

Recall that $a_n = \psi(n) n^{1/(\alpha \wedge 2)}$ is the normalizing sequence for τ_n (cf. (3.150)-(3.151)).

Proposition 3.18. *If $\alpha < 1$, then $\sup_{N,M \in \mathbb{N}} U_{N,M} < +\infty$.*

If $\alpha \geqslant 1$, $\mu < +\infty$, then set $\rho := 1 - \min(\alpha, 2)^{-1} \in [0, 1/2]$. We have

$$\begin{aligned} U_{N,N} &\sim 2c_\alpha \sum_{n=1}^{N} \frac{1}{a_n} \to +\infty && \text{as } N \to \infty, \\ \hat{U}(\lambda) &\sim \frac{2^{1-\rho} c_\alpha}{\Gamma(1+\rho)} \sum_{n=1}^{1/\lambda} \frac{1}{a_n} && \text{as } \lambda \downarrow 0,\ (1/\lambda \in \mathbb{N}), \end{aligned} \tag{3.158}$$

with $c_\alpha = \int_0^\infty c_\alpha(t)^2 \mathrm{d}t$, $c_\alpha(t)$ being the constant appearing in Theorem 3.17.

Moreover, if $\alpha \geqslant 1, \mu < +\infty$

$$\sum_{n=1}^{N} \frac{1}{a_n} \sim \widetilde{L}(N) N^\rho \qquad \text{as } N \to \infty, \tag{3.159}$$

for some slowly varying function $\widetilde{L}(\cdot)$, with $\widetilde{L}(N)/\log N \to +\infty$ if $\alpha = 1, \mu < +\infty$.

As a consequence, we have that $\nu = \tau \cap \tau'$ is terminating if $\alpha < 1$, and persistent if $\alpha \geqslant 1, \mu < +\infty$.

Proof of Proposition 3.18. Let us start with the case $\alpha < 1$. Notice that by symmetry, for $M \geqslant N$,

$$U_{N,M} \leqslant U_{M,M} \leqslant 2 \sum_{n=1}^{M} \sum_{k=0}^{M-n} \mathbf{P}((n, n+k) \in \tau)^2 .$$

We therefore need to control the last sum. Let us denote

$$W_n := \sum_{k \geqslant 0} \mathbf{P}((n, n+k) \in \tau)^2 .$$

Using Theorem 3.16 (and properties of slowly varying functions), we get that there is a constant c such that for all $n \geqslant 1$

$$W_n \leqslant c \sum_{k=1}^{n} L(n)^{-2} n^{-2(2-\alpha)} + c \sum_{k \geqslant n} L(n)^{-2} n^{-2(2-\alpha)} (k/n)^{-2} \leqslant C' L(n)^{-2} n^{2\alpha-3},$$

where we used that $\sum_{k \geqslant n} (k/n)^{-2} \sim n \int_1^\infty x^{-2}$ as $n \to \infty$. Therefore, since $\alpha < 1$, we get that

$$\sup_{N,M} U_{N,M} \leqslant \sum_{n=1}^{+\infty} W_n < +\infty .$$

We now turn to the case $\alpha \geqslant 1$, $\mu < +\infty$. We start by showing that $\sum_{n=1}^{N} \frac{1}{a_n}$ diverges.

If $\alpha > 1$, this is immediate since $a_n \sim \psi(n) n^{\frac{1}{\min(\alpha,2)}}$, see (3.151): it directly gives (3.159).

If $\alpha = 1, \mu < +\infty$, we show that $\psi(n) = a_n/n \to 0$. Therefore, $D_N := \sum_{n=1}^{N} (a_n)^{-1} = \sum_{n=1}^{n} \psi(n)^{-1} n^{-1}$, diverges to $+\infty$ as a slowly varying function, and $D_N / \log N \to +\infty$.

Since $L(n)n^{-1}$ is asymptotically decreasing, there exists a slowly varying function $\mathbf{L}$ and a constant c such that $\mathbf{L}(n)n^{-1} \leqslant L(n)n^{-1} \leqslant c\mathbf{L}(n)n^{-1}$, and with $\mathbf{L}(n)n^{-1}$ non-increasing in n. Hence, because $\sum_{n \geqslant 1} L(n)n^{-1} < +\infty$ (since $\mu < +\infty$), we have that $\sum_{n \geqslant 1} \mathbf{L}(n)n^{-1} < +\infty$, and because $\mathbf{L}(n)n^{-1}$ is non-increasing we get that

$$k \times \frac{\mathbf{L}(2k)}{2k} \leqslant \sum_{n=k}^{2k} \mathbf{L}(n)n^{-1} \to 0, \quad \text{as } k \to \infty.$$

We therefore get that $\mathbf{L}(n) \to 0$ as $n \to \infty$, which yields that $L(n) \to 0$, and $\psi(n) \to 0$ because of (3.150).

We now prove (3.158). We fix $\varepsilon > 0$, and denote, in complement to the definition of W_n above

$$W_n^{(\varepsilon)} := \sum_{k=0}^{\lfloor \frac{1}{\varepsilon} a_n \rfloor} \mathbf{P}((n, n+k) \in \tau)^2 .$$

As a preliminary, we show that there exists some n_ε such that, provided that $n \geqslant n_\varepsilon$

$$(1-\varepsilon)\frac{c_\alpha}{a_n} \leqslant W_n^{(\varepsilon)} \leqslant W_n \leqslant (1+\varepsilon)\frac{c_\alpha}{a_n}. \tag{3.160}$$

Note that we also have that $W_n \leqslant \sum_{k=0}^{+\infty} \mathbf{P}((n, n+k) \in \tau) \leqslant 1$ for any n.

To prove (3.160), we use Theorem 3.17 to get that uniformly for $0 \leqslant k \leqslant \frac{1}{\varepsilon}a_n$, $\mathbf{P}((n,n+k) \in \tau)^2 \sim (a_n)^{-2}c_\alpha(k/a_n)^2$. Hence, provided that n is large enough, we get that

$$a_n W_n^{(\varepsilon)} \geqslant (1-\varepsilon^2)\frac{1}{a_n}\sum_{k=0}^{\lfloor \frac{1}{\varepsilon}a_n \rfloor} c_\alpha(k/a_n)^2 \geqslant (1-2\varepsilon^2)\int_0^{1/\varepsilon} c_\alpha(t)^2 \mathrm{d}t\,,$$

the last inequality holding by Riemann-sum approximation. Note that a similar upper bound, with $1-2\varepsilon^2$ replaced with$1+2\varepsilon^2$ holds. Now, thanks to (3.155), there exists a constant $c>0$ such that

$$a_n(W_n - W_n^{(\varepsilon)}) = a_n \sum_{k>\frac{1}{\varepsilon}a_n}^{+\infty} \mathbf{P}((n,n+k) \in \tau)^2 \leqslant c\frac{1}{a_n}\sum_{k>\frac{1}{\varepsilon}a_n}\left(\frac{k}{a_n}\right)^{-3} \leqslant c'\varepsilon^2,$$

where the last inequality also comes from a Riemann-sum approximation. Finally, note that $c_\alpha - \int_0^{1/\varepsilon} c_\alpha(t)^2\mathrm{d}t$ is positive, and thanks to (3.155) smaller than $\int_{1/\varepsilon} ct^{-3}\mathrm{d}t \leqslant c''\varepsilon^2$. In the end, we get that, provided that n is large enough,

$$(1-2\varepsilon^2)(c_\alpha - c''\varepsilon^2) \leqslant a_n W_n^{(\varepsilon)} \leqslant a_n W_n \leqslant (1+2\varepsilon^2)c_\alpha + c'\varepsilon^2, \tag{3.161}$$

which gives (3.160) provided that ε has been fixed small enough.

We are now ready to estimate $U_{N,N}$. We write

$$U_{N,N} = 2\sum_{n=0}^{N}\sum_{k=0}^{N-n}\mathbf{P}((n,n+k)\in\tau)^2 - \sum_{n=0}^{N}\mathbf{P}((n,n)\in\tau)^2.$$

The second sum is negligible compared to $\sum_{n=1}^{N}\frac{1}{a_n}$, since $\mathbf{P}((n,n)\in\tau)^2 \sim c(a_n)^{-2}$, with $a_n \to +\infty$. We therefore focus on the first sum.

An upper bound is simply

$$\sum_{n=0}^{N}\sum_{k=0}^{N-n}\mathbf{P}((n,n+k)\in\tau)^2 \leqslant \sum_{n=0}^{N} W_n,$$

and since we have that $W_n \sim c_\alpha/a_n$ and $\sum\frac{1}{a_n}$ is diverging, we get that for n large enough

$$\sum_{n=0}^{N}\sum_{k=0}^{N-n}\mathbf{P}((n,n+k)\in\tau)^2 \leqslant (1+2\varepsilon)c_\alpha\sum_{n=1}^{N}\frac{1}{a_n}$$

A lower bound is

$$\sum_{n=0}^{N}\sum_{k=0}^{N-n}\mathbf{P}((n,n+k)\in\tau)^2 \geqslant \sum_{n=n_\varepsilon}^{(1-\varepsilon)N} W_n^{(\varepsilon)} \geqslant (1-2\varepsilon)c_\alpha\sum_{n=1}^{(1-\varepsilon)N}\frac{1}{a_n} \geqslant (1-c\varepsilon)c_\alpha\sum_{n=1}^{N}\frac{1}{a_n},$$

where we used the lower bound (3.160) valid for n large enough, together with the fact that $\sum\frac{1}{a_n}$ is diverging, as a regularly varying function for the last inequality.

We now turn to estimating $\hat{U}(\lambda)$ as $\lambda \to 0$. By symmetry, we can write that

$$\hat{U}(\lambda) = 2\sum_{n=0}^{+\infty}e^{-2\lambda n}\sum_{k=0}^{+\infty}e^{-\lambda k}\mathbf{P}((n,n+k)\in\tau)^2 - \sum_{n=0}^{+\infty}e^{-2\lambda n}\mathbf{P}((n,n)\in\tau)^2.$$

Again, the second term is negligible compared to $\sum_{n=1}^{1/\lambda} \frac{1}{a_n}$ as $\lambda \to 0$, since $\sum_{n=0}^{N} \mathbf{P}((n,n) \in \tau)^2$ is negligible compared to $\sum_{n=1}^{N} \frac{1}{a_n}$, by standard properties of Laplace transforms. We therefore focus on the first term.

First of all, an upper bound is

$$\sum_{n=0}^{+\infty} e^{-2\lambda n} \sum_{k=0}^{+\infty} e^{-\lambda k} \mathbf{P}((n,n+k) \in \tau)^2 \leqslant \sum_{n=0}^{+\infty} e^{-2\lambda n} W_n \,.$$

Since $\sum_{n=0}^{N} W_n \sim c_\alpha \sum_{n=1}^{N} \frac{1}{a_n}$ and diverges as a regularly varying function with exponent $\rho = 1 - 1/\min(\alpha, 2)$, we get by standard properties of Laplace transforms (see Corollary 1.7.3 in [14]) that

$$\sum_{n=0}^{+\infty} e^{-2\lambda n} W_n \sim \frac{c_\alpha}{\Gamma(1+\rho)} \sum_{n=1}^{1/(2\lambda)} \frac{1}{a_n} \sim \frac{c_\alpha}{2^\rho \Gamma(1+\rho)} \sum_{n=1}^{1/\lambda} \frac{1}{a_n} \qquad \text{as } \lambda \downarrow 0 \,.$$

For a lower bound, we get that

$$\sum_{n=0}^{+\infty} e^{-2\lambda n} \sum_{k=0}^{+\infty} e^{-\lambda k} \mathbf{P}((n,n+k) \in \tau)^2 \geqslant \sum_{n=0}^{+\infty} e^{-2\lambda n} e^{-\lambda a_n/\varepsilon} W_n^{(\varepsilon)} \,.$$

Now, we use that there is some n_ε such that for $n \geqslant n_\varepsilon$ we have that $W_n^{(\varepsilon)} \geqslant (1-\varepsilon) c_\alpha / a_n$ (see (3.160)), and that $a_n/\varepsilon \leqslant \varepsilon n$, since $a_n/n \to 0$ (it is obvious if $\alpha > 1$, and has been shown above if $\alpha = 1$ and $\mu < +\infty$). We therefore get that

$$\begin{aligned} &\sum_{n=0}^{+\infty} e^{-2\lambda n} \sum_{k=0}^{+\infty} e^{-\lambda k} \mathbf{P}((n,n+k) \in \tau)^2 \\ &\qquad \geqslant (1-\varepsilon) c_\alpha \sum_{n=n_\varepsilon}^{+\infty} e^{-2(1+\varepsilon)\lambda n} \frac{1}{a_n} \overset{\lambda \to 0}{\sim} \frac{(1-\varepsilon) c_\alpha}{(2(1+\varepsilon))^\rho \Gamma(1+\rho)} \sum_{n=1}^{1/\lambda} \frac{1}{a_n} \,, \end{aligned}$$

where we used again Corollary 1.7.3 in [14] for the last asymptotics.

By letting ε go to 0, we obtain matching upper and lower bound, so that (3.158) is proved.

□

Remark 3.19. The case $\alpha = 1$, $\mu = +\infty$ has been left aside, because of the lack of results on the renewal mass function $\mathbf{P}((n,m) \in \tau)$. We believe that, close to the diagonal, we should have

$$\mathbf{P}((n,n+m) \in \tau) \asymp \frac{\min(k_n, a_{k_n})}{(a_{k_n})^2}, \qquad k_n = n\mu^\star(n), \tag{3.162}$$

for any $|m| \leqslant a_{k_n}$. Here, $\mu^\star(n)$ denotes the conjugate slowly varying function of $\mu(n) := \mathbf{E}[X_1 \wedge n]$, that is such that $\mu(n\mu^\star(n)) \sim \mu^\star(n)^{-1}$ (In particular, $\mu(n) \sim \mu(n)^{-1}$ for most reasonable slowly varying functions, such as $(\log n)^\gamma$ or $e^{(\log n)^\kappa}$). With the definition $k_n = n\mu^\star(n)$, k_n is the typical number of renewals before length n, $\tau_{k_n} \approx n$.

If $L(\cdot)$ is bounded, then $a_k \leqslant ck$ because of (3.150) and one obtains $1/a_{k_n}$ in (3.162). If $L(n) \to +\infty$, one obtains $k_n/(a_{k_n})^2$ in (3.162).

We note that $a_{k_n} \leqslant cn$ for all n (and we believe that $a_{k_n} = o(n)$, as it is the case in all the reasonable examples I can come up with): indeed, k_n is the typical number of jumps needed to reach n ($\tau_{k_n} \asymp n$), and a_{k_n} has the order of $\max_{k \leqslant k_n}(\tau_k - \tau_{k-1})$, so a_{k_n} it is necessarily smaller than $\tau_{k_n} \approx n$.

Also in the $\alpha = 1$ case, the main contribution to $U_{N,N}$ should come from the terms close to the diagonal, that is

$$U_{N,N} \asymp 2 \sum_{n=1}^{N} \sum_{k=0}^{a_{k_n}} \mathbf{P}((n, +k) \in \tau)^2 \asymp \sum_{n=1}^{N} \frac{\min(k_n, a_{k_n})^2}{(a_{k_n})^3}.$$

As seen above, if $L(n)$ is bounded, then we obtain

$$U_{N,N} \asymp \sum_{n=1}^{N} \frac{\max(k_n, a_{k_n})^2}{(a_{k_n})^3} \asymp \sum_{n=1}^{N} \frac{1}{a_{k_n}} \geqslant c \log N,$$

where we used that $a_{k_n} \leqslant cn$ as remarked above. One should therefore get that ν is persistent if $L(\cdot)$ is bounded. On the other hand, when $L(n) \to +\infty$, we should have that $U_{N,N} \asymp \sum_{n=1}^{N} (k_n)^2 / a_{k_n}$, and there is no obvious simplification of this expression. In the end, we expect that we have the following criterion:

$$\nu = \tau \cap \tau' \text{ is persistent} \quad \Leftrightarrow \quad \sum_{n \geqslant 1} \frac{(k_n)^2}{(a_{k_n})^3} = +\infty.$$

(The sum is indeed infinite when $L(n)$ is bounded.)

For example, if $L(n) = (\log n)^\gamma$ with $\gamma \geqslant 0$, then $\mu(n) \sim (\log n)^{1+\gamma}$ and $k_n \sim n(\log n)^{-(1+\gamma)}$. Moreover, $a_n \sim n(\log n)^\gamma$, and hence $a_{k_n} \sim n(\log n)^{-1}$. We get that $(k_n)^2/(a_{k_n})^3 \sim (\log n)^{1-2\gamma} n^{-1}$, and we should have that $\nu = \tau \cap \tau'$ is persistent if and only if $\gamma \leqslant 1$. From this example, one may wonder if the necessary and sufficient condition for persistence could be $\sum_{n=1}^{+\infty} \frac{1}{a_n} = +\infty$, which is equivalent to the recurrence of the random walk $\tau_k - \mu(k, k)$.

We now use Proposition 3.18, and in particular the estimate of the Laplace transform $\hat{U}(\lambda)$, to obtain estimates on the tail probability of the intersection renewal $\nu = \tau \cap \tau'$. More precisely, we define $\underline{\nu} := \nu^{(1)} + \nu^{(2)}$ and estimate $\mathbf{P}^{\otimes 2}(\underline{\nu}_1 > N)$.

Lemma 3.20. *Assume that $\alpha \geqslant 1$ and $\nu < +\infty$. Then recalling that $\rho = 1 - \min(\alpha, 2)^{-1} \in [0, 1/2]$, we get that*

$$\mathbf{P}^{\otimes 2}(\underline{\nu}_1 > N) \overset{N \to \infty}{\sim} \frac{2^\rho \sin(\pi\rho)}{\pi\rho} (U_{N,N})^{-1}. \tag{3.163}$$

Proof. Recall the definition of $\hat{U}(\lambda) = \sum_{n,m \geq 0} e^{-\lambda(n+m)} \mathbf{P}^{\otimes 2}((n, m) \in \nu)$. We also set, for any $\lambda > 0$,

$$\hat{K}(\lambda) := \sum_{n,m \geq 1} e^{-\lambda(n+m)} \mathbf{P}^{\otimes 2}(\nu_1 = (n, m)) = \sum_{k \geq 2} e^{-\lambda k} \mathbf{P}^{\otimes 2}(\underline{\nu}_1 = k). \tag{3.164}$$

The key idea of this proof is the following identity

$$\hat{U}(\lambda) = 1 + \hat{K}(\lambda)\hat{U}(\lambda) \quad \Leftrightarrow \quad 1 - \hat{K}(\lambda) = \frac{1}{\hat{U}(\lambda)}, \tag{3.165}$$

which can be obtained from the identity

$$\mathbf{P}^{\otimes 2}\left((n,m) \in \nu\right) = \mathbf{1}_{n=m=0} + \sum_{i=1}^{n}\sum_{j=1}^{m} \mathbf{P}^{\otimes 2}(\nu_1 = (i,j))\mathbf{P}^{\otimes 2}((n-i,n-j) \in \nu). \tag{3.166}$$

Now, since we know the behavior of $\hat{U}(\lambda)$ as $\lambda \downarrow 0$, we get the behavior of $\hat{K}(\lambda)$, from which we should be able to infer that of $\mathbf{P}^{\otimes 2}(\underline{\nu}_1 > N)$. Let us develop here how we proceed: we use Corollary 1.7.3 and Theorem 8.7.3 in [14]. Viewing $\underline{\nu}$ as a renewal process with inter-arrival $\mathbf{P}^{\otimes 2}(\underline{\nu}_1 = k) = \mathbf{P}^{\otimes 2}(\nu_1^{(1)} + \nu_1^{(2)} = k)$ which is indeed a probability distribution, we set $u_n := \mathbf{P}^{\otimes 2}(n \in \underline{\nu})$ and interpret $\hat{U}(\lambda) = \sum_{n=0}^{\infty} e^{-\lambda n} u_n$, in view of (3.165). Now, [14, Corollary 1.7.3] tells that since $\hat{U}(\lambda)$ is regularly varying with exponent $-\rho$ (recall $\rho = 1 - \min(\alpha, 2)^{-1}$), we have that $\sum_{n=0}^{N} u_n \sim \Gamma(1+\rho)\hat{U}(1/N) \sim 2^{-\rho} U_{N,N}$ (where we used (3.158)). In turn [14, Theorem 8.7.3] gives that

$$\mathbf{P}^{\otimes 2}(\underline{\nu}_1 > N) \overset{N\to\infty}{\sim} \frac{(2^{-\rho}U_{N,N})^{-1}}{\Gamma(1+\rho)\Gamma(1-\rho)},$$

and we are done. □

We conclude this Appendix with a useful Lemma.

Lemma 3.21. *Given $\delta > 0$ there exists $\varepsilon > 0$ such that for N sufficiently large and $M \sim \gamma N$ we have*

$$\mathbf{P}\Big(|\tau \cap (0,N] \times (0,M]| \geq \varepsilon N^{\alpha}/L(N)\Big) \geq 1 - \delta\,. \tag{3.167}$$

Proof. Set $n = n(\varepsilon, N) = \varepsilon N^{\alpha}/L(N)$ and $B_{N,M} := (0,N] \times (0,M]$. We want to prove

$$\mathbf{P}\left(\tau_n \notin B_{N,M}\right) \leq \delta\,. \tag{3.168}$$

Let us set

$$\widetilde{\tau}_n^{(1)} := \sum_{i=1}^{n} (\tau_i^{(1)} - \tau_{i-1}^{(1)}) \mathbf{1}_{\{\tau_i^{(1)} - \tau_{i-1}^{(1)} \leq N\}}\,. \tag{3.169}$$

$$\widetilde{\tau}_n^{(2)} := \sum_{i=1}^{n} (\tau_i^{(2)} - \tau_{i-1}^{(2)}) \mathbf{1}_{\{\tau_i^{(2)} - \tau_{i-1}^{(2)} \leq M\}}\,. \tag{3.170}$$

Then we have

$$\mathbf{P}\left(\tau_n \notin B_{N,M}\right) \leq \mathbf{P}\left(\widetilde{\tau}_n \notin B_{N,M}\right) + \mathbf{P}\left(\exists i \leq n; (\tau_i - \tau_{i-1}) \notin B_{N,M}\right)\,. \tag{3.171}$$

The second term is bounded above by

$$\begin{aligned} & n\Big(\mathbf{P}(\tau_1^{(1)} > N, \tau_1^{(2)} \leq M) + \mathbf{P}(\tau_1^{(1)} \leq N, \tau_1^{(2)} > M) + \mathbf{P}(\tau_1^{(1)} > N, \tau_1^{(2)} > M)\Big) \\ & \leq n\left(\sum_{s>N}\sum_{t=1}^{M} K(s+t) + \sum_{s=1}^{N}\sum_{t>M} K(s+t) + \sum_{s>N}\sum_{t>M} K(s+t)\right) \\ & \leq \varepsilon C_{\alpha,\gamma} \leq \delta/2\,, \end{aligned} \tag{3.172}$$

for $\varepsilon = \delta/(2C_{\alpha,\gamma})$. The first term in (3.171) is bounded by $\mathbf{P}(\widetilde{\tau}_n^{(1)} > N) + \mathbf{P}(\widetilde{\tau}_n^{(2)} > M)$. Observe that for every choice of $\lambda_1 > 0$ we have

$$\mathbf{P}(\widetilde{\tau}_n^{(1)} > N) \le e^{-\lambda_1 N}\mathbf{E}[e^{\lambda_1 \widetilde{\tau}_n^{(1)}}] \le e^{n \log \mathbf{E}[e^{\lambda_1 \widetilde{\tau}_1^{(1)}}] - \lambda_1 N}. \tag{3.173}$$

The marginals $\tau_n^{(1)}$ and $\tau_n^{(2)}$ have the same distribution: as $N \to \infty$ we have

$$\mathbf{P}(\tau_1^{(1)} = N) = \mathbf{P}(\tau_1^{(2)} = N) \sim \frac{1}{(1+\alpha)} L(N) N^{-(1+\alpha)}. \tag{3.174}$$

Using the fact that for any $s \ge 1$

$$\mathbf{E}[(\widetilde{\tau}_1^{(1)})^s \mathbf{1}_{\widetilde{\tau}_1^{(1)} < N}] \le N^{s-1}\mathbf{E}[(\widetilde{\tau}_1^{(1)})\mathbf{1}_{\widetilde{\tau}_1^{(1)} < N}] \le c_\alpha L(N) N^{-\alpha} N^s, \tag{3.175}$$

so that

$$\log \mathbf{E}[e^{\lambda_1 \widetilde{\tau}_1^{(1)}} \mathbf{1}_{\widetilde{\tau}_1^{(1)} < N}] \le \log[1 + c_\alpha L(N) N^{-\alpha}(e^{\lambda_1 N} - 1)] \le c_\alpha L(N) N^{-\alpha} e^{\lambda_1 N}. \tag{3.176}$$

We choose $\lambda_1 = 2N^{-1} \log \delta^{-1}$, then we get

$$n \log \mathbf{E}[e^{\lambda_1 \widetilde{\tau}_1^{(1)}}] \le c_\alpha \varepsilon \delta^{-2}. \tag{3.177}$$

Therefore by choosing $\varepsilon = \delta^2$, we obtain (with different c_α)

$$\mathbf{P}(\widetilde{\tau}_n^{(1)} > N) \le c_\alpha \delta^2. \tag{3.178}$$

Using the same reasoning and choosing $\lambda_2 = 2M^{-1} \log \delta^{-1}$, we have that there exists $c_{\alpha,\gamma}$ such that

$$\mathbf{P}(\widetilde{\tau}_n^{(2)} > M) \le c_{\alpha,\gamma} \delta^2. \tag{3.179}$$

The proof is therefore complete by taking $\varepsilon = \min\{\delta/(2C_{\alpha,\gamma}), \delta^2\}$. □

3.B Irrelevance via a replica-coupling trick

We prove first the irrelevance of disorder for $\alpha \in (0,1)$ based on the method of quadratic replica-coupling method that is used in [88] for the univariate disordered pinning model to derive a lower bound on the free energy and prove disorder irrelevance. This method was generalized in [91] to the case of infinitely divisible disorder.

Proof. The upper bound follows directly from (3.12), it suffices then to prove the lower bound. Let $\Delta > 0$ and observe that

$$\mathrm{F}_\gamma(\beta, -\beta^2/2 + \Delta) = \mathrm{F}_\gamma(0, \Delta) + \lim_{\substack{N,M\to\infty \\ M/N\to\gamma}} \Phi_{N,M,\Delta}(1,\beta), \tag{3.180}$$

where

$$\Phi_{N,M,\Delta}(t,\beta) := \frac{1}{N}\mathbb{E} \log \mathbf{E}_{N,M,\Delta}\left[e^{\sum_{n=1}^N \sum_{m=1}^M (\beta\sqrt{t}\omega_{n,m} - t\beta^2/2)\delta_{n,m}}\right], \tag{3.181}$$

and for a $\mathbf{P}$-measurable function $f(\tau)$, we have

$$\mathbf{E}_{N,M,\Delta}\left[f(\tau)\right] = \frac{\mathbf{E}\left[e^{\Delta\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}} f\delta_{N,M}\right]}{\mathbf{E}\left[e^{\Delta\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}}\delta_{N,M}\right]}. \tag{3.182}$$

Using the gaussian integration by parts formula (if ω is a centered standard Gaussian variable)

$$\mathbb{E}\left[\omega f(\omega)\right] = \mathbb{E}\left[f'(\omega)\right], \tag{3.183}$$

for every differentiable function such that $\lim_{|x|\to\infty} e^{-\frac{x^2}{2}} f(x) = 0$, then for $0 < t < 1$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_{N,M,\Delta}(t,\beta) = \frac{-\beta^2}{2N}\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbb{E}\left(\frac{\mathbf{E}_{N,M,\Delta}\left[\delta_{n,m}e^{\sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\sqrt{t}\omega_{n,m}-t\beta^2/2)\delta_{n,m}}\right]}{\mathbf{E}_{N,M,\Delta}\left[e^{\sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\sqrt{t}\omega_{n,m}-t\beta^2/2)\delta_{n,m}}\right]}\right)^2. \tag{3.184}$$

For notational convenience, we define for $\lambda \geq 0$

$$\Psi_{N,M,\Delta}(t,\lambda,\beta) := \frac{1}{2N}\mathbb{E}\log\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right], \tag{3.185}$$

with

$$H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau') = \sum_{n=1}^{N}\sum_{m=1}^{M}(\beta\sqrt{t}\omega_{n,m} - t\beta^2/2)(\delta_{n,m}+\delta'_{n,m}) + \lambda\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}, \tag{3.186}$$

and τ, τ' are two independent bivariate renewal process under the law $\mathbf{P}^{\otimes 2}$.

Using again the Gaussian integration by parts formula for $\Psi_{N,M,\Delta}$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Psi_{N,M,\Delta}(t,\lambda,\beta) &= \frac{\beta^2}{2N}\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbb{E}\left(\frac{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[\delta_{n,m}\delta'_{n,m}e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}\right) \\
&\quad \frac{-\beta^2}{4N}\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbb{E}\left(\frac{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[(\delta_{n,m}+\delta'_{n,m})e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}\right)^2 \\
&\leq \frac{\beta^2}{2N}\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbb{E}\left(\frac{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[\delta_{n,m}\delta'_{n,m}e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}{\mathbf{E}^{\otimes 2}_{N,M,\Delta}\left[e^{H_{N,M,\Delta}(t,\lambda,\beta,\tau,\tau')}\right]}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_{N,M,\Delta}(t,\lambda,\beta).
\end{aligned} \tag{3.187}$$

The above implies that for every $t \in [0,1]$ and $\lambda \geq 0$

$$\Psi_{N,M,\Delta}(t,\lambda,\beta) \leq \Psi_{N,M,\Delta}(0,\lambda+t,\beta). \tag{3.188}$$

Comparing (3.184) and (3.187), and using convexity and monotonicity of $\Psi_{N,M,\Delta}(t,\lambda,\beta)$ with respect to λ and the fact that $\Psi_{N,M,\Delta}(t,0,\beta) = \Phi_{N,M,\Delta}(t,\beta)$, one gets

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\Phi_{N,M,\Delta}(t,\beta) &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_{N,M,\Delta}(t,\lambda,\beta)\Big|_{\lambda=0} \\
&\leq \frac{\Psi_{N,M,\Delta}(t,2-t,\beta) - \Phi_{N,M,\Delta}(t,\beta)}{2-t} \leq \Psi_{N,M,\Delta}(0,2,\beta) - \Phi_{N,M,\Delta}(t,\beta),
\end{aligned} \tag{3.189}$$

where in the last inequality we used $(2-t) \geq 1$ and (3.188). Integrating this inequality between 0 and 1 and observing that $\Phi_{N,M,\Delta}(0) = 0$, we get

$$\Phi_{N,M,\Delta}(t,\beta) \geq (1-e)\Psi_{N,M,\Delta}(0,2,\beta)\,. \tag{3.190}$$

Using Holder's inequality and p, q such that $\frac{1}{p} + \frac{1}{q} = 1$, we obtain

$$\begin{aligned}\Psi_{N,M,\Delta}(0,2,\beta) &= -\mathrm{F}_{\gamma,N}(0,\Delta) + \\ &\frac{1}{2N}\log \mathbf{E}^{\otimes 2}\left[e^{2\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}+\Delta\sum_{n=1}^{N}\sum_{m=1}^{M}(\delta_{n,m}+\delta'_{n,m})}\delta_{(N,M)}\delta'_{(N,M)}\right] \\ &\leq -\mathrm{F}_{\gamma,N}(0,\Delta) + \frac{\mathrm{F}_{\gamma,N}(0,q\Delta)}{q} + \frac{1}{2Np}\log \mathbf{E}^{\otimes 2}\left[e^{2p\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}}\right],\end{aligned} \tag{3.191}$$

where $\mathrm{F}_{\gamma,N}(0,\Delta) := \frac{1}{N}\log \mathbf{E}\left[e^{\Delta\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}}\delta_{N,M}\right]$. One finds then

$$\begin{aligned}\limsup_{\substack{N,M\to\infty\\ M/N\to\gamma}} \Psi_{N,M,\Delta}(0,2,\beta) &\leq \limsup_{\substack{N,M\to\infty\\ M/N\to\gamma}} \frac{1}{2Np}\log \mathbf{E}^{\otimes 2}\left[e^{2p\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}}\right] \\ &\quad + \mathrm{F}_{\gamma}(0,\Delta)\left(\frac{1}{q}\frac{\mathrm{F}_{\gamma}(0,q\Delta)}{\mathrm{F}_{\gamma}(0,\Delta)} - 1\right).\end{aligned} \tag{3.192}$$

By the property of slow variation and from the critical behavior of the free energy in the homogeneous case (see Theorem 1.2 in [51]), for every $q > 0$ we have

$$\lim_{\Delta\searrow 0}\frac{\mathrm{F}_{\gamma}(0,q\Delta)}{\mathrm{F}_{\gamma}(0,\Delta)} = q^{1/\alpha}\,. \tag{3.193}$$

Choosing $q = q(\varepsilon)$ sufficiently close to 1 (then $p(\varepsilon) = \frac{q(\varepsilon)}{q(\varepsilon)-1} < \infty$) and $\Delta_0(\varepsilon) > 0$ sufficiently small one has, uniformly on $\beta \geq 0$ and on $0 < \Delta \leq \Delta_0(\varepsilon)$

$$\begin{aligned}\limsup_{\substack{N,M\to\infty\\ M/N\to\gamma}} \Psi_{N,M,\Delta}(0,2,\beta) &\leq \frac{\varepsilon}{e-1}\mathrm{F}_{\gamma}(0,\Delta) \\ &\quad + \limsup_{\substack{N,M\to\infty\\ M/N\to\gamma}} \frac{1}{2Np(\varepsilon)}\log \mathbf{E}^{\otimes 2}\left[e^{2p(\varepsilon)\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}}\right].\end{aligned} \tag{3.194}$$

Now observe that from Proposition 3.18, the renewal process $\nu := \tau \cap \tau'$ is terminating under the law $\mathbf{P}^{\otimes 2}$ if $\alpha \in (0,1)$. Therefore, there exists $\beta_1 > 0$ such that

$$\sup_{N}\mathbf{E}^{\otimes 2}\left[e^{2p\beta^2\sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{n,m}\delta'_{n,m}}\right] < \infty\,, \tag{3.195}$$

for every $\beta^2 p(\varepsilon) \leq \beta_1^2$. Then by (3.194) and (3.180), we get

$$\mathrm{F}_{\gamma}(\beta, -\beta^2/2 + \Delta) \geq (1-\varepsilon)\mathrm{F}_{\gamma}(0,\Delta)\,, \tag{3.196}$$

for $\beta^2 \leq \beta_0{}^2(\varepsilon) := \beta_1{}^2/p(\varepsilon)$.

□

Bibliography

[1] K. S. Alexander, *The effect of disorder on polymer depinning transitions*, Commun. Math. Phys. **279** (2008), 117-146.

[2] K. S. Alexander and V. Sidoravicius, *Pinning of polymers and interfaces by random potentials*, Ann. Appl. Probab., **16** (2006), 636-669.

[3] K. S. Alexander and N. Zygouras, *Quenched and annealed critical points in polymer pinning models*, Comm. Math. Phys. **291** (2009), 659-689.

[4] A. Bar, A. Kabakçıŏglu, D. Mukamel, *Denaturation of circular DNA: supercoil mechanism*, Phys. Rev. E **84** (2011), 041935.

[5] A. Bar, A. Kabakçıŏglu and D. Mukamel, *Denaturation of circular DNA: supercoils and overtwist*, Phys. Rev. E **86** (2012), 061904.

[6] Q. Berger, *Bivariate renewal theory: renewal theorems and local large deviation estimates*, work in progress.

[7] Q. Berger, F. Caravenna, J. Poisat, R. Sun and N. Zygouras, *The critical curve of the random pinning and copolymer models at weak coupling*, Commun. Math. Phys. **326** (2014), 507-530.

[8] Q. Berger and H. Lacoin, *Sharp critical behavior for pinning models in a random correlated environment*, Stochastic Process. Appl. **122** (2012), 1397-1436.

[9] Q. Berger and H. Lacoin, *The effect of disorder on the free-energy for the Random Walk Pinning Model: smoothing of the phase transition and low temperature asymptotics*, J. Stat. Phys. **42** (2011), 322-341.

[10] Q. Berger and H. Lacoin, *Pinning on a defect line: characterization of marginal disorder relevance and sharp asymptotics for the critical point shift*, J. Inst. Math. Jussieu, Firstview 1-42 (2016).

[11] Q. Berger and H. Lacoin, *The high-temperature behavior of the directed polymer in dimension* $1+2$, Ann. Inst. Henri Poincaré Probab. Stat., to appear.

[12] Q. Berger and J. Poisat, *On the critical curve of the pinning and copolymer models in correlated Gaussian environment*, Elect. J. Probab. **20** (2015), 35 pages.

[13] Q. Berger and F. Toninelli, *On the critical point of the Random Walk Pinning Model in dimension d = 3*, Electron. J. Probab. **15** (2010), 654-683

[14] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation*, Cambridge University Press, Cambridge (1987).

[15] M. Birkner and R. Sun, *Annealed vs quenched critical points for a random walk pinning model*, Ann. Inst. H. Poincaré **46** (2010), 414-441.

[16] M. Birkner and R. Sun, *Disorder relevance for the random walk pinning model in dimension 3*, Ann. Inst. H. Poincaré **47** (2011), 259-293.

[17] R. D. Blake and S. G. Delcourt, *Thermal stability of DNA*, Nucleic Acids Research **26** (1998), 3323-3332.

[18] R. D. Blake, J. W. Bizzaro, J. D. Blake, G. R. Day, S. G Delcourt, J. Knowles, K. A. Marx, K.A. and J. Jr SantaLucia, *Statistical Mechanical Simulation of Polymeric DNA Melting with MELTSIM*, Bioinformatics **15** (1999), 370-375.

[19] A.A. Borovkov, K.A. Borovkov, *On probabilities of large deviations for random walks. I. Regularly varying distribution tails*, Theory Probab. Appl. **46** (2000), 193-213.

[20] A. A. Borovkov and A. A. Mogul'skiĭ, *The second rate function and the asymptotic problems of renewal and hitting the boundary for multidimensional random walks*, Sibirsk. Mat. Zh. **37** (1996) No. 4, 745-782.

[21] A. A. Borovkov and A. A. Mogul'skiĭ, *Integro-local limit theorems including large deviations for sums of random vectors II*, Theory Probab.Appl. **45** (2001), 3-22.

[22] A. A. Borovkov and A. A. Mogul'skiĭ, *On large deviations of sums of independent random vectors on the boundary and outside of the cramér zone I*, Theory Probab.Appl. **53** (2009), No. 2, 301-311.

[23] R. Brak, A. L. Owczarek, T. Prellberg and A. J Guttmann, *Finite-length scaling of collapsing directed walks*, Phys. Rev. E **48** (1993), 2386-2396.

[24] R. Bundschuh and T. Hwa, *Statistical mechanics of secondary structures formed by random RNA sequences*, Phys. Rev. E **65** (2002), 031903 (22 pages).

[25] F. Caravenna, F. den Hollander, and N. Pétrélis, *Lectures on random polymers* **15**, Clay Mathematics Proceedings, 2012, Proceedings of the Clay Mathematics Institute

Summer School and XIV Brazilian School of Probability (Buzios, Brazil). Edited by D. Ellwood and C. Newman and V. Sidoravicius and W. Werner.

[26] F. Caravenna and F. den Hollander, *A general smoothing inequality for disordered polymers*, Electron. Commun. Probab. **18** (2013),1-15.

[27] F. Caravenna, G. Giacomin and L. Zambotti, *Sharp asymptotic behavior for wetting models in (1+1)-dimension*, Elect. J. Probab. **11** (2006), 345-362.

[28] F. Caravenna, F. L. Toninelli and N. Torri , *Universality for the pinning model in the weak coupling regime*, arXiv:1505.04927

[29] P. Carmona, N. Pétrélis and G. B. Nguyen, *Interacting partially self-avoiding walk. From phase transition to the geometry of the collapsed phase*, Ann. Probab., to appear, http://arxiv.org/abs/1306.4887

[30] E. Carlon, E. Orlandini et A.L. Stella, *Roles of Stiffness and Excluded Volume in DNA Denaturation*, Phys. Rev. Letters **88** (2002), 198101.

[31] D. Cheliotis and F. den Hollander, *Variational characterization of the critical curve for pinning of random polymers*, Ann. Probab. **41** (2013), 1767-1805.

[32] F. Comets, *Weak disorder for low dimensional polymers: the model of stable laws. Markov Process*, Markov Process. Related Fields **13** (2007), 681-696 .

[33] D. Cule and T. Hwa, *Denaturation of Heterogeneous DNA*, Phys. Rev. Lett. **79** (1997), 2375 .

[34] P. G. De Gennes, *Scaling concepts in polymer physics*, Cornell University Press, Ithaca, NY (1979).

[35] B. Derrida, G. Giacomin, H. Lacoin and F. L. Toninelli, *Fractional moment bounds and disorder relevance for pinning models*, Commun. Math. Phys. **287** (2009), 867-887 .

[36] B. Derrida, V. Hakim, and J. Vannimenus, *Effect of disorder on two- dimensional wetting*, J. Stat. Phys. **66** (1992),1189-1213.

[37] B. Derrida and M. Retaux, *The depinning transition in presence of disorder: a toy model*, J. Stat. Phys. **156** (2014), 268-290.

[38] M. Doi and S. F. Edwards, *The theory of Polymer dynamics*, Oxford University Press (1986).

[39] R. A. Doney, *An analogue of the renewal theorem in higher dimensions*, Proc. London Math. Soc. (3) **16** (1966), 669-684.

[40] R. A. Doney, *One-sided local large deviation and renewal theorems in the case of infinite mean*, Probab. Theory Relat. Fields, **107** (1997), 451-465.

[41] T. R. Einert, H. Orland and R. R. Netz, *Secondary structure formation of homopolymeric single-stranded nucleic acids including force and loop entropy: implications for DNA hybridization*, Eur. Phys. J. E **34** (2011), 55 (15 pages).

[42] M. E. Fisher, *Walks, walls, wetting, and melting*, J. Statist. Phys. **34** (1984), 667-729.

[43] P. Flory, *Principles of Polymer Chemistry*, Cornell University Press (1953).

[44] G. Forgacs, J. M. Luck, T. M. Nieuwenhuizen, and H. Orland, *Wetting of a disordered substrate: exact critical behavior in two dimensions*, Phys. Rev. Lett.**57** (1986), 2184-2187.

[45] T. Garel and C. Monthus, *Numerical study of the disordered Poland-Scheraga model of DNA denaturation*, Journal of Statistical Mechanics : Theory and Experiment (2005), P06004.

[46] T. Garel and H. Orland, *On the role of mismatches in DNA denaturation*, arXiv:cond-mat/0304080

[47] T. Garel and H. Orland, *Generalized Poland-Scheraga model for DNA hybridization*, Biopolymers **75** (2004), 453-467.

[48] G. Giacomin, *Random polymer models*, Imperial College Press, World Scientific, 2007.

[49] G. Giacomin, *Renewal convergence rates and correlation decay for homogeneous pinning models*, Elect. J. Probab. **13** (2008), 513-529.

[50] G. Giacomin, *Disorder and critical phenomena through basic probability models*,Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010. Lecture Notes in Mathematics **2025**, Springer, 2011.

[51] G. Giacomin and M. Khatib, *Generalized Poland Sheraga denaturation model and two dimensional renewal processes*, arXiv:1510.07996

[52] G. Giacomin, H. Lacoin and F. L. Toninelli, *Hierarchical pinning models, quadratic maps and quenched disorder*, Probab. Theor. Rel. Fields **147** (2010), 185-216.

[53] G. Giacomin, H. Lacoin and F. L. Toninelli, *Marginal relevance of disorder for pinning models*, Comm. Pure Appl. Math. **63** (2010), 233-265.

[54] G. Giacomin, H. Lacoin and F. L. Toninelli, *Disorder relevance at marginality and critical point shift*, Ann. Inst. H. Poincaré **47** (2011), 148-175.

[55] G. Giacomin and F. L. Toninelli, *The localized phase of disordered copolymers with adsorption*, ALEA-Latin American J. Probab. Math. Stat. **1** (2006), 149-180.

[56] G. Giacomin and F. L. Toninelli, *Smoothing effect of quenched disorder on polymer depinning transitions*, Commun. Math. Phys. **266** (2006), 1-16.

[57] G. Giacomin and F. L. Toninelli, *The localized phase of disordered copolymers with adsorption*, ALEA-Latin American Journal of Probability and Mathematical Statistics **1** (2006), 149-180.

[58] A. B. Harris, *Effect of random defects on the critical behaviour of ising models*, J. Phys. C **7** (1974), 1671-1692.

[59] F. den Hollander, *Random polymers*, Lectures from the 37th Probability Summer School held in Saint-Flour, 2007. Lecture Notes in Mathematics **1974**, Springer-Verlag, 2009.

[60] J. Hunter, *Renewal theory in two-dimensions: asymptotic results*, Advances in Applied Probability **6** (1974), 546–562 .

[61] A. Kabakçıoğlu, A. Bar and D. Mukamel, *Macroscopic loop formation in circular DNA denaturation*, Phys. Rev. E **85** (2012), 051919.

[62] Y. Kafri, D. Mukamel, *Griffiths singularities in unbinding of strongly disordered polymers*, Phys. Rev. Lett. **91** (2003), 038103.

[63] Y. Kafri, D. Mukamel and L. Peliti, *Why is the DNA denaturation transition first order?*, Phys. Rev. Lett. **85** (2000), 4988-4991.

[64] L. Kaup and B. Kaup, *Holomorphic functions of several variables*. Studies in Mathematics. Walter de Gruyter (1983).

[65] J. F.C. Kingman, *Subadditive ergodic theory*, Ann. Probab. **1**(1973), 882–909.

[66] S. G. Krantz and H. R. Parks, *A primer of real analytic functions*, Second edition, Birkhäuser Advanced Texts: : Basel Textbooks. Birkhäuser Boston, Inc., Boston, MA, 2002.

[67] H. Kunz, R. Livi, *DNA denaturation and wetting in the presence of disorder*, Eur. Phys. Lett. **99** (2012), 30001.

[68] H. Lacoin, *New bounds for the free energy of directed polymers in dimension 1 + 1 and 1 + 2*, Commun. Math. Phys. **294** (2010), 471-503.

[69] H. Lacoin, *The martingale approach to disorder irrelevance for pinning models*, Elec. Comm. Probab. **15** (2010), 418-427.

[70] A. Litan, *A Statistical Mechanical Treatment of the Open Ends in a Double-Stranded Polynucleotide Molecule*, Biopolymers **2** (1964), 279-282.

[71] S. T. Lovett, *Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences*, Mol. Microbiol. **52** (2004), 1243.

[72] R. A. Neher and U. Gerland, *Intermediate phase in DNA melting*, Phys. Rev. E **73** (2006), 030902R.

[73] G. B. Nguyen and N. Pétrélis, *A variational formula for the free energy of the partially directed polymer collapse*, J. Stat. Phys. **151** (2013), 1099-1120.

[74] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons and H. E. Stanley *Long-range correlations in nucleotide sequences* Nature **356** (1992), 168-170.

[75] I. F. Pinelis, *A problem on large deviations in a space of trajectories*, Theory Probab. Appl. **26** (1981), 69-84.

[76] D. Poland and H. A. Schegara, *Phase transitions in one dimension and the HelixCoil transition in polyamino acids*, J. Chem. Phys. **45** (1966), 1456.

[77] D. Poland and H. A. Schegara, *Occurence of a phase transition in nucleic acid models*, J. Chem. Phys. **45** (1966), 1464–1469.

[78] D. Poland and H. A. Scheraga, *Theory of helix-coil transitions in biopolymers;: Statistical mechanical theory of order-disorder transitions in biological macromolecules*, Academic Press, 1970.

[79] C. Richard and A. J. Guttmann, *Poland-Scheraga Models and the DNA Denaturation Transition*, J. Statist. Phys. **115** (2004), 925-947.

[80] E. Rouchka, *RNA secondary structure prediction*, Lecture notes (2003).

[81] S. Shneer and V. Wachtel, *A unified approach to the heavy-traffic analysis of the maximum of random walks*, Theory Probab. Appl. **55** (2011), 332-341.

[82] K. Sture J. Nordholm and S. A. Rice, *Phase transitions and end effects in models of biopolymers*, J. Chem. Phys. **59** (1973), 5605-5613.

[83] M. V. Tamm and S. K. Nechaev , *Unzipping of two random heteropolymers: Ground state energy and finite size effects*, Phys. Rev. E **78** (2008), 011903.

[84] F. Tanaka and H. Takahashi, *Elastic theory of supercoiled DNA*, J. Chem. Phys. **83** (1985), 6017.

[85] R. Thomas, *The denaturation of deoxyribonucleic acids*, Trans. Faraday Soc **50** (1954), 304.

[86] I. Tobias, B. D. Coleman, and W. K. Olson, *The dependence of DNA tertiary structure on end conditions : Theory and implications for topological transitions*, J. Chem. Phys. **101** (1994), 10990-10996.

[87] F. L. Toninelli, *Critical properties and finite-size estimates for the depinning transition of directed random polymers*, J. Stat. Phys. **126** (2007), 1025-1044.

[88] F. L. Toninelli, *A replica-coupling approach to disordered pinning models*, Commun. Math. Phys. **280** (2008), 389-401.

[89] F. L. Toninelli, *Disordered pinning models and copolymers: beyond annealed bounds*, Ann. Appl. Probab. **18** (2008), 1569-1587.

[90] T. Vojta and R. Sknepnek, *Critical points and quenched disorder: from Harris criterion to rare regions and smearing*, Phys. Stat. Sol. B **241** (2004), 2118-2127.

[91] F. Watbled, *Sharp asymptotics for the free energy of 1+1 dimensional directed polymers in an infinitely divisible environment*, Elec. Commun. Probab. **17** (2012), 1-9.

[92] J. D. Watson and F. H. Crick, *Molecular structure of nucleic acids*, Nature **171** (1953), 737-738.

[93] J. H. White and W. R. Bauer, *Calculation of the twist and the writhe for representative models of dna*, Journal of Molecular Biology, **189** (1986), 329–341.

[94] R. Wei, *On the Long-range directed polymer model*, arXiv:1510.02593

[95] E. Yeramian, *Complexity and Tractability. Statistical Mechanics of Helix-Coil Transitions in Circular DNA as a Model-Problem*, Europhysics Letterr **25** (1964), 49-55.