



**HAL**  
open science

# Perturbations of adsorption transition in polymer models

Legrand Alexandre

► **To cite this version:**

Legrand Alexandre. Perturbations of adsorption transition in polymer models. Probability [math.PR]. Université de Nantes, 2021. English. NNT: . tel-03358904

**HAL Id: tel-03358904**

**<https://hal.science/tel-03358904>**

Submitted on 29 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

DE L'UNIVERSITÉ DE NANTES  
COMUE UNIVERSITÉ BRETAGNE LOIRE

Ecole Doctorale N° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Mathématiques et leurs Interactions*

Par

## ALEXANDRE LEGRAND

### Perturbations de la transition d'adsorption dans des modèles de polymères

Thèse présentée et soutenue à l'UNIVERSITÉ DE NANTES, le 16 septembre 2021  
Unité de recherche : Laboratoire de Mathématiques Jean Leray (LMJL)

#### Rapporteurs :

**Yueyun HU**, Professeur, Université Sorbonne Paris Nord  
**Yvan VELENIK**, Professeur, Université de Genève

#### Jury :

Président : **Philippe CARMONA**, Professeur, Université de Nantes  
Examineurs : **Mireille BOUSQUET-MÉLOU**, Directrice de recherche, Université de Bordeaux  
**Giambattista GIACOMIN**, Professeur, Université de Paris  
**Yueyun HU**, Professeur, Université Sorbonne Paris Nord  
**Yvan VELENIK**, Professeur, Université de Genève  
Directeurs de thèse : **Nicolas PÉTRÉLIS**, Maître de conférence, Université de Nantes  
**Quentin BERGER**, Maître de conférence, Sorbonne Université



# Remerciements

Cette thèse est l'aboutissement d'un voyage de plusieurs années, au cours duquel de nombreuses personnes m'ont apporté leur soutien et leurs encouragements, aussi bien dans les meilleurs moments que dans les passages plus cahoteux. Mon investissement et mon engouement pour mon travail n'auraient pu être les mêmes sans elles, et je souhaite leur consacrer ici quelques mots pour les remercier.

D'abord, mes directeurs de thèse, envers qui je ne pourrai jamais exprimer toute ma gratitude. Quentin, c'est grâce à toi que j'ai fait mes premiers pas dans la recherche académique, ton application et ta bienveillance m'ont assuré que j'avais choisi le bon chemin. Nicolas, ta passion et ta gentillesse m'ont poussé à donner le meilleur de mes capacités tout au long de ce voyage. Le temps et l'attention que vous m'avez accordés pendant ces quelques années m'ont profondément touché, m'ont permis de considérablement progresser, et me lancent sur la bonne voie pour devenir un chercheur aussi talentueux et aimable que vous. Je suis fier et reconnaissant de vous avoir eus comme encadrants.

Ensuite, mes rapporteurs Yvan et Yueyun, et mon jury de thèse Mireille, Giambattista et Philippe. Que ce soit par la relecture de mon manuscrit, le comité de suivi de thèse ou votre présence à ma soutenance, je suis gratifié par l'intérêt que vous portez à mon travail et le temps que vous lui avez accordé, et je vous en remercie sincèrement.

Retraçant le parcours qui m'a amené à finir cette thèse et les gens qui l'ont rendue possible, je pense à toutes les personnes qui m'ont accompagné lors de mon séjour à Nantes. Une ancienne tradition des doctorants nantais veut qu'ils remercient exhaustivement leurs camarades, et je ne souhaite pas y déroger.

Je commence par celles et ceux plus anciens, qui m'ont chaleureusement accueilli dans leur laboratoire. Solène, qui m'a (littéralement) ouvert sa porte et a partagé avec moi de nombreuses parties de jeux où elle était bien meilleure que moi. Caro, dont le sourire et la gaieté m'ont tout de suite aidé à m'intégrer au groupe. Germain, avec qui j'ai échangé des discussions aussi impromptues que passionnantes. Matthieu, dont je me suis rapproché par le heureux hasard de nos goûts musicaux. Côme et Amiel, qui m'ont impressionné, voire rendu jaloux, par leur flegme naturel sans égal (non sans me remémorer la K-Fêt). Maha, Zeinab, Claire et Arthur, avec qui chaque conversation ou pause-café fut toujours très joviale et plaisante. Et Thomas, notre aîné à tous, qui a répondu à tant de mes questions de nouvel arrivant et m'a également donné le courage de me lancer dans une de mes passions.

Suivent celles et ceux qui sont arrivés en même temps que moi ou après. D'abord, mon éternel co-bureau, Fabien, avec qui j'ai tant partagé : de nos blocages de jeunes chercheurs à notre passion pour nos enseignements, de nos frustrations à la lecture de nos mails à nos éclats de rire en soirée, je suis heureux d'avoir fait une bonne partie de ce trajet avec toi. Mon nouveau co-bureau, Adrien, qui est rapidement devenu un très bon ami au labo et en-dehors, et que je regrette de devoir quitter si tôt. Antoine, qui m'a ravi en me rejoignant au sport mais aussi au bar. Adrian, avec son entrain permanent et ses pointes d'humour (Double-pile aura sa vengeance). Anh, sans qui nous n'aurions pu organiser les Rencontres Doctorales avec succès. Anthony, qui a eu la gentillesse de m'accompagner et me remplacer aux conseils de labo. Samuel, dont les curiosités mathématiques ont initié maintes discussions autour d'un tableau et d'un café. Azeddine, toujours très souriant et chaleureux. Trung, Meïssa, Maël, Mohamad, qui m'ont accueilli amicalement lors de mon arrivée à Nantes. Et enfin Jérôme, Silvère, Hai et tous les nouveaux qui prennent notre relève et continueront de souder notre joyeux groupe de doctorants.

J'adresse également un mot à toutes les personnes qui font vivre ce magnifique laboratoire. Vincent, Annick, Brigitte, Stéphanie, Alexandra, Béatrice, Anaïs, Erwan, Eric, Saïd, Claude, Anh, Colette et beaucoup d'autres : que ce soit pour les nombreuses tâches administratives dont je n'ai vu que la surface, pour la vie scientifique et l'organisation des séminaires et autres évènements, pour l'installation et le bon fonctionnement de notre matériel, pour la mise à disposition d'une machine à café à l'étage du bas, ou simplement l'amabilité et la bonne humeur qui planent tous les jours entre ces murs, je vous remercie d'avoir créé ensemble un environnement de travail aussi accueillant et agréable pour nous tous.

Cette étape à Nantes m'a enfin permis de rencontrer ma meilleure amie Marine : ta bienveillance et ta générosité m'ont été très précieuses cette année. Que ce soit au marché ou au bord de l'Erdre, devant un film ou autour d'un thé, chacun de nos moments m'enchantent et m'encouragent ; alors j'espère qu'ils continueront encore en dépit de la distance.

Je remonte le temps d'encre quelques années pour parler de mon temps à Paris, et de toutes les personnes qui m'entourent depuis le début du doctorat ou même avant. Mon passage au LPSM fut un peu bref, mais j'ai là-aussi été accueilli par des personnes remarquables, William, Michel, Alexandra, Chenguang et tous les autres. Par pure coïncidence nous sommes plusieurs à nous installer à Toulouse l'an prochain, et j'ai hâte de vous rejoindre là-bas. J'adresse un mot en particulier à mes compagnons de Master, Felix, Isao, Colin et Lucas. C'est une joie de parcourir la même route avec vous depuis plusieurs années, et ce ne sont pas quelques postdocs outre-mer qui nous feront nous perdre de vue.

Enfin, de tous mes anciens voisins parisiens, certains d'entre vous restent aujourd'hui parmi mes amis les plus chers. Arthur, dont je me suis beaucoup rapproché par le plus curieux des hasards, mais depuis j'ai toujours pu compter sur ta bienveillance et ton soutien indéfectible. Pierre et Marion, qui m'invitent encore à la Dhuy et à Paris, et c'est toujours une grande joie de vous revoir quand je peux venir. Clémence, qui m'a convié à de superbes weekends à la plage. Mais aussi Etienne, Mathilde, Babeth, Charles, Marion et toutes celles et ceux que je retrouve avec plaisir à chacune de nos occasions.

---

Je ne peux pas clore ces remerciements sans adresser quelques mots aux personnes qui suivent mon voyage depuis le début. À Papa, qui très tôt m’as transmis ta passion des maths alors que ni toi ni moi ne soupçonnions qu’elle me mènerait aussi loin. À Maman, qui m’as appris à faire tant de choses que j’adore en-dehors des maths, et même me conseille encore en pâtisserie. Votre soutien et votre amour ont été inestimables pendant toutes ces années, surtout envers le fils maladroit et inapte du calendrier que je suis. À Pierre-Louis, avec qui je suis heureux de toujours partager autant de centres d’intérêt et de sujets de conversation depuis des années. À Paul, qui malgré son plaisir à me provoquer aux repas de famille reste un petit frère admirable et adorable. À mes cousins-cousines aussi, Natacha, Carole, Sophie et toute la famille, pour tous les bons moments qu’on a passés ensemble.

À vous, qui m’avez accompagné pendant mon travail, m’avez apporté votre soutien ou votre amitié, ou m’avez seulement entouré quand j’en avait besoin ; vous avez contribué, parfois sans vous en rendre compte, à la réalisation de cette thèse. “Merci” est un mot bien ordinaire, mais qui porte toute ma gratitude et mon affection envers vous tous et toutes. Alors, merci à vous.



# Table des matières / Table of Contents

<b>Introduction et vue d'ensemble</b>	<b>9</b>
0.1 IPDSAW au voisinage d'un mur : un modèle à deux interactions . . . . .	15
0.2 Le modèle généralisé de Poland-Scheraga pour la dénaturation de l'ADN .	31
<b>Presentation and outline of the thesis</b>	<b>57</b>
<b>I The Interacting Partially Directed Self-Avoiding Walk</b>	<b>65</b>
<b>1 Surface transition in the IPDSAW adsorbed along a wall</b>	<b>67</b>
1.1 Introduction . . . . .	67
1.2 Description of the model and phase diagram . . . . .	68
1.3 Collapsed phase: restriction to the single-bead model . . . . .	73
1.4 Proof of Theorem 1.3.2 . . . . .	79
1.5 Proof of Theorem 1.2.2 and Proposition 1.3.1 . . . . .	103
1.6 Proofs of technical results . . . . .	108
<b>Appendices to Chapter 1</b>	<b>121</b>
1.A The wetting model . . . . .	121
1.B An FKG inequality . . . . .	122
<b>2 Sharp asymptotics for the collapsed IPDSAW</b>	<b>125</b>
2.1 Introduction . . . . .	125
2.2 Main results . . . . .	127
2.3 Preparations . . . . .	130
2.4 Proof of Theorem 2.2.1 . . . . .	135
2.5 Proof of Proposition 2.4.5 . . . . .	141
2.6 Proof of Theorem 2.2.2 . . . . .	155
<b>Appendices to Chapter 2</b>	<b>158</b>
2.A Local limit estimates, proof of Lemma 2.5.12 . . . . .	158
2.B Functional estimates, proof of Proposition 2.5.4 . . . . .	161



<b>II</b>	<b>The generalized Poland-Scheraga model</b>	<b>163</b>
<b>3</b>	<b>The gPS model: the question of disorder (ir)-relevance</b>	<b>165</b>
3.1	Introduction . . . . .	165
3.2	Presentation of the results: the question of disorder relevance . . . . .	170
3.3	Disorder irrelevance . . . . .	178
3.4	Disorder relevance: smoothing of the phase transition when $\mathbb{P} \neq \mathbb{P}_{\pm x}$ . . . . .	187
3.5	Disorder relevance: shift of the critical point when $\mathbb{P} \neq \mathbb{P}_{\pm x}$ . . . . .	191
3.6	Disorder relevance: shift of the critical point when $\mathbb{P} = \mathbb{P}_{\pm x}$ . . . . .	202
	<b>Appendices to Chapter 3</b>	<b>209</b>
3.A	Estimates on the homogeneous gPS model . . . . .	209
3.B	Some properties of bivariate renewals . . . . .	212
3.C	Computation of the second moment in the Gaussian case . . . . .	213
<b>4</b>	<b>Scaling limit of the gPS model</b>	<b>217</b>
4.1	Introduction . . . . .	217
4.2	Convergence of the field $M_n$ to $\mathcal{M}$ : proof of Theorem 4.1.2 . . . . .	227
4.3	Covariance measure of $\mathcal{M}$ and stochastic integral . . . . .	235
4.4	Convergence of terms of order $k \geq 1$ to an integral . . . . .	249
	<b>Appendices to Chapter 4</b>	<b>255</b>
4.A	Technical estimates on $\overline{M}_n, n \geq 1$ . . . . .	255
	<b>Conclusion and perspectives</b>	<b>257</b>
	<b>Appendix: Notation</b>	<b>261</b>
	<b>Bibliography</b>	<b>263</b>

# Introduction et vue d'ensemble

Cette thèse s'intéresse aux modèles mathématiques de polymères. Cette grande famille de molécules, qui inclut les brins d'ADN, les protéines ou les plastiques pour n'en citer que quelques-unes, occupe une place importante dans la littérature de physique, de chimie ou de biologie, car celles-ci sont au centre de nombreux phénomènes observables d'intérêt scientifique. Les étudier d'un point de vue mathématique nous permet d'approfondir notre compréhension de ces phénomènes, en exhibant les mécanismes qui les engendrent et en développant des modèles et des cadres d'étude les plus généraux possibles.

Ces dernières années la communauté mathématique a beaucoup étudié une famille de modèles appelés modèles *d'accrochage* ou de *mouillage* (souvent dénommés *pinning* et *wetting* respectivement), voir en particulier [84, 92, 100]. Ceux-ci ont l'intérêt de représenter assez fidèlement un phénomène physico-chimique de *transition de phase*, tout en restant abordables (voire résolubles) dans un cadre mathématique. Nous allons d'abord présenter le phénomène en question au travers d'un exemple particulier de modèle de wetting, la *marche auto-évitante, partiellement dirigée au dessus d'un mur dans  $\mathbb{Z}^2$* . Ensuite, nous donnerons un aperçu d'un ensemble plus général de modèles présentant la même transition d'adsorption, et nous nous interrogerons enfin sur les éléments de la modélisation qui peuvent l'affecter et à quelle échelle.

## Construction d'un modèle de wetting de polymère : PDSAW près d'un mur

Les *polymères* sont des macro-molécules, composées d'un grand nombre de briques élémentaires d'un ou plusieurs atomes, appelés *monomères*, arrangées selon une structure de *chaîne* linéaire : un des exemples les plus connus est le brin d'ADN, composé à partir de bases habituellement nommées A, T, C et G. Pour les modéliser dans un cadre mathématique, un objet naturellement adapté est la *marche sur un réseau*, chaque site occupé par la marche correspondant à un monomère, et sa trajectoire complète représentant la géométrie du polymère dans l'espace. Considérons par exemple une trajectoire de longueur  $L \in \mathbb{N}$  dans  $\mathbb{Z}^2$  : les  $L + 1$  sites par lesquels elle passe désignent les positions des monomères successifs constituant la molécule<sup>†</sup>. Comme deux vrais monomères ne peuvent pas physiquement

---

<sup>†</sup>. Afin de rester consistante entre les parties en français et en anglais, cette thèse utilise les notations mathématiques anglo-saxonnes :  $\log$  désigne le logarithme népérien,  $\mathbb{N}$  l'ensemble des entiers naturels non nuls... Voir le chapitre Appendix : Notation pour plus de détails.

occuper le même site dans l'espace, on peut naturellement se restreindre aux trajectoires de marches *auto-évitantes*, ou *Self-Avoiding Walks* (SAW). Nous pouvons bien sûr jouer avec cette première modélisation en imposant des contraintes géométriques supplémentaires, par exemple nous limiter aux marches *partiellement dirigées* dans une direction, disons la droite, i.e. ne réalisant aucun pas dans la direction opposée, dans ce cas-ci la gauche (on parle de *Partially Directed Self-Avoiding Walk*, PDSAW); ou tenir compte d'autres éléments dans l'environnement du polymère, comme un objet ou un mur que le polymère ne peut pas traverser (voir Fig. 0.1).

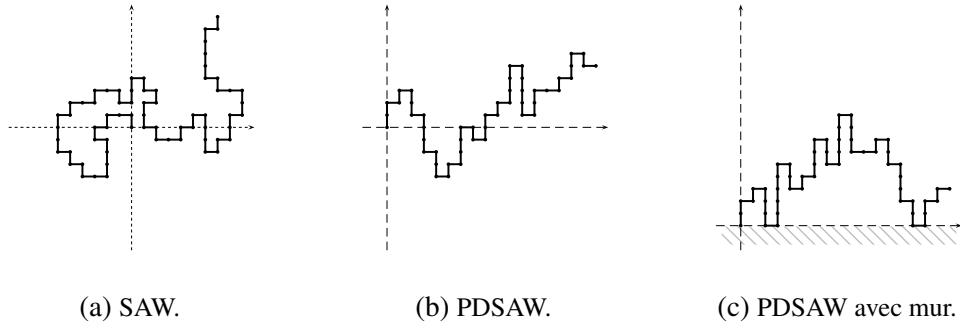


FIGURE 0.1: Représentation de polymères par des marches dans  $\mathbb{Z}^2$ . La première est auto-évitante (SAW), la seconde est additionnellement partiellement dirigée vers la droite (PDSAW), et la dernière est en plus contrainte à rester au-dessus d'un mur dur.

L'objectif d'un modèle de wetting est d'étudier le comportement d'un polymère situé à proximité d'un autre objet (comme la paroi d'une cellule, ou une ligne d'impuretés dans l'environnement) et plus précisément de déterminer sa *localisation* : le polymère s'éloigne-t-il de l'objet, ou reste-t-il à proximité ? À quelle distance se positionne-t-il ? Combien de fois touche-t-il l'objet ? Afin d'apporter des réponses à ces questions, la troisième représentation proposée ci-dessus (voir Fig. 0.1c) est naturellement adaptée, et sera même l'objet d'étude du Chapitre 1 de cette thèse.

Considérons donc les trajectoires de marche dans  $\mathbb{Z}^2$  de longueur  $L \in \mathbb{N}$  fixée, auto-évitantes, partiellement dirigées vers la droite et contraintes à rester au-dessus d'un mur horizontal. L'ensemble  $\Omega_L^{\text{wet}}$  des trajectoires satisfaisant ces hypothèses peut être écrit rigoureusement,

$$\Omega_L^{\text{wet}} := \left\{ (x_i)_{i=0}^L \in (\mathbb{Z}^2)^{L+1} \left| \begin{array}{l} x_0 = (0, 0), x_i \neq x_j \quad \forall 0 \leq i < j \leq L \\ x_i - x_{i-1} \in \{\uparrow, \downarrow, \rightarrow\} \quad \forall 1 \leq i \leq L, \\ x_i \in \mathbb{Z} \times \mathbb{N} \quad \forall 1 \leq i \leq L \end{array} \right. \right\}, \quad (0.0.1)$$

où on note  $\uparrow = (0, 1)$ ,  $\downarrow = (0, -1)$  et  $\rightarrow = (1, 0)$ .

Pour aller plus loin dans notre modélisation, nous devons nous intéresser aux *interactions* physico-chimiques entre le polymère et le mur. Plus précisément, supposons que le polymère interagit avec le mur à chaque *contact*, c'est-à-dire à chaque segment  $[x_{i-1}, x_i]$ ,  $1 \leq i \leq L$ ,

qui longe le mur :  $[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}$ . Chacune de ces interactions fait intervenir une *énergie d'interaction* que nous noterons  $\delta \in \mathbb{R}$  : si  $\delta > 0$ , on peut la voir comme une récompense accordée au polymère pour chaque contact avec le mur, et si  $\delta < 0$  il s'agit plutôt d'un coût à payer pour maintenir le contact (notons que cette convention de signe est opposée à celle utilisée le plus souvent dans la littérature physique, mais simplifie légèrement les formules ultérieures). Pour tout  $x \in \Omega_L^{\text{wet}}$ , notons  $H_{L,\delta}^{\text{wet}}(x)$  le total des énergies d'interaction du polymère ayant pour trajectoire  $x$ , soit

$$H_{L,\delta}^{\text{wet}}(x) := \sum_{i=1}^L \delta \mathbf{1}_{\{[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}}. \quad (0.0.2)$$

Avec cette convention de signe, la loi de Boltzmann nous affirme que dans un système physique, plus la récompense en énergie  $H_{L,\delta}^{\text{wet}}(x)$  associée à une trajectoire (ou *configuration*)  $x$  est élevée, plus celle-ci sera privilégiée par le système par rapport aux autres. Plus précisément dans notre modèle, la trajectoire du polymère est donnée par une variable aléatoire distribuée selon la loi

$$\mathbf{P}_{L,\delta}^{\text{wet}}(x) := \frac{e^{H_{L,\delta}^{\text{wet}}(x)}}{Z_{L,\delta}^{\text{wet}}}, \quad x \in \Omega_L^{\text{wet}}, \quad (0.0.3)$$

qu'on appelle *mesure de Gibbs* ou *mesure polymère*, où on définit le terme de renormalisation

$$Z_{L,\delta}^{\text{wet}} := \sum_{x \in \Omega_L^{\text{wet}}} e^{H_{L,\delta}^{\text{wet}}(x)}, \quad (0.0.4)$$

qu'on appelle *fonction de partition* du système.

## Le phénomène de transition de phase

Notre objectif est d'étudier la localisation du polymère dans notre modélisation, donc en particulier la localisation des trajectoires sous la mesure de probabilité  $\mathbf{P}_{L,\delta}^{\text{wet}}$ . Pour cela nous utilisons un outil courant de mécanique statistique : le taux de croissance exponentielle de  $Z_{L,\delta}^{\text{wet}}$ , aussi appelé *énergie libre* du système,

$$\mathbf{F}^{\text{wet}}(\delta) := \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\delta}^{\text{wet}}. \quad (0.0.5)$$

Dans cette introduction nous admettons que cette fonction est bien définie. Il est également possible de démontrer qu'elle vérifie un certain nombre de propriétés : elle est convexe, croissante, continue et positive sur  $\mathbb{R}$ . De plus, une étude un peu plus poussée permet de remarquer que si  $\delta$  tend vers  $-\infty$ , alors  $\mathbf{F}^{\text{wet}}(\delta)$  est localement constante (car  $Z_{L,\delta}^{\text{wet}}$  ne varie que très peu en  $-\delta$  grand). Inversement si  $\delta$  tend vers  $+\infty$ , alors  $\mathbf{F}^{\text{wet}}(\delta)$  tend vers  $+\infty$ . En rassemblant toutes ces observations, on en conclut que la fonction  $\delta \mapsto \mathbf{F}^{\text{wet}}(\delta)$  vérifie une autre propriété très singulière : il existe un nombre réel, notons-le  $\delta_c^{\text{wet}}$ , tel que

- $\delta \mapsto \mathbf{F}^{\text{wet}}(\delta)$  est constante sur  $(-\infty, \delta_c^{\text{wet}})$ ,

—  $\delta \mapsto F^{\text{wet}}(\delta)$  est strictement croissante sur  $(\delta_c^{\text{wet}}, +\infty)$ .

Un tel point de rupture de régularité dans l'énergie libre est appelé un *point critique* du système (voir Fig. 0.2b plus bas). Remarquer l'existence de ce point critique —et le changement d'expression de  $F$  de part et d'autre— prend toute son importance quand on observe également que la fonction  $F$  est intimement liée aux propriétés de la mesure  $\mathbf{P}_{L,\delta}^{\text{wet}}$ . Plus précisément, pour un polymère de longueur  $L \in \mathbb{N}$ , considérons  $N_L^{\text{wet}}(\delta)$  la proportion moyenne du polymère en contact avec le mur :

$$N_L^{\text{wet}}(\delta) := \frac{1}{L} \mathbf{E}_{L,\delta}^{\text{wet}} \left[ \sum_{i=1}^L \mathbf{1}_{\{[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}} \right]. \quad (0.0.6)$$

On peut alors démontrer que pour tout  $\delta \neq \delta_c^{\text{wet}}$ , on a

$$N_L^{\text{wet}}(\delta) \xrightarrow{L \rightarrow \infty} (F^{\text{wet}})'(\delta). \quad (0.0.7)$$

Cela découle d'un calcul direct, au moins là où  $F^{\text{wet}}$  est dérivable. Pour ne pas complexifier cette introduction, nous ne parlons pas du cas  $\delta = \delta_c^{\text{wet}}$  pour l'instant. Nous pouvons déjà identifier, en associant les propriétés de  $F^{\text{wet}}$  (voir Fig. 0.2b) et (0.0.7), que deux comportements très différents sont possibles pour le polymère sous la mesure  $\mathbf{P}_{L,\delta}^{\text{wet}}$  :

- Soit  $\delta < \delta_c^{\text{wet}}$  ; alors pour les trajectoires typiques sous  $\mathbf{P}_{L,\delta}^{\text{wet}}$ , la fraction des monomères en contact avec le mur tend vers 0 quand  $L$  devient grand, donc le nombre de contacts avec le mur est négligeable devant la taille du polymère. On appelle  $\mathcal{D} := (-\infty, \delta_c^{\text{wet}})$  la *phase désorbée*, ou *délocalisée*.
- Soit  $\delta > \delta_c^{\text{wet}}$  ; alors pour les trajectoires typiques sous  $\mathbf{P}_{L,\delta}^{\text{wet}}$ , le nombre de monomères en contact avec le mur est proportionnel à la longueur du polymère. On appelle  $\mathcal{A} := (\delta_c^{\text{wet}}, +\infty)$  la *phase adsorbée*, ou *localisée*.

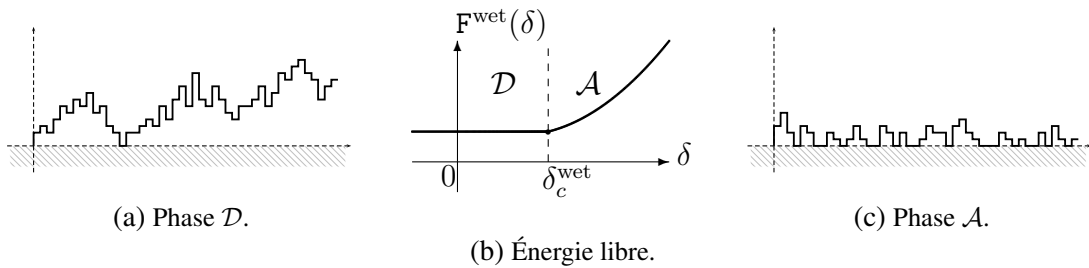


FIGURE 0.2: Diagramme de phase du modèle de wetting, et représentation d'une trajectoire typique dans chacune des deux phases  $\mathcal{D}$  et  $\mathcal{A}$ . Dans  $\mathcal{D}$ , le polymère ne touche que rarement le mur. Dans  $\mathcal{A}$ , le polymère réalise un grand nombre de contacts, et en conséquence s'éloigne très peu du mur.

Un tel changement de comportement du système au point critique  $\delta_c^{\text{wet}}$  —c'est-à-dire au point où l'énergie libre  $F^{\text{wet}}$  change d'expression— est appelé une *transition de phase*. Dans notre modèle, le polymère passe d'un régime  $\mathcal{D}$  où il n'interagit que rarement avec le

mur, en particulier il reste à distance de celui-ci ; à un régime  $\mathcal{A}$  où il effectue de nombreux contacts et donc reste proche du mur (voir Fig. 0.2). On parle de *transition d'adsorption* ou *d'accrochage* du polymère, et une majeure partie de cette thèse s'intéressera à ce phénomène.

## Généralisation : les modèles de pinning

Comme mentionné plus haut, le modèle que nous venons de construire est appelé un *modèle de wetting*. Avec le temps, des modèles similaires ont été formalisés dans un cadre de plus en plus général, jusqu'à former la famille des modèles dits de *pinning* (voir [73] puis [84, Chapter 1]), en délaissant la description du polymère par une marche dans un réseau, au profit de l'utilisation des *processus de renouvellement*. L'idée derrière cette nouvelle représentation est que, plutôt que de vouloir décrire toute la géométrie du polymère dans l'espace par une marche, il nous suffit seulement de connaître les indices des monomères qui interagissent avec le mur.

Un processus de renouvellement  $\tau = (\tau_i)_{i \geq 1}$  est une suite de variables aléatoires telle que ses incréments  $(\tau_i - \tau_{i-1})$ ,  $i \geq 1$ , sont i.i.d. à valeur dans  $\mathbb{N} \cup \{+\infty\}$  (avec la convention  $\tau_0 = 0$ ). Par habitude, on le note aussi de manière ensembliste  $\tau \subset \mathbb{N}$ . En particulier, si on considère  $(S_n)_{n \geq 1}$  une marche aléatoire assez générique dans  $\mathbb{Z}^d$ ,  $d \geq 1$ , et si on note  $\tau$  la suite de ses temps de retour, disons sur la ligne  $\mathbb{Z} \times \{0\}^{d-1}$ , alors  $\tau$  est un processus de renouvellement. Cette structure de renouvellement apparaîtra très généralement, même si on choisit une marche (partiellement) dirigée, ou si on rajoute certaines contraintes ou conditionnements (comme rester dans le demi-plan positif en dimension  $d = 2$ ). Alors, on peut s'en servir pour construire un modèle très similaire à celui présenté ci-dessus, en attribuant une énergie d'interaction  $\delta \in \mathbb{R}$  à chaque point de renouvellement (c'est-à-dire à chaque monomère qui est en contact avec la ligne, ou "accroché"). Pour une marche de longueur  $L \in \mathbb{N}$  réalisant la suite de temps de contacts  $\tau$ , on lui associe alors l'énergie totale

$$H_{L,\delta}^{\text{pin}}(\tau) := \sum_{i=1}^L \delta \mathbf{1}_{\{i \in \tau\}},$$

et on définit la mesure de Gibbs  $\mathbf{P}_{L,\delta}^{\text{pin}}$  pour laquelle la probabilité d'une suite  $\tau$  est proportionnelle à  $e^{H_{L,\delta}^{\text{pin}}(\tau)} \mathbf{P}(\tau)$ , où  $\mathbf{P}$  est la loi initiale du processus de renouvellement ; en particulier, cela nous permet également de définir une fonction de partition  $Z_{L,\delta}^{\text{pin}}$  et une énergie libre  $\mathbf{F}^{\text{pin}}$  dans ce cadre, similairement à celles du PDSAW près d'un mur.

Les modèles obtenus de cette façon (pour n'importe quels processus de renouvellement  $\tau$  dans  $\mathbb{N}$ ) forment la famille des modèles de *pinning*. Avec un peu de réécriture (que nous ne détaillons pas ici), nous pouvons reformuler le modèle PDSAW près d'un mur défini plus haut dans ces termes, en construisant un processus de renouvellement sous-jacent à partir des indices de contacts avec le mur (soit  $\tau(x) = \{i \in \mathbb{N}, [x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}$ ,  $x \in \Omega_L^{\text{wet}}$ ).

Le premier résultat important concernant les modèles de pinning est que, sous des hypothèses minimales (i.e. la queue de  $\tau_1$  doit être "à variations régulières", nous expliciterons cette hypothèse plus bas), ils sont tous sujets à la même transition de phase de *localisation*

ou *d'accrochage* — que nous avons appelé “transition d'adsorption” dans le modèle PDSAW avec mur—, entre une phase où l'énergie libre  $F^{\text{pin}}$  est constante (précisément égale à 0), et une autre où elle est strictement croissante. Bien sûr, ses détails comme la valeur du point critique  $\delta_c^{\text{pin}}$  ou le comportement asymptotique de  $F^{\text{pin}}(\delta)$  autour de  $\delta_c^{\text{pin}}$  peuvent varier avec la loi  $\mathbf{P}$  du renouvellement ; mais on les connaît explicitement (voir [84, Theorem 2.1]), et le même phénomène de (dé)-localisation apparaît dans tous les modèles : si  $\delta < \delta_c^{\text{pin}}$ , le nombre de contacts réalisés par le polymère (i.e. de points dans  $\tau \cap [1, L]$  sous  $\mathbf{P}_{L,\delta}^{\text{pin}}$ ) est négligeable devant sa longueur  $L$  ; et si  $\delta > \delta_c^{\text{pin}}$ , alors il vaut au premier ordre  $(F^{\text{pin}})'(\delta)L$  pour  $L$  grand avec  $(F^{\text{pin}})'(\delta) > 0$  (voir [84, Theorem 2.4]).

### Problématique : quelles sont les perturbations possibles de cette transition de localisation ?

Le phénomène de transition de phase présenté ci-dessus se retrouve donc dans une très grande famille de modèles de polymères, et ses caractéristiques précises (point critique, comportement critique de l'énergie libre) sont même très bien identifiées aujourd'hui. Devant de tels résultats, il est naturel de se poser la question de leur robustesse si on altère ces modèles en leur rajoutant d'autres éléments ou détails : supposons que le polymère peut être soumis à d'autres interactions, avec un autre mur ou objet, ou même d'une toute autre nature ; supposons que toutes les interactions de contact ne se valent pas, à cause d'irrégularité à la surface du mur, ou d'inhomogénéités dans la composition du polymère ; supposons que la trajectoire du polymère est fortement contrainte par d'autres obstacles ou restrictions dans l'espace environnant... Sous ces hypothèses supplémentaires, observe-t-on toujours une transition d'accrochage ? A-t-elle lieu au même point critique, et l'énergie libre a-t-elle le même comportement critique ?

Cette thèse s'intéresse à ces questionnements à travers l'étude de deux modèles particuliers, chacun introduisant un phénomène supplémentaire différent qui va avoir des répercussions sur la transition d'adsorption ou d'accrochage.

Pour notre première approche, nous nous concentrerons sur le modèle PDSAW avec mur présenté plus haut, dans lequel nous supposerons qu'un autre type d'interaction a lieu : si deux monomères, non-consécutifs dans le polymère, se retrouvent dans des sites adjacents de  $\mathbb{Z}^2$ , alors ils interagissent entre eux. Cela donne un modèle à deux interactions très différentes (l'une est avec un objet fixe, situé sur la ligne horizontale ; l'autre est entre le polymère et lui-même, et peut avoir lieu n'importe où), qui peuvent chacune influencer drastiquement la géométrie et la localisation du polymère et de manières distinctes. Nous montrerons que pour cette perturbation, la transition de phase d'adsorption peut devenir une *transition de surface*. Nous verrons également que ce modèle fait naturellement apparaître un phénomène dit de *pre-wetting*, une autre perturbation du wetting qui intéresse beaucoup la littérature de mécanique statistique.

Pour notre seconde approche, nous utiliserons le formalisme des modèles de pinning avec des processus de renouvellements, et nous nous intéresserons plutôt à une perturbation à l'échelle microscopique du système, plus précisément la présence d'inhomogénéités dans

la composition du polymère, et donc des énergies d'interaction variables d'un monomère à l'autre. Une autre motivation de ce travail sera son application à l'étude de la molécule d'ADN, qui est sujette à un phénomène dit de *dénaturation* lequel peut être représenté par un modèle de pinning, et dont les brins, comme mentionné plus haut, sont composés à partir de plusieurs types de monomères distincts. Nous modéliserons cela avec un modèle *désordonné*, dans lequel le *désordre* sera muni de corrélations non-triviales jouant un rôle décisif dans son influence sur la transition d'accrochage.

La suite de cette introduction est séparée en deux parties, une pour chaque approche, présentant l'état de l'art, les motivations et nos résultats sur le modèle en question.

## 0.1 IPDSAW au voisinage d'un mur : un modèle à deux interactions

Comme annoncé ci-dessus, ce premier travail s'intéresse à un modèle de polymère à deux interactions : les monomères interagissent à la fois avec un mur dur (que la molécule ne peut pas traverser), et avec les autres monomères adjacents. Celui-ci peut se voir comme la superposition exacte de deux modèles de polymère différents : celui de wetting présenté précédemment, et un modèle de polymère interagissant avec son solvant.

### 0.1.1 Le modèle IPDSAW et la transition d'effondrement

Commençons par présenter le modèle de polymère interagissant avec son solvant, introduit dans [155] (voir aussi [40, 131]) sous le nom *Interacting Partially Directed Self-Avoiding Walk* (IPDSAW). L'objectif de ce modèle est de représenter un polymère plongé dans un "solvant pauvre", c'est-à-dire avec lequel les monomères qui le composent ont une faible affinité chimique. Pour minimiser ses interactions avec le solvant, le polymère a la possibilité de se replier sur lui-même, empêchant ainsi le solvant de s'introduire entre ses monomères et diminuant la pénalisation énergétique subie. Par conséquent, on peut modéliser ce phénomène de manière équivalente en considérant que le polymère interagit avec lui-même à chaque *auto-contact*, plus précisément qu'il est récompensé à chaque fois que deux monomères, non-successifs dans le polymère, sont adjacents dans l'espace.

#### Construction du modèle

Comme le nom IPDSAW l'indique, nous nous intéressons aux trajectoires partiellement dirigées et auto-évitantes dans  $\mathbb{Z}^2$ , comme représentées dans la Figure 0.1b plus haut. L'ensemble des trajectoires de longueur  $L \in \mathbb{N}$  satisfaisant ces contraintes s'écrit explicitement,

$$\left\{ (x_i)_{i=0}^L \in (\mathbb{Z}^2)^{L+1} \mid \begin{array}{l} x_0 = (0, 0), x_i \neq x_j \quad \forall 0 \leq i < j \leq L \\ x_i - x_{i-1} \in \{\uparrow, \downarrow, \rightarrow\} \quad \forall 1 \leq i \leq L, \end{array} \right\}, \quad (0.1.1)$$



où on rappelle  $\uparrow = (0, 1)$ ,  $\downarrow = (0, -1)$  et  $\rightarrow = (1, 0)$  (remarquons qu'il s'obtient simplement en levant la contrainte imposée par le mur dans  $\Omega_L^{\text{wet}}$  en (0.0.1)).

On remarque qu'une telle trajectoire est composée d'une suite de bandes verticales, appelées *stretches*, orientées vers le haut ou vers le bas ou de longueur nulle, et séparées les unes des autres par les pas horizontaux (vers la droite) du polymère. Cette description d'une trajectoire nous sera très pratique par la suite; c'est pourquoi nous décrivons  $\Omega_L$ , l'ensemble des trajectoires auto-évitantes partiellement dirigées vers la droite, en décomposant une trajectoire selon son nombre de pas horizontaux et les longueurs (signées) des stretches verticaux entre ceux-ci :

$$\Omega_L := \bigcup_{N \geq 1} \mathcal{L}_{N,L} := \bigcup_{N \geq 1} \left\{ (\ell_i)_{i=1}^N \in \mathbb{Z}^N ; \sum_{i=1}^N |\ell_i| = L - N \right\}, \quad (0.1.2)$$

où pour tout  $1 \leq N \leq L$ ,  $\mathcal{L}_{N,L}$  désigne l'ensemble des trajectoires comportant exactement  $N$  pas horizontaux, et pour tout  $\ell \in \mathcal{L}_{N,L}$  et  $1 \leq i \leq N$ ,  $\ell_i \in \mathbb{Z}$  est la longueur et l'orientation du stretch vertical précédant le  $i$ -ème pas horizontal (voir Fig. 0.3). Notons que dans cette décomposition, on a contraint le dernier pas de la trajectoire à être horizontal, mais cela n'affecte les résultats suivants qu'à la marge.

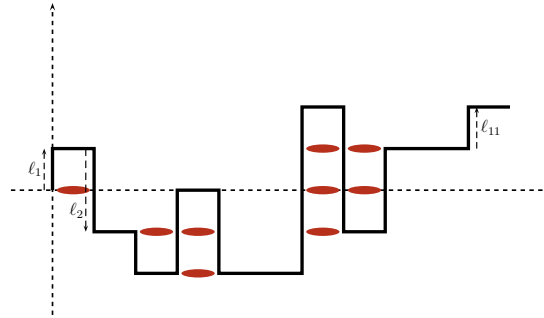


FIGURE 0.3: Représentation d'une trajectoire IPDSAW de longueur  $L = 29$  et d'extension horizontale  $N = 11$ , terminant par un pas horizontal et comportant 11 stretches verticaux. Lorsque deux monomères non-successifs sont adjacents dans  $\mathbb{Z}^2$ , ils réalisent un auto-contact (représenté par une ellipse rouge)

À présent, intéressons-nous au nombre d'*auto-contacts* réalisés par une trajectoire  $\ell \in \Omega_L$ , c'est-à-dire le nombre de paires de monomères non-consécutifs mais adjacents dans  $\mathbb{Z}^2$ . On remarque qu'avec la description ci-dessus, les auto-contacts ne peuvent avoir lieu qu'entre deux stretches verticaux successifs, et à condition qu'ils soient de sens opposés (voir Fig. 0.3). Plus précisément, pour  $1 \leq N \leq L$ ,  $\ell \in \mathcal{L}_{N,L}$ , et  $1 \leq i \leq N - 1$ , le nombre de contacts entre les stretches  $\ell_i$  et  $\ell_{i+1}$  est donné par  $\ell_i \tilde{\wedge} \ell_{i+1} := \min(|\ell_i|, |\ell_{i+1}|) \mathbf{1}_{\{\ell_i \ell_{i+1} \leq 0\}}$ . Par conséquent, le nombre total d'auto-contacts réalisé par une trajectoire  $\ell \in \mathcal{L}_{N,L}$  vaut  $\sum_{i=1}^{N-1} (\ell_i \tilde{\wedge} \ell_{i+1})$ . Soit  $\beta \geq 0$  l'énergie d'interaction accordée à chaque auto-contact du

polymère ; alors, pour un polymère de longueur  $L \in \mathbb{N}$ , la mesure polymère (obtenue à partir de la loi de Boltzmann) s'écrit

$$\mathbf{P}_{L,\beta}(\ell) := \frac{1}{Z_{L,\beta}} \exp \left( \beta \sum_{i=1}^{N_L(\ell)-1} (\ell_i \tilde{\wedge} \ell_{i+1}) \right), \quad \ell \in \Omega_L, \quad (0.1.3)$$

où  $N_L(\ell)$  dénote l'extension horizontale de la trajectoire  $\ell \in \Omega_L$  (i.e.  $\ell \in \mathcal{L}_{L,N_L(\ell)}$ ), et la fonction de partition  $Z_{L,\beta}$  est donnée par

$$Z_{L,\beta} := \sum_{N=1}^L \sum_{\ell \in \mathcal{L}_{L,N}} \exp \left( \beta \sum_{i=1}^{N-1} (\ell_i \tilde{\wedge} \ell_{i+1}) \right). \quad (0.1.4)$$

Comme dans le modèle PDSAW avec mur, on peut définir l'énergie libre du modèle IPDSAW par

$$F(\beta) := \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\beta}. \quad (0.1.5)$$

Cette limite est bien définie, voir [52, 130]. De plus c'est la limite d'une suite de fonctions convexes et croissantes, donc  $F$  est convexe, croissante et continue.

### Transition d'effondrement

Avant d'énoncer un résultat sur ce modèle, proposons une petite étude heuristique. Considérons un polymère de taille  $L \in \mathbb{N}$  très grande. Si  $\beta$  est petit voire nul, on s'attend à observer le polymère s'étendant dans son solvant, réalisant au moins  $cL$  pas horizontaux pour un  $c > 0$  ; car sous ces conditions, une trajectoire typique sous  $\mathbf{P}_{L,\beta}$  n'est pas très différente de celle d'une marche aléatoire partiellement dirigée dans  $\mathbb{Z}^2$ , ayant une probabilité strictement positive de faire un pas vers la droite à chaque instant. En revanche si  $\beta$  devient très grand, le polymère sera incité à réaliser le plus grand nombre d'auto-contacts possibles. On remarque facilement que pour ce faire, il doit se composer de stretches verticaux très longs et de signes alternés : plus précisément, on constate que le nombre d'auto-contacts réalisables par une trajectoire est majoré par  $L$ , et si  $\sqrt{L} \in \mathbb{N}$  cette borne est atteinte (au premier ordre) par la trajectoire  $\ell^* \in \mathcal{L}_{\sqrt{L},L}$  définie par  $\ell_i^* = (-1)^{i-1}(\sqrt{L}-1)$ ,  $1 \leq i \leq \sqrt{L}$ . Si, pour  $\beta$  grand, le polymère privilégie les trajectoires de la même allure que  $\ell^*$ , il aura alors une extension horizontale  $O(\sqrt{L})$  très petite comparée au  $O(L)$  mentionné plus haut pour  $\beta$  petit.

Formulons les conséquences de ces observations sur l'énergie libre. La trajectoire  $\ell^*$  définie plus haut nous donne immédiatement une borne inférieure sur  $Z_{L,\beta}$ ,

$$Z_{L,\beta} \geq \exp \left( \beta \sum_{i=1}^{\sqrt{L}-1} (\sqrt{L}-1) \right) = e^{\beta(\sqrt{L}-1)^2} = e^{\beta L + o(L)},$$

ce qui implique

$$F(\beta) \geq \beta, \quad \forall \beta \geq 0. \quad (0.1.6)$$

Pour  $\beta$  grand, on s'attend à observer  $F(\beta) = \beta$ , car les trajectoires contribuant le plus à  $Z_{L,\beta}$  seront similaires à  $\ell^*$  et sont trop peu nombreuses dans  $\Omega_L$  pour porter de l'entropie. En revanche pour  $\beta = 0$ , on remarque

$$Z_{L,\beta} = |\Omega_L| \geq 2^{L-1},$$

car  $\Omega_L$  contient les marches dans  $\mathbb{Z}^2$  ne faisant que des pas vers le haut et la droite (et terminant par un pas vers la droite); et on en déduit  $F(0) \geq \log 2 > 0$ . Par conséquent, on pose

$$\beta_c := \sup\{\beta \geq 0; F(\beta) > \beta\}, \quad (0.1.7)$$

le point critique où  $F$  perd sa régularité, et on anticipe  $0 < \beta_c < +\infty$ , et donc que le modèle est bien sujet à une transition de phase en  $\beta_c$ , entre un régime étendu (*Extended*)  $\mathcal{E}$  où le polymère s'étend dans le solvant, et un régime effondré (*Collapsed*)  $\mathcal{C}$  où le polymère se replie sur lui-même.

Cette transition de phase, dite *d'effondrement* (*Collapse*) a été démontrée dans [130, Theorem 1.3] (voir aussi [52, Theorem A]), notamment  $\beta_c$  peut être calculé explicitement, et l'extension horizontale du polymère dans chaque phase a été calculée dans [53, Theorem 2.2] (voir aussi [54] pour une vue d'ensemble du modèle). Nous synthétisons ces résultats dans le théorème suivant.

**Théorème 0.1.1** (IPDSAW, transition d'effondrement). *(i)  $\beta_c$  est l'unique solution en  $\beta > 0$  de l'équation  $e^{-\beta \frac{1+e^{-\beta/2}}{1-e^{-\beta/2}}} = 1$ .*

*(ii) Pour tout  $\beta < \beta_c$  il existe  $a_\varepsilon(\beta) \in (0, 1)$  tel que, pour tout  $\eta > 0$ ,*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L,\beta} \left( \left| \frac{N_L(\ell)}{L} - a_\varepsilon(\beta) \right| > \eta \right) = 0. \quad (0.1.8)$$

*(iii) Pour tout  $\beta > \beta_c$  il existe  $a_c(\beta) > 0$  tel que, pour tout  $\eta > 0$ ,*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L,\beta} \left( \left| \frac{N_L(\ell)}{\sqrt{L}} - a_c(\beta) \right| > \eta \right) = 0. \quad (0.1.9)$$

Le cas critique  $\beta = \beta_c$  est également traité dans [53, Theorem 2.2] (dans ce cas l'extension horizontale est d'ordre  $L^{2/3}$ ), mais nous ne nous attarderons pas dessus dans cette thèse. Avec ces résultats, il est clair qu'on peut définir

$$\begin{aligned} \mathcal{E} &:= \{\beta; F(\beta) > \beta\} = \{\beta; \beta < \beta_c\}, \\ \text{et } \mathcal{C} &:= \{\beta; F(\beta) = \beta\} = \{\beta; \beta \geq \beta_c\}. \end{aligned} \quad (0.1.10)$$

De plus,  $\beta_c$  est bien le point critique d'une transition de phase (car l'énergie libre change d'expression analytique en  $\beta_c$ ), et on peut distinguer la phase  $\mathcal{E}$ , dite *Extended*, où l'extension horizontale du polymère est d'ordre  $L$ , et la phase  $\mathcal{C}$ , dite *Collapsed*, où son extension horizontale est d'ordre  $\sqrt{L}$  (voir Fig. 0.4). Dans la seconde, le polymère se replie sur lui-même pour former un *globule* macroscopique, plein, et de hauteur et largeur comparables.

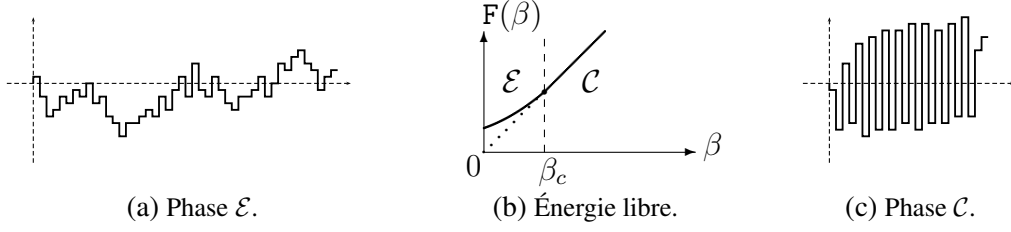


FIGURE 0.4: Diagramme de phase du modèle IPDSAW, et représentation d'une trajectoire typique dans chacune des deux phases  $\mathcal{E}$  et  $\mathcal{C}$ . Dans  $\mathcal{E}$ , le polymère se déploie horizontalement sur une extension d'ordre  $L$ , et réalise beaucoup de petits stretches verticaux. Dans  $\mathcal{C}$ , le polymère se replie sur lui-même, réalisant une extension horizontale d'ordre  $\sqrt{L}$ , et une succession de stretches verticaux de signes alternés et de longueur d'ordre  $\sqrt{L}$ .

En ce qui concerne la valeur de  $\beta_c$ , on peut déduire de Théorème 0.1.1-(i) que  $X = e^{\beta_c/2}$  est l'unique solution dans  $(1, +\infty)$  de l'équation de degré trois  $X^3 - X^2 - X - 1 = 0$ . On peut donc l'exprimer explicitement et même la calculer numériquement, ce qui nous donne  $\beta_c \approx 1.219$ .

En somme, le modèle IPDSAW, tout comme le modèle PDSAW en interaction avec un mur introduit précédemment, est sujet à un phénomène de transition de phase, lequel change drastiquement les propriétés des trajectoires typiques du polymère en fonction de l'énergie d'interaction  $\beta$ . Dans la suite de cette partie, nous nous intéresserons à un modèle associant les deux : une trajectoire IPDSAW située près d'un mur avec lequel elle interagit.

## 0.1.2 Le modèle joint : IPDSAW près d'un mur

Le modèle que nous allons présenter ici a déjà été étudié dans la littérature physique (voir [1, 75, 76], et plus récemment [123, 137, 133]), mais n'a, à notre connaissance, pas encore été abordé dans la littérature mathématique. Son étude est l'objet du Chapitre 1.

### Définition de l'IPDSAW près d'un mur

Nous nous intéressons donc aux trajectoires auto-évitantes et partiellement dirigées dans  $\mathbb{Z}^2$ , contraintes à rester au-dessus d'un mur, interagissant répulsivement avec le solvant (ou, de manière équivalente, en auto-interaction) et attractivement à chaque contact avec le mur. Comme pour le modèle IPDSAW, nous pouvons décrire ces trajectoires par une succession de stretches verticaux séparés par des pas vers la droite. On pose donc

$$\Omega_L^+ := \bigcup_{N \geq 1} \mathcal{L}_{N,L}^+ := \bigcup_{N \geq 1} \left\{ (\ell_i)_{i=1}^N \in \mathbb{Z}^N ; \sum_{i=1}^N |\ell_i| = L - N, \sum_{i=1}^k \ell_i \geq 0, \forall k \leq N \right\}, \quad (0.1.11)$$

où la dernière contrainte  $\sum_{i=1}^k \ell_i \geq 0$  pour tout  $1 \leq k \leq N_L(\ell)$  est celle imposée par le mur, avec  $N_L(\ell)$  l'extension horizontale de  $\ell \in \Omega_L^+$ ; notons que comme dans (0.1.2) on impose à la trajectoire de finir par un pas horizontal (voir Fig. 0.5).

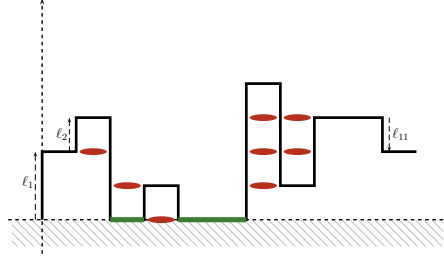


FIGURE 0.5: Représentation d'une trajectoire IPDSAW à proximité d'un mur dur. Lorsque deux monomères non-successifs sont adjacents dans  $\mathbb{Z}^2$ , ils réalisent un auto-contact (représenté par une ellipse rouge). Lorsqu'un segment du polymère longe le mur, il réalise un contact avec celui-ci (représenté par un segment vert).

Soit  $\beta \geq 0$  l'énergie d'interaction de chaque auto-contact, et  $\delta \geq 0$  celle de chaque contact avec le mur. L'énergie totale  $H_L^{\beta, \delta}(\ell)$  récupérée par une trajectoire  $\ell \in \Omega_L^+$  est donc la somme des énergies des deux types d'interactions, soit

$$H_L^{\beta, \delta}(\ell) := \beta \sum_{i=1}^{N_L(\ell)-1} (\ell_i \tilde{\wedge} \ell_{i+1}) + \delta \sum_{k=1}^{N_L(\ell)} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}}, \quad (0.1.12)$$

où on rappelle la définition  $x \tilde{\wedge} y := \min(|x|, |y|) \mathbf{1}_{\{xy \leq 0\}}$ ,  $x, y \in \mathbb{Z}$ . Alors on écrit naturellement la mesure polymère et la fonction de partition associées à ce modèle,

$$\mathbf{P}_{L, \beta, \delta}^+(\ell) := \frac{e^{H_L^{\beta, \delta}(\ell)}}{Z_{L, \beta, \delta}^+}, \quad \ell \in \Omega_L^+, \quad \text{où } Z_{L, \beta, \delta}^+ := \sum_{\ell \in \Omega_L^+} e^{H_L^{\beta, \delta}(\ell)}. \quad (0.1.13)$$

Comme précédemment, l'énergie libre est le taux de croissance exponentielle de  $Z_{L, \beta, \delta}^+$ , soit

$$\mathbf{F}(\beta, \delta) := \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L, \beta, \delta}^+, \quad (0.1.14)$$

(nous montrerons qu'elle est bien définie dans la preuve du Theorem 1.2.2).

### Motivations et objectifs : la transition de surface

Avant d'énoncer nos résultats sur le modèle IPDSAW près d'un mur, proposons ici quelques observations préliminaires. Ce modèle est une superposition exacte d'un modèle de wetting, le PDSAW interagissant avec un mur, et du modèle IPDSAW (sans mur). Chacun d'entre eux est sujet à une transition de phase, d'*adsorption* pour le PDSAW près d'un mur (représentée en Fig. 0.2), et d'*effondrement* pour l'IPDSAW (Fig. 0.4). La première question naturelle concernant le modèle joint consiste à déterminer si ces deux transitions de phase persistent ou non.

Un a priori raisonnable est que ces deux transitions ont toujours lieu, la première en  $\delta$  et la seconde en  $\beta$  : que le polymère soit effondré en globule (Collapsed) ou étendu (Extended), il peut toujours adhérer au mur (Adsorbed) ou s'en éloigner (Desorbed). Nous conjecturons donc que le *diagramme de phase* (i.e. les régions de  $(\mathbb{R}_+)^2$  où  $F$  est analytique) a l'allure décrite en Fig. 0.6.

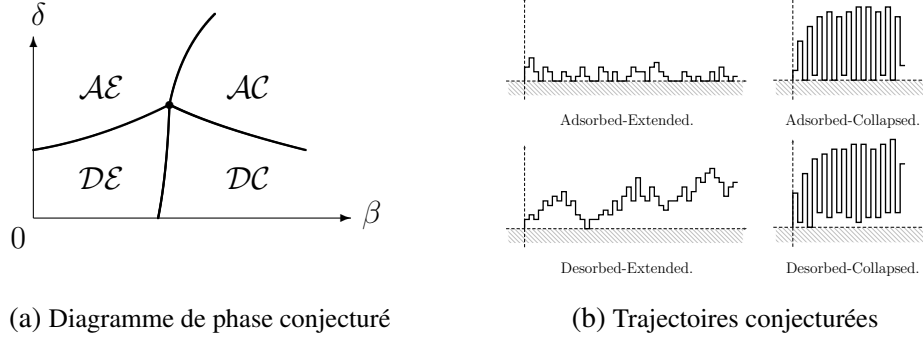


FIGURE 0.6: Première conjecture du diagramme des phases de l'IPDSAW près d'un mur, et des trajectoires typiques dans chaque phase. Ces trajectoires peuvent être Collapsed ou Extended quand  $\beta$  varie, et Adsorbed ou Desorbed quand  $\delta$  varie.

L'énergie libre perdrait son analyticité le long de courbes délimitant quatre régions du quadrant  $(\mathbb{R}_+)^2$  :  $\mathcal{AE}$ ,  $\mathcal{AC}$ ,  $\mathcal{DE}$  et  $\mathcal{DC}$ . Chacune de ces courbes correspond à une transition d'effondrement (entre des régimes  $\mathcal{E}$  et  $\mathcal{C}$ ) ou d'adsorption (entre  $\mathcal{D}$  et  $\mathcal{A}$ ), comme dans les deux modèles d'origine.

Toutefois, on peut observer dans cette conjecture que la transition supposée entre les régimes  $\mathcal{AC}$  et  $\mathcal{DC}$  présente une certaine spécificité par rapport aux autres. Chacune des autres transitions supposées (de  $\mathcal{E}$  à  $\mathcal{C}$ , et de  $\mathcal{AE}$  à  $\mathcal{DE}$ ) traduit un changement de comportement d'une partie macroscopique du polymère : de l'ordre de  $L$  monomères s'accrochent au mur de  $\mathcal{DE}$  à  $\mathcal{AE}$ , ou se replient en globule de  $\mathcal{E}$  à  $\mathcal{C}$ . En revanche, la transition d'accrochage entre  $\mathcal{AC}$  et  $\mathcal{DC}$  concerne essentiellement *l'enveloppe inférieure* du globule, qui nous le rappelons est de longueur d'ordre  $\sqrt{L}$ , donc négligeable devant la taille  $L$  du polymère.

Pour aller plus loin, pour chaque trajectoire  $\ell$  d'extension horizontale  $a_c\sqrt{L}$ ,  $a_c > 0$ , la contribution de l'interaction de contact avec le mur au terme  $e^{H_L^{\beta,\delta}(\ell)}$  de la fonction de partition est bornée par le facteur  $e^{\delta a_c\sqrt{L}}$ . En appliquant la fonction  $\frac{1}{L} \log(\cdot)$  et en prenant la limite  $L \rightarrow \infty$ , on remarque que cette contribution disparaît dans l'énergie libre. Il est donc clair que, si le polymère est dans un régime Collapsed, les interactions de contact entre le polymère et le mur n'apparaissent pas dans l'énergie libre : *il ne peut pas y avoir de transition de phase* (i.e. de rupture de régularité de l'énergie libre  $F$ ) *entre deux régimes  $\mathcal{AC}$  et  $\mathcal{DC}$ .*

Mais cela ne signifie pas qu'une transition de *régimes  $\mathcal{AC}$ - $\mathcal{DC}$*  est impossible : rappelons que l'énergie libre n'est que le premier terme du développement logarithmique de la fonction

de partition, i.e.  $Z_{L,\beta,\delta}^+ = e^{F(\beta,\delta)L+o(L)}$  pour  $L$  grand. Ce que nous venons de constater, est que l'accrochage de l'enveloppe inférieure d'un polymère effondré ne concerne qu'une partie de taille  $\sqrt{L}$  du polymère. Par conséquent, dans la phase effondrée  $\mathcal{C}$ , la contribution des interactions entre le globule et le mur à la fonction de partition ne seront visibles qu'au terme suivant de son développement logarithmique :

$$Z_{L,\beta,\delta}^+ = e^{F(\beta,\delta)L + G(\beta,\delta)\sqrt{L} + o(\sqrt{L})}, \quad \text{pour } L \rightarrow \infty, \quad (0.1.15)$$

avec  $G(\beta, \delta)$  un coefficient inconnu. En somme, si nous souhaitons prouver l'existence d'une transition d'accrochage au sein de la phase Collapsed  $\mathcal{C}$ , entre deux sous-régimes  $\mathcal{AC}$  et  $\mathcal{DC}$ , nous devons prouver un résultat de la forme (0.1.15) et étudier la fonction  $G$ , en particulier ses ruptures d'analyticité.

Dans la littérature physique, ce type de phénomène —c'est-à-dire une transition de régime n'affectant qu'une petite partie du système comme sa surface, et donc invisible dans l'énergie libre— a déjà été observé et étudié dans plusieurs systèmes, et porte dans ce cas le nom de *transition de surface* : en particulier, cela a été conjecturé pour des modèles de polymères près d'un mur dans [123, 137]. Notre objectif est donc de démontrer que l'IPDSAW près d'un mur est sujet à ce phénomène : plus précisément que la transition d'adsorption présentée dans le modèle PDSAW avec mur est toujours présente, même lorsque l'interaction avec le solvant affecte drastiquement la géométrie du polymère en le repliant en globule, mais que sous ces contraintes elle prend effet sur une partie plus petite du système, i.e. la surface du globule, et n'est donc plus une *transition de phase* à proprement parler mais de *surface*.

### 0.1.3 Résultats sur l'IPDSAW près d'un mur

Nous énonçons dans cette section nos résultats principaux sur le modèle IPDSAW près d'un mur, et sur une version contrainte que nous appellerons *modèle à une perle*.

#### Transition d'effondrement

Notre premier résultat concerne l'existence de la transition entre deux régimes respectivement étendu et effondré, similairement au modèle sans mur. En remarquant que la trajectoire  $\ell^*$  définie précédemment est dans  $\Omega_L^+$ , on en déduit rapidement  $F(\beta, \delta) \geq \beta$  comme en (0.1.6); et en s'inspirant du Théorème 0.1.1 pour le modèle sans mur, on pose

$$\begin{aligned} \mathcal{E} &:= \{(\beta, \delta) \in (\mathbb{R}_+)^2; F(\beta, \delta) > \beta\}, \\ \text{et } \mathcal{C} &:= \{(\beta, \delta) \in (\mathbb{R}_+)^2; F(\beta, \delta) = \beta\}, \end{aligned} \quad (0.1.16)$$

(on se permet d'utiliser les mêmes notations que pour le modèle sans mur). Rappelons aussi du Théorème 0.1.1 que  $\beta_c$  est défini comme l'unique solution positive de l'équation  $e^{-\beta} \frac{1+e^{-\beta/2}}{1-e^{-\beta/2}} = 1$ . Alors nous démontrons le théorème suivant.

**Théorème 0.1.2** (IPDSAW avec mur, transition d'effondrement). *On a*

$$\mathcal{C} = \left\{ (\beta, \delta) \in (\mathbb{R}_+)^2; \beta \geq \beta_c, \delta \leq \delta_c(\beta) \right\},$$

où pour tout  $\beta \geq \beta_c$ ,

$$\delta_c(\beta) := \log \left( \frac{\sinh(\beta) + \sqrt{\sinh(\beta)^2 + 1 - e^\beta}}{1 - e^{-\beta}} \right). \quad (0.1.17)$$

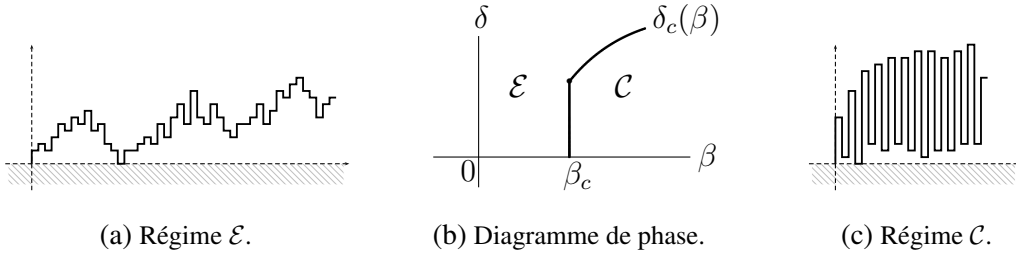


FIGURE 0.7: Diagramme de phase du modèle IPDSAW près d'un mur, avec une transition entre deux régimes *Extended* et *Collapsed*; et représentation des trajectoires typiques attendues dans chaque régime.

Ce résultat nous permet de caractériser complètement les régimes *Extended*  $\mathcal{E}$  et *Collapsed*  $\mathcal{C}$  du modèle avec mur, du moins en les ayant définies à partir de l'énergie libre; en ce qui concerne les estimations géométriques (i.e. l'ordre de l'extension horizontale), on s'attend bien sûr à obtenir des résultats similaires au Théorème 0.1.1-(ii)-(iii) (voir Fig. 0.7). L'extension horizontale du modèle à une perle dans sa phase effondrée sera calculée précisément plus bas, mais nous n'avons pas encore étendu nos résultats au modèle complet; quant au régime étendu, il demande une étude à part que nous avons commencée, mais nous n'avons pas encore obtenu de résultat notable à ce jour. À présent, nous nous intéressons à l'existence d'une transition d'adsorption au niveau du mur dans la phase effondrée  $\mathcal{C}$ .

Remarquons que  $\mathcal{C}$  est bien une *phase* proprement définie, au sens où l'énergie libre ne change pas d'expression analytique en son sein ( $F(\beta, \delta) = \beta$  partout dans  $\mathcal{C}$ ). Cela confirme nos observations et conjectures précédentes, c'est-à-dire qu'une transition de régime *AC-DC*, si elle existe effectivement, n'est pas une transition de phase à proprement parler.

### Modèle à une perle et transition de surface

Pour démontrer le phénomène de transition de surface conjecturé, nous avons vu en (0.1.15) qu'il nous faut obtenir des estimées précises sur la fonction de répartition  $Z_{L,\beta,\delta}^+$ , et en particulier sur son développement logarithmique pour  $L$  grand. Toutefois le modèle IPDSAW avec mur s'avère très complexe mathématiquement, et des estimées précises sont difficiles à obtenir directement. C'est pourquoi nous introduisons une version simplifiée appelée *modèle à une perle* (ou *single-bead model*).





résultats) par :

$$\Omega_L^{\circ,+} := \bigcup_{N \geq 1} \mathcal{L}_{N,L}^{\circ,+} := \bigcup_{N \geq 1} \left\{ \begin{array}{l} (\ell_i)_{i=1}^{2N} \in \mathbb{Z}^{2N} ; \sum_{i=1}^{2N} |\ell_i| = L - 2N , \sum_{i=1}^{2N} \ell_i = 0 , \\ \sum_{i=1}^k \ell_i \geq 0, \forall k \leq 2N , \ell_i \ell_{i+1} < 0, \forall 1 \leq i < 2N \end{array} \right\}, \quad (0.1.18)$$

où on a remarqué que sous ces hypothèses, l'extension horizontale d'une trajectoire est nécessairement paire. Alors la fonction de partition du modèle à une perle se définit par

$$Z_{L,\beta,\delta}^{\circ,+} := \sum_{\ell \in \Omega_L^{\circ,+}} e^{H_L^{\beta,\delta}(\ell)}, \quad (0.1.19)$$

où on rappelle la définition de  $H_L^{\beta,\delta}(\cdot)$  en (0.1.12). De même, son énergie libre se définit par

$$F^\circ(\beta, \delta) := \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\beta,\delta}^{\circ,+}, \quad (0.1.20)$$

où nous prouverons que cette limite existe dans la Proposition 1.3.1. Encore une fois, on remarque que  $\ell^* \in \Omega_L^{\circ,+}$  si  $\sqrt{L} \in 2\mathbb{N}$ , donc  $F^\circ(\beta, \delta) \geq \beta$  comme en (0.1.6).

Nous caractérisons d'abord la phase effondrée de ce modèle à une perle comme pour le modèle complet (on rappelle Théorème 0.1.2, Fig. 0.7), même si la courbe la délimitant est légèrement différente.

**Proposition 0.1.3** (Une perle, transition d'effondrement). *On a*

$$\mathcal{C}_{\text{bead}} := \{(\beta, \delta); F^\circ(\beta, \delta) = \beta\} = \{(\beta, \delta); \beta \geq \beta_c, \delta \leq \delta_c^\circ(\beta)\},$$

où pour tout  $\beta \geq \beta_c$ ,  $\delta_c^\circ(\beta)$  est connu explicitement.

Ensuite, nous séparons la phase  $\mathcal{C}_{\text{bead}}$  en deux sous-régimes le long de la courbe

$$\tilde{\delta}_c : \begin{array}{ll} [\beta_c, +\infty) & \longrightarrow \mathbb{R}_+ \\ \beta & \longmapsto -\log(1 - e^{-\beta/2}) \end{array}, \quad (0.1.21)$$

en posant

$$\begin{aligned} \mathcal{DC} &:= \{(\beta, \delta); \beta \geq \beta_c, \delta \leq \tilde{\delta}_c(\beta)\}, \\ \mathcal{AC} &:= \{(\beta, \delta); \beta \geq \beta_c, \tilde{\delta}_c(\beta) \leq \delta \leq \delta_c^\circ(\beta)\}. \end{aligned} \quad (0.1.22)$$

Ajoutons qu'on observera  $\tilde{\delta}_c(\beta_c) = \delta_c(\beta_c) = \delta_c^\circ(\beta_c) > 0$ , et  $0 < \tilde{\delta}_c(\beta) < \delta_c(\beta) < \delta_c^\circ(\beta)$  pour tout  $\beta > \beta_c$ ; en particulier la courbe  $\tilde{\delta}_c$  sépare bien  $\mathcal{C}_{\text{bead}}$  (et même  $\mathcal{C}$ ) en deux régions bien définies. Alors nous démontrons le théorème suivant, qui est le premier résultat important de cette thèse. Pour deux suites  $(a_L)_{L \geq 1}$ ,  $(b_L)_{L \geq 1}$ , on note  $a_L \asymp b_L$  s'il existe  $C > 0$  tel que  $\frac{1}{C}a_L \leq b_L \leq Ca_L$  pour tout  $L$  assez grand (voir (5.0.1) dans Appendix : Notation).

**Théorème 0.1.4** (Une perle, transition d'adsorption de surface). *Soit  $\beta > \beta_c$ .*

(i) *Pour tout  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ , on a*

$$Z_{L,\beta,\delta}^{\circ,+} \asymp \frac{1}{L^{3/4}} e^{\beta L + \Phi(\beta,\delta)\sqrt{L}}, \quad \text{quand } L \rightarrow \infty; \quad (0.1.23)$$

(ii) *pour tout  $\delta \in (0, \tilde{\delta}_c(\beta)]$  et  $\varepsilon > 0$ , il existe  $C > 0$ ,  $L_0 \in \mathbb{N}$  tels que pour tout  $L \geq L_0$ ,*

$$e^{\beta L + \Phi(\beta,0)\sqrt{L} + (\Psi(\beta) - \varepsilon)L^{1/6}} \leq Z_{L,\beta,\delta}^{\circ,+} \leq \frac{C}{L^{3/4}} e^{\beta L + \Phi(\beta,0)\sqrt{L}}; \quad (0.1.24)$$

(iii) *pour  $\delta = 0$ , on a*

$$Z_{L,\beta,0}^{\circ,+} = e^{\beta L + \Phi(\beta,0)\sqrt{L} + \Psi(\beta)L^{1/6}(1+o(1))}, \quad \text{quand } L \rightarrow \infty, \quad (0.1.25)$$

où  $\Phi(\beta, \delta)$  et  $\Psi(\beta)$  sont connus explicitement.

Nous n'avons pas écrit les formules explicites de  $\Phi(\beta, \delta)$  et  $\Psi(\beta)$  ici pour alléger la présentation, mais elles sont détaillées dans le Chapitre 1. Ces estimées nous permettent notamment d'obtenir des résultats géométriques sur les trajectoires à une perle dans chacun des régimes  $\mathcal{AC}$  et  $\mathcal{DC}$ , en particulier sur leur nombre de contacts avec le mur.

**Corollaire 0.1.5** (Adsorption de surface, estimées géométriques). *Soit  $\beta > \beta_c$ . Alors la fonction  $\delta \in [0, +\infty) \mapsto \Phi(\beta, \delta)$  est  $\mathcal{C}^1$  sur les intervalles  $[0, \tilde{\delta}_c(\beta))$  et  $(\tilde{\delta}_c(\beta), +\infty)$ , constante sur le premier et strictement croissante sur le second. De plus,*

(i) *pour tout  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$  et  $\varepsilon > 0$ , on a*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( \sum_{k=1}^{N_L(\ell)} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}} \in \left[ \partial_\delta \Phi(\beta, \delta) - \varepsilon, \partial_\delta \Phi(\beta, \delta) + \varepsilon \right] \sqrt{L} \right) = 1; \quad (0.1.26)$$

(ii) *pour tout  $\delta \in [0, \tilde{\delta}_c(\beta))$ , il existe  $K > 0$  (qui dépend de  $\beta$ ) tel que*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( \sum_{k=1}^{N_L(\ell)} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}} \leq KL^{1/6} \right) = 1. \quad (0.1.27)$$

Le Théorème 0.1.4 démontre, dans le modèle à une perle, le phénomène de transition de surface que nous avons conjecturé. En effet pour tout  $\beta > \beta_c$ , le coefficient du terme de surface (i.e. en  $\sqrt{L}$ ) dans le développement logarithmique de  $Z_{L,\beta,\delta}^{\circ,+}$  perd son analyticit  au point  $\tilde{\delta}_c(\beta)$ , ce qui se traduit par des estimations g om etriques tr s diff erentes dans le Corollaire 0.1.5 sur le nombre de contacts entre le polym re et le mur (voir Fig. 0.9).

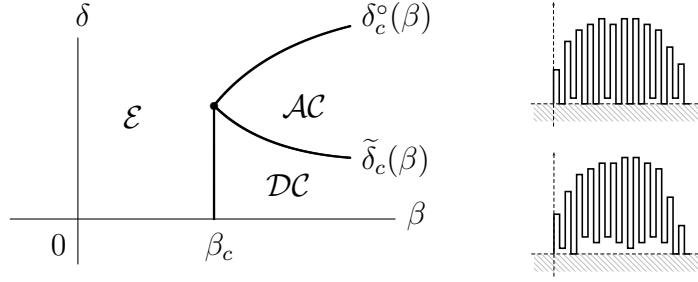


FIGURE 0.9: Diagramme de phase et trajectoires typiques prouvés pour le modèle à une perle, avec une transition de surface entre les sous-régimes  $\mathcal{AC}$  (en haut à droite) et  $\mathcal{DC}$  (en bas à droite) le long de la courbe  $\tilde{\delta}_c(\cdot)$ . Dans le premier l'enveloppe inférieure du globule s'accroche au mur, alors qu'elle est décrochée dans le second. Le même diagramme est conjecturé pour le modèle complet.

### Structure de la preuve : représentation probabiliste et apparition d'un *pre-wetting*

Nous présentons ici la structure de la preuve du Théorème 0.1.4 et ses idées principales. Celle-ci s'inspire d'abord de [130], où les auteurs réécrivent les trajectoires du modèle IPDSAW sous la mesure polymère à l'aide d'une marche aléatoire auxiliaire. Nous adaptons cette méthode ici pour décrire les *enveloppes* du globule.

Pour une trajectoire donnée  $\ell \in \mathcal{L}_{L,N}^{\circ,+}$  d'extension horizontale  $2N$ , nous définissons ses enveloppes supérieure  $S = (S_i)_{i=0}^{N+1}$  et inférieure  $I = (I_i)_{i=0}^N$  par

$$\begin{aligned} S_k &= \sum_{i=1}^{2k-1} \ell_i, \quad 1 \leq k \leq N+1, \\ I_k &= \sum_{i=1}^{2k} \ell_i, \quad 1 \leq k \leq N, \end{aligned} \tag{0.1.28}$$

et  $S_0 = I_0 = 0$ , où on pose  $\ell_0 = \ell_{2N+1} = 0$ . Comme les stretches de  $\ell$  sont contraints à alterner leur signe, on remarque  $\ell_{2i} > 0$  et  $\ell_{2i-1} < 0$  pour tout  $1 \leq i \leq N$ , et donc  $S$  reste au-dessus de  $I$  : plus précisément  $S_i > \max(I_{i-1}, I_i)$ ,  $1 \leq i \leq N$ . De plus, la longueur  $L$  du polymère contraint les deux marches  $S$  et  $I$  à entourer une région d'aire fixée (donnée par le nombre de pas verticaux dans la trajectoire, soit  $L - 2N$ ) entre elles, voir Fig. 0.10.

Avec un peu de réécriture, nous reformulerons  $Z_{L,\beta,\delta}^{\circ,+}$ , pour un  $N$  fixé, comme une espérance sur deux marches aléatoires  $S$  et  $I$  indépendantes et de même loi  $\mathbf{P}_\beta$  (qui dépend de  $\beta > 0$ ), contraintes à rester l'une au-dessus de l'autre, au-dessus du mur, à entourer une aire fixée  $L - 2N$ , et dont celle du bas peut s'accrocher au mur (gagnant une énergie  $\delta$  à chaque contact).

La deuxième étape consiste à découpler les deux marches, en fixant (i.e. en conditionnant) par l'enveloppe inférieure et en intégrant sur l'enveloppe supérieure. Pour  $N \approx \sqrt{L}$ , elle est contrainte à envelopper une aire anormalement grande  $L - 2N + A_N(I) \approx N^2 + A_N(I)$ , où  $A_N(I) = \sum_{i=1}^N I_i$  est l'aire arithmétique de l'enveloppe inférieure. Avec des

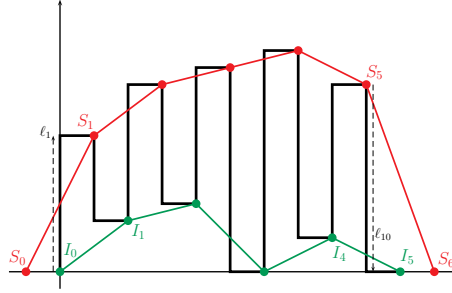


FIGURE 0.10: Représentation des enveloppes supérieure  $S$  et inférieure  $I$  d'une trajectoire à une perle. Les signes alternés des stretches imposent à  $S$  de rester au dessus de  $I$ . La longueur  $L$  du polymère contraint l'aire géométrique située entre les deux marches.

estimées précises dans ce régime de grande déviations, nous montrons que sa contribution à la fonction de partition est d'ordre  $\frac{1}{N^2} e^{-cN - \gamma \frac{1}{N} A_N(I)}$ .

En réinjectant ce résultat dans la réécriture de  $Z_{L,\beta,\delta}^{\circ,+}$ , on observe d'une part un premier facteur à l'ordre de la surface  $e^{-cN}$  (qui donnera en partie la fonction  $\Phi(\cdot, \cdot)$ ), et d'autre part qu'il nous reste à étudier la fonction de partition d'une marche aléatoire affectée à la fois par un wetting  $\exp(\delta \sum_{i=1}^N 1_{\{I_i=0\}})$ , et par un terme de *pre-wetting*  $\exp(-\gamma \frac{1}{N} A_N(I))$  (notons que  $\frac{1}{N} A_N(I)$  est la hauteur moyenne de la marche  $I$ ).

L'influence d'un terme de pre-wetting sur une marche aléatoire en dimension  $1 + 1$  a déjà été étudiée dans [102] et [106] (voir aussi [107] pour une vue d'ensemble). L'étude de ces problèmes est notamment motivée par la littérature physique, par exemple pour une l'interface liquide-gaz quand un gaz stable est en contact avec un substrat (mur dur), alors que ce dernier préfère interagir avec le liquide, et que la température décroît vers le point critique liquide-gaz. Notre étude dévoile un autre système où ces phénomènes de pre-wetting apparaissent naturellement, avec une enveloppe inférieure  $I$  qui est contrainte à rester au-dessus du mur, mais la présence du globule et la tension de l'enveloppe supérieure (qui est lourdement pénalisée si elle doit monter très haut) exercent une "pression" sur l'enveloppe inférieure, sous la forme du pre-wetting.

La suite de la preuve est l'étude du modèle avec wetting et pre-wetting : nous résumons nos observations principales ici sans entrer dans les détails. On note  $\tilde{\delta}_c = \tilde{\delta}_c(\beta)$  le point critique de ce modèle de wetting sans pre-wetting (lequel est très bien connu, voir encore [84, Chapters 1-2]). On observe d'une part dans le régime  $\delta > \tilde{\delta}_c$  que la marche sans pre-wetting est déjà accrochée au mur, donc l'ajout de la pénalisation  $-\frac{1}{N} A_N(I)$  ne change pas significativement les trajectoires ; alors sa fonction de partition est  $\approx e^{g(\delta)N}$  avec  $g(\delta) > 0$  l'énergie libre du modèle de wetting, ce qui apporte un second facteur à l'ordre de la surface dans le calcul de  $Z_{L,\beta,\delta}^{\circ,+}$ .

D'autre part dans le régime  $\delta < \tilde{\delta}_c$ , le modèle sans pre-wetting est décroché du mur, et la présence de pre-wetting n'est pas suffisante pour l'accrocher ; mais elle le contraint tout de même à ne monter qu'à une hauteur  $N^{1/3}$  (contrairement aux déviations standard

$N^{1/2}$ ), et sa fonction de partition devient de l'ordre de  $e^{hN^{1/3}(1+o(1))}$ . Ceci est à l'origine de l'apparition du terme  $\Psi(\beta)L^{1/6} \approx \Psi(\beta)N^{1/3}$  dans le calcul de  $Z_{L,\beta,\delta}^{\circ,+}$ , et implique également qu'il n'y a pas de second facteur à l'ordre de la surface dans  $Z_{L,\beta,\delta}^{\circ,+}$  pour  $\delta < \tilde{\delta}_c$  contrairement à  $\delta > \tilde{\delta}_c$  : d'où la rupture d'analyticité de la fonction  $\Phi(\beta, \delta)$  le long de la courbe  $\tilde{\delta}_c(\cdot)$ . En réalité, nous n'obtenons d'estimation précise que lorsque  $\delta = 0$  (ce qui nous suffit), mais nous conjecturons qu'elle reste valable pour  $\delta \in (0, \tilde{\delta}_c(\beta))$ .

Enfin, la preuve se termine en prouvant que les seules extensions horizontales contribuant significativement à  $Z_{L,\beta,\delta}^{\circ,+}$  sont les  $N \in [a_1, a_2]\sqrt{L}$ ,  $a_1 < a_2 \in \mathbb{R}$ ; puis en optimisant sur la valeur de  $\frac{N}{\sqrt{L}}$ , ce qui donne les expressions de  $\Phi$  et  $\Psi$  comme les solutions d'un problème variationnel.

### 0.1.4 Améliorations pour l'IPDSAW sans mur

La précision de nos estimées sur le modèle avec mur —en particulier nos théorèmes limite locaux dans un régime de grande déviations— nous a également permis d'améliorer les résultats antérieurs concernant le modèle sans mur dans sa phase effondrée. En effet, nous donnons ici un équivalent exact de la fonction de partition de l'IPDSAW sans mur. On rappelle que  $Z_{L,\beta}$  et  $\beta_c$  sont définis en (0.1.4) et (0.1.7).

**Théorème 0.1.6** (IPDSAW sans mur, équivalent de  $Z_{L,\beta}$ ). *Pour tout  $\beta > \beta_c$ , on a*

$$Z_{L,\beta} \underset{L \rightarrow \infty}{\sim} \frac{K_\beta}{L^{3/4}} e^{\beta L + \Upsilon(\beta)\sqrt{L}}, \quad (0.1.29)$$

où  $K_\beta$  et  $\Upsilon(\beta)$  sont connus explicitement.

Les expressions explicites de  $K_\beta$  et  $\Upsilon(\beta)$  sont présentées dans le Chapitre 2. Ce résultat se base sur la décomposition d'une trajectoire en perles similairement à la Figure 0.8, et sur le calcul d'un équivalent exact de la fonction de partition du modèle à une perle (améliorant [52, Proposition 4.2] qui détermine le développement logarithmique mais pas le pré-facteur polynomial). Une conséquence directe de cette estimée est le constat que, à l'exception d'un nombre fini d'entre eux, tous les monomères du modèle IPDSAW dans sa phase effondrée sont contenus dans une seule grande perle macroscopique. Plus précisément, la taille de la plus grande perle d'une trajectoire  $\ell \in \Omega_L$  se définit par

$$|I_{\max}(\ell)| := \max \left\{ \sum_{i=u}^v (1 + |\ell_i|); 1 \leq u \leq v \leq N_L(\ell), \ell_i \ell_{i+1} < 0 \ \forall u \leq i \leq v-1 \right\}, \quad (0.1.30)$$

et on obtient le résultat suivant.

**Théorème 0.1.7** (IPDSAW sans mur, perle macroscopique). *Pour tout  $\beta > \beta_c$ ,*

$$\lim_{k \rightarrow \infty} \liminf_{L \rightarrow \infty} \mathbf{P}_{L,\beta}(|I_{\max}(\ell)| \geq L - k) = 1. \quad (0.1.31)$$

C'est une amélioration de [52, Theorem C], où un nombre  $O((\log L)^4)$  de monomères pouvaient être en dehors de la perle. Cela confirme l'observation qu'une marche IPDSAW effondrée (sans mur) n'est composée que d'une seule perle macroscopique. En outre, si ce résultat est aussi valable pour le modèle avec mur (ce qui semble être une conjecture raisonnable), cela indiquerait que sa fonction de partition est très proche du modèle à une seule perle (ne diffère que d'un facteur  $O(1)$ ); nos estimées du Théorème 0.1.4 devraient donc rester vraies pour le modèle avec mur complet.

### 0.1.5 Synthèse et questions ouvertes

L'objectif que nous nous étions initialement fixé était d'observer une perturbation de la transition d'adsorption du modèle PDSAW près d'un mur. Par l'introduction dans le modèle d'une interaction d'auto-contacts (qui, rappelons-le, représente la mauvaise qualité d'un solvant), nous avons constaté que cette dernière peut affecter considérablement la géométrie du polymère en l'incitant à former un globule compact; mais que même sous ces contraintes la transition d'adsorption est toujours présente, en revanche elle ne concerne plus qu'une petite partie du modèle, la surface du globule. En particulier, notre résultat le plus intéressant dans cette partie —et également le plus difficile— est le Théorème 0.1.4, qui exhibe dans un cadre mathématique un phénomène de transition de surface sur un modèle de polymère, comme précédemment énoncé dans la littérature physique [123, 137] (on rappelle Fig. 0.9).

Cependant, de nombreuses questions restent encore ouvertes sur ce phénomène, et plus largement sur le modèle IPDSAW près d'un mur. En premier lieu, rappelons que nous n'avons prouvé la transition de surface que pour le modèle à une perle, et plus précisément une perle commençant et finissant près du mur. Nous nous attendons à observer le même phénomène dans le modèle complet, en particulier avec le même développement logarithmique et la même courbe critique  $\beta \mapsto \tilde{\delta}_c(\beta)$  que le modèle à une perle; d'où la conjecture suivante.

**Conjecture 0.1.8.** *Soit  $\beta > \beta_c$ . Alors pour tout  $\delta \in (\tilde{\delta}_c(\beta), \delta_c(\beta))$ , on a*

$$Z_{L,\beta,\delta}^+ \asymp \frac{1}{L^{3/4}} e^{\beta L + \Phi(\beta,\delta)\sqrt{L}}, \quad \text{quand } L \rightarrow \infty; \quad (0.1.32)$$

*et pour tout  $\delta \in [0, \tilde{\delta}_c(\beta))$ , on a*

$$Z_{L,\beta,\delta}^+ = e^{\beta L + \Phi(\beta,0)\sqrt{L} + \Psi(\beta)L^{1/6} + O(\log L)}, \quad \text{quand } L \rightarrow \infty, \quad (0.1.33)$$

*où  $\Phi$  et  $\Psi$  sont donnés par le Théorème 0.1.4.*

(En particulier, notre étude nous permet de conjecturer que le terme  $\Psi(\beta)L^{1/6}$  prouvé en Théorème 0.1.4-(iii) est le dernier terme du développement de  $\log Z$  et reste le même dans tout le régime  $DC$  des deux modèles.) Pour obtenir un tel résultat, nous devons obtenir des estimées précises pour *toutes* les perles (y compris celles éloignées du mur) puis les utiliser dans la décomposition en perle de la trajectoire complète. Cela représente un défi

technique important, mais devrait être accessible en améliorant nos résultats et fait l'objet d'un travail en cours.

Mentionnons également les questions ouvertes suivantes :

- Nos résultats du Théorème 0.1.4-(ii)-(iii) se sont pas pleinement satisfaisants : nous conjecturons que le résultat pour  $\delta = 0$  s'applique à tout  $\delta \in [0, \tilde{\delta}_c(\beta))$ , et qu'il n'y a pas d'autre terme dans le développement logarithmique de la fonction de partition après  $\Psi(\beta)L^{1/6}$  — cela remplacerait  $L^{1/6}$  par  $\log L$  dans le Corollaire 0.1.5-(ii). Pour obtenir ces améliorations, la clef est d'obtenir des estimées plus précises sur le modèle de wetting additionnellement pénalisé par un pre-wetting mentionné plus haut, notamment dans le cas  $0 < \delta < \tilde{\delta}_c$ . En particulier, ces améliorations sont nécessaires pour obtenir des résultats aussi précis que ceux annoncés dans la Conjecture 0.1.8.
- Nous n'avons pas traité le régime étendu du modèle dans cette thèse, la raison principale étant que nos outils, notamment l'introduction des "enveloppes" du polymère, sont assez mal adaptés au régime  $\mathcal{E}$ . Il ne devrait pas être trop difficile de montrer l'existence d'une transition de phase entre deux régions  $\mathcal{AE}$  et  $\mathcal{DE}$ , avec deux expressions analytiques différentes de l'énergie libre ; mais déterminer explicitement la courbe critique associée est une question difficile. Une seconde question qui suit naturellement est le raccordement des courbes critiques : y a-t-il un point quadruple, comme représenté dans notre conjecture Fig. 0.6a, ou bien deux points triples ? La littérature physique semble privilégier la première hypothèse (voir [123, 137]) ce qui est une conjecture très raisonnable, mais cela reste à prouver dans notre cadre.
- Dans [52, Theorem B], il est prouvé que la transition d'effondrement du modèle IPDSAW sans mur est d'ordre  $3/2$ , c'est-à-dire  $F(\beta - \varepsilon) \approx \beta + c\varepsilon^{3/2}$ ,  $\varepsilon \searrow 0$ ). Qu'en est-il du modèle avec mur ? On peut conjecturer qu'entre les régimes  $\mathcal{DE}$  et  $\mathcal{DC}$ , on retrouve le même exposant critique  $3/2$  ; mais en ce qui concerne la transition de  $\mathcal{AE}$  à  $\mathcal{AC}$ , c'est-à-dire entre deux régimes énergétiques très différents, cela n'est pas clair. Notons aussi que s'il s'avère que le diagramme de phase a deux points triples plutôt qu'un point quadruple, alors il existerait une autre transition d'effondrement  $\mathcal{AE}-\mathcal{DC}$  ou  $\mathcal{DE}-\mathcal{AC}$ , laquelle demanderait une étude spécifique.
- Une dernière question ouverte est celle des régimes critiques. D'une part, y a-t-il une transition d'adsorption "de surface" le long de la courbe  $(\beta, \delta) \in \overline{\mathcal{C}} \cap \overline{\mathcal{E}}$  (a priori à l'échelle  $L^{2/3}$ , voir [53, Theorem 2.2] ; et remarquons que cette question peut être reliée à la précédente) ? D'autre part, à quoi ressemblent la fonction de partition et les trajectoires typiques dans le cas  $(\beta, \delta) \in \overline{\mathcal{AC}} \cap \overline{\mathcal{DC}}$  ?

## 0.2 Le modèle généralisé de Poland-Scheraga pour la dénaturation de l'ADN

À présent nous allons nous intéresser à une toute autre approche, en considérant des possibles inhomogénéités dans la composition du polymère, ou la présence d'impuretés sur



l'objet auquel il s'adsorbe, ce qui se traduit par une perturbation des énergies d'interactions entre le polymère et l'objet fixe. En particulier, ces perturbations peuvent être formulées indépendamment des détails fins du modèle choisi, puisqu'elles ne dépendent pas du comportement du polymère loin de l'objet ; par conséquent nous pouvons nous placer ici dans le formalisme très général des modèles de pinning comme introduit plus haut. Dans cette thèse nous nous concentrerons plus précisément sur une version un peu étendue du modèle de pinning, que nous appellerons *modèle de Poland-Scheraga généralisé* ; mais pour exposer clairement nos motivations et objectifs, nous présentons d'abord les résultats déjà connus sur le premier. Avertissons également le lecteur que nous adaptons légèrement nos notations dans cette partie, afin qu'elles correspondent à celles déjà existantes dans la littérature.

### 0.2.1 Les modèles de pinning désordonnés

Rappelons qu'un processus de renouvellement  $\tau = (\tau_i)_{i \geq 1}$  est une suite de variables aléatoires telle que ses incréments  $(\tau_i - \tau_{i-1})$ ,  $i \geq 1$ , sont i.i.d. à valeur dans  $\mathbb{N} \cup \{+\infty\}$ , avec la convention  $\tau_0 = 0$ . Nous faisons ici l'hypothèse très commune que la queue de  $\tau_1$  est donnée par une fonction à variations régulières, plus précisément il existe  $\alpha \geq 0$  et  $L$  une fonction à variations lentes (en  $+\infty$ ) tels que

$$\mathbf{P}(\tau_1 = n) = \frac{L(n)}{n^{\alpha+1}}, \quad n \in \mathbb{N}. \quad (0.2.1)$$

Si le processus  $\tau$  est construit à partir des temps de retours d'une marche aléatoire, cette hypothèse sera très généralement vérifiée (par exemple  $\alpha = 1/2$  pour les retours en 0 de la marche aléatoire simple dans  $\mathbb{Z}$ ). Dans la suite nous supposons que  $\tau_1 < +\infty$   $\mathbf{P}$ -p.s. par commodité pour les formules, mais retirer cette hypothèse ne fait que traduire les résultats ultérieurs d'une constante (voir par exemple [84] pour les détails) donc nous ne perdons pas en généralité.

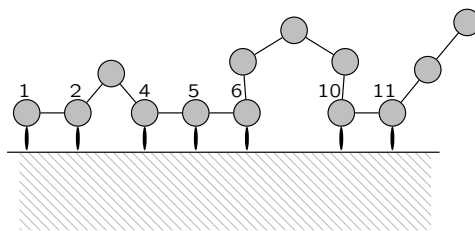


FIGURE 0.11: Représentation d'un polymère s'accrochant à un objet. La suite des indices des monomères interagissant avec le mur est donnée par le processus de renouvellement  $\tau = \{1, 2, 4, \dots\}$ .

## Introduction des inhomogénéités, ou “désordre”

Intéressons-nous à présent aux interactions entre le polymère et l'objet, en considérant qu'elles peuvent différer d'un monomère à l'autre. Pour chaque monomère s'accrochant à l'objet, c'est-à-dire chaque point de renouvellement  $i \in \tau$  (voir Fig. 0.11), nous lui attribuons une récompense énergétique

$$\beta\omega_i + h, \quad (0.2.2)$$

où  $h \in \mathbb{R}$  est un paramètre *homogène*,  $\omega_i$  est une variable aléatoire réelle et  $\beta \geq 0$  un paramètre d'intensité. La suite  $\omega = (\omega_i)_{i \in \mathbb{N}}$  est appelée *désordre*, et elle peut représenter au choix les inhomogénéités de la composition du polymère (par exemple pour un brin d'ADN, le fait qu'un monomère soit une base A, T, C ou G) ou des impuretés sur l'objet à l'endroit où le monomère s'accroche. Le paramètre  $\beta$  est appelé *intensité du désordre*, et en particulier si  $\beta = 0$  on retrouve un modèle parfaitement homogène, récompensant chaque interaction par la même énergie  $h$ . Dans un premier temps nous supposons la suite  $\omega$  i.i.d., quitte à modifier  $h$  et  $\beta$  on peut aussi la prendre centrée et de variance 1, et nous supposons également  $\mathbb{E}[e^{\beta\omega_1}] < +\infty$  pour certains  $\beta > 0$ . Nous notons  $\mathbb{P}$  la loi du désordre  $\omega$ , pour la différencier de celle du renouvellement  $\tau$ , notée  $\mathbf{P}$ .

À présent nous pouvons construire un modèle de pinning associé aux énergies d'interaction (0.2.2). Soit un polymère de longueur  $n \in \mathbb{N}$ , et considérons une réalisation  $\omega \in \mathbb{R}^{\mathbb{N}}$  du désordre. Alors nous définissons la mesure polymère *quenched* (ou “trempé”) et *contrainte* par

$$\frac{d\mathbf{P}_{n,h}^{\beta,\omega,q}}{d\mathbf{P}}(\tau) := \frac{1}{Z_{n,h}^{\beta,\omega,q}} \exp\left(\sum_{i=1}^n (\beta\omega_i + h) \mathbf{1}_{\{i \in \tau\}}\right) \mathbf{1}_{\{n \in \tau\}}, \quad (0.2.3)$$

avec  $Z_{n,h}^{\beta,\omega,q}$  la fonction de partition *quenched* et *contrainte*,

$$Z_{n,h}^{\beta,\omega,q} := \mathbf{E}\left[\exp\left(\sum_{i=1}^n (\beta\omega_i + h) \mathbf{1}_{\{i \in \tau\}}\right) \mathbf{1}_{\{n \in \tau\}}\right], \quad (0.2.4)$$

où nous avons également rajouté une *contrainte*  $\{n \in \tau\}$  —c'est-à-dire le dernier monomère est accroché— dans les définitions précédentes, et où le terme *quenched* signifie que nous avons préalablement fixé une réalisation de  $\omega$ . Bien sûr nous pouvons définir les mesure polymère et fonction de partition *libres* en retirant la contrainte,

$$\begin{aligned} \frac{d\mathbf{P}_{n,h}^{\beta,\omega,\text{free}}}{d\mathbf{P}}(\tau) &:= \frac{1}{Z_{n,h}^{\beta,\omega,\text{free}}} \exp\left(\sum_{i=1}^n (\beta\omega_i + h) \mathbf{1}_{\{i \in \tau\}}\right), \\ \text{et } Z_{n,h}^{\beta,\omega,\text{free}} &:= \mathbf{E}\left[\exp\left(\sum_{i=1}^n (\beta\omega_i + h) \mathbf{1}_{\{i \in \tau\}}\right)\right]. \end{aligned} \quad (0.2.5)$$

Dans la suite de notre travail, il sera souvent plus commode d'utiliser la version contrainte du modèle, d'où notre choix de notation. Remarquons qu'on peut aussi définir une version

conditionnée de la mesure polymère par  $\mathbf{P}_{n,h}^{\beta,\omega,\text{cond}}(\cdot) := \mathbf{P}_{n,h}^{\beta,\omega,\text{free}}(\cdot | n \in \tau)$ , mais nous ne nous en servons pas ici.

Il est important de remarquer que la mesure  $\mathbf{P}_{n,h}^{\beta,\omega,q}$  dépend fortement de la réalisation  $\omega$ , donc est elle-même aléatoire —même si dans la suite, nous omettrons l'indice  $\omega$  pour alléger la notation. En particulier, définir son énergie libre et vérifier la convergence demande d'être plus vigilant que pour un modèle homogène. Cela a notamment été fait dans [84, Theorem 4.1], résultat que nous réécrivons ici.

**Proposition 0.2.1** (Pinning désordonné, énergie libre). *Soient  $\beta \geq 0$  et  $h \in \mathbb{R}$ . Il existe un nombre  $F(\beta, h) \in \mathbb{R}$  tel que*

$$\begin{aligned} F(\beta, h) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,h}^{\beta,q} = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \log Z_{n,h}^{\beta,q} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,h}^{\beta,\text{free}} = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \log Z_{n,h}^{\beta,\text{free}}, \end{aligned} \quad (0.2.6)$$

où les convergences ci-dessus sont au sens  $\mathbb{P}$ -p.s. et  $L^1(\mathbb{P})$ . De plus la fonction  $(\beta, h) \mapsto F(\beta, h)$  est convexe, positive, et est croissante en chaque coordonnée.

Par conséquent l'énergie libre du modèle (désordonné) est bien définie. De plus, un calcul assez direct de la dérivée partielle  $\partial_h F(\beta, h)$ , là où elle est définie (voir [91, Theorem 2.1]), donne

$$\partial_h F(\beta, h) = \lim_{n \rightarrow \infty} \mathbf{E}_{n,h}^{\beta,q} \left[ \frac{\#(\tau \cap \llbracket 1, n \rrbracket)}{n} \right], \quad (0.2.7)$$

où  $\#(\tau \cap \llbracket 1, n \rrbracket)$  désigne le nombre de renouvellements avant  $n$  —soit le nombre de monomères accrochés. Ce résultat est très similaire à (0.0.7), et montre bien que, selon si la fonction  $h \mapsto F(\beta, h)$  est localement constante ou strictement croissante, la localisation du polymère peut drastiquement changer. Nous allons donc étudier la question de l'existence d'une transition de phase d'accrochage dans ce modèle, aussi bien quand il est désordonné ( $\beta > 0$ ) que homogène ( $\beta = 0$ ). En particulier, le problème qui nous intéresse dans cette thèse est de déterminer dans quelle mesure l'introduction du désordre  $(\beta\omega_i)_{i \in \mathbb{N}}$  affecte cette transition.

Comme nous cherchons à étudier l'existence d'une transition de phase d'accrochage en fonction du paramètre  $h$  pour un  $\beta \geq 0$  fixé, nous définissons

$$h_c^q(\beta) = h_c(\beta) := \sup\{h \in \mathbb{R}; F(\beta, h) = 0\}, \quad (0.2.8)$$

et nous observons grâce à la Proposition 0.2.1 et à (0.2.7) qu'une transition d'accrochage a lieu si  $h_c(\beta) \in \mathbb{R}$ , ce qui implique que le polymère est *localisé* (ou *accroché*) pour  $h > h_c(\beta)$ , et *délocalisé* (ou *décroché*) pour  $h < h_c(\beta)$ .

Comme mentionné plus haut, dans le cas  $\beta = 0$  le modèle est homogène —en particulier la fonction de partition  $Z_{n,h} := Z_{n,h}^{0,q}$  est constante  $\mathbb{P}$ -p.s.. La transition d'accrochage est alors très bien connue, y compris l'équivalent de  $F$  autour du point critique qui est donné par le résultat ci-dessous et démontré dans [84, Theorem 2.1].

**Théorème 0.2.2** (Pinning homogène, comportement critique). *Soit  $\beta = 0$ . Alors  $h_c(0) = 0$ . De plus, on a pour tout  $h > 0$ ,*

$$F(0, h) = h^{1/\min(\alpha, 1)} L_\alpha(1/h), \quad (0.2.9)$$

où  $L_\alpha$  est une fonction à variations lentes (si  $\alpha = 0$ , cela signifie  $F(0, h) = o(h^\gamma)$  quand  $h \searrow 0$  pour tout  $\gamma > 0$ ).

L'expression de la fonction  $L_\alpha$  est détaillée dans [84, Theorem 2.1]. Les éléments qui nous intéresseront particulièrement dans ce théorème sont la valeur du *point critique* homogène  $h_c(0) = 0$ , et l'*exposant critique*  $1/\min(\alpha, 1)$  (aussi appelé *ordre*) de la transition de phase.

Avant de donner des résultats sur l'influence du désordre sur le modèle de pinning, il nous reste une notion à introduire : le modèle *annealed* (ou “recuit”), obtenu en prenant l'espérance de la fonction de partition :

$$\begin{aligned} Z_{n,h}^{\beta,a} &:= \mathbb{E} \left[ Z_{n,h}^{\beta,q} \right] = \mathbb{E} \mathbf{E} \left[ \exp \left( \sum_{i=1}^n (\beta \omega_i + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] \\ &= \mathbf{E} \left[ \exp \left( \sum_{i=1}^n (\lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] = Z_{n, \lambda(\beta) + h}, \end{aligned} \quad (0.2.10)$$

où nous avons utilisé le théorème de Fubini-Tonelli et que les  $(\beta \omega_i + h)_{i \in \mathbb{N}}$  sont i.i.d, en posant  $\lambda(\beta) := \log \mathbb{E}[e^{\beta \omega_1}] < +\infty$ . Dans le cas du pinning, ce modèle *annealed* correspond exactement au modèle homogène translaté de  $\lambda(\beta)$ . En particulier il a une énergie libre

$$F^a(\beta, h) := \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,h}^{\beta,a} = F(0, \lambda(\beta) + h), \quad (0.2.11)$$

et un point critique  $h_c^a(\beta) := \sup\{h \in \mathbb{R}; F^a(\beta, h) = 0\} = -\lambda(\beta)$ . Nous pouvons alors comparer les points critiques des modèles quenched et annealed, d'une part avec l'inégalité de Jensen  $\mathbb{E} \log Z_{n,h}^{\beta,q} \leq \log Z_{n,h}^{\beta,a}$  qui implique  $F(\beta, h) \leq F^a(\beta, h)$ , et d'autre part  $F(\beta, h) \geq F(0, h)$  car  $F$  est croissante en  $\beta$ . Ces comparaisons impliquent

$$-\lambda(\beta) = h_c^a(\beta) \leq h_c(\beta) \leq h_c(0) = 0. \quad (0.2.12)$$

De plus, il est montré dans [84, Theorem 5.2] que l'inégalité de droite est stricte pour  $\beta > 0$  : en effet, même si  $\mathbb{E}[\omega_i] = 0$ , le comportement du polymère est changé par le désordre, et en privilégiant de s'accrocher sur les  $\omega_i, i \in \mathbb{N}$  les plus élevés, il gagne un peu plus d'énergie que  $h$  ; la présence de désordre va systématiquement encourager la localisation. En revanche pour l'inégalité de gauche, le fait qu'elle soit stricte ou non est directement relié à la question de l'influence du désordre sur le modèle, ce que nous discutons ci-dessous.

### La question de la (non)-pertinence du désordre

Rappelons que la problématique que nous nous sommes fixée est d'étudier l'effet de perturbations —ici le désordre— sur la transition d'accrochage d'un modèle de polymère.

Les inégalités (0.2.12) nous permettent d'affirmer que la transition de phase a toujours lieu dans le modèle désordonné, et à présent nous souhaitons la comparer à son analogue homogène, en particulier les valeurs de son point critique (en comparaison au modèle annealed) et de son exposant critique. Il s'agit d'une question classique et généralement difficile de mécanique statistique, qu'on appelle la *pertinence* ou *non-pertinence* du désordre. Plus précisément, si l'introduction d'un désordre de n'importe quelle intensité  $\beta > 0$  affecte les valeurs critiques, voire fait disparaître la transition de phase, le désordre est qualifié de *pertinent* (ou *relevant*). Mais si un désordre de faible intensité n'affecte pas les valeurs critiques, notamment l'exposant, alors il est dit *non-pertinent* (ou *irrelevant*).

Pour le modèle de pinning, cette question a été complètement résolue ces dernières années par la contribution de nombreux auteurs, voir [6, 20, 64, 90, 84, 147] parmi d'autres. En particulier, il a été montré que le modèle de pinning est conforme au *critère de Harris* [98] pour la (non)-pertinence du désordre, critère qui a d'abord été formulé dans la littérature physique et conjecture que la (non)-pertinence du désordre peut se lire dans l'exposant critique de la longueur de corrélation du modèle homogène.

La longueur de corrélation du modèle homogène au voisinage du point critique est donnée par l'inverse de l'énergie libre (voir [93, (1.16)]), ce qui donne un exposant critique pour la longueur de corrélation  $\nu = 1/\min(\alpha, 1)$ . Alors le critère de Harris prédit que le désordre est *pertinent* si  $\nu < 2/d$  (ici  $\alpha > 1/2$ ) et *non-pertinent* si  $\nu > 2/d$  (ici  $\alpha < 1/2$ ), où  $d$  est la dimension du système (pour le pinning  $d = 1$ ). Dans notre cas cela implique que la réponse à cette question dépend seulement de l'exposant  $\alpha$  de la loi du renouvellement  $\mathbf{P}$ .

Nous résumons les résultats principaux dans le théorème suivant. Dans cet énoncé, nous notons  $f(\beta) \underset{\text{s.v.}}{\asymp}^0 g(\beta)$  si  $\frac{f(\beta)}{g(\beta)}$  est à *variations lentes* quand  $\beta \searrow 0$  : voir (5.0.2-5.0.3) dans Appendix : Notation pour plus de détails.

**Théorème 0.2.3** (Pinning, (non)-pertinence du désordre). (i) [6, 147] Soit  $0 \leq \alpha < 1/2$ . Alors il existe  $\beta_1 > 0$  tel que  $h_c(\beta) = h_c^a(\beta) = -\lambda(\beta)$  pour tout  $\beta \in [0, \beta_1)$ . De plus pour tout  $\varepsilon > 0$ , il existe  $\beta_\varepsilon > 0$  tels que pour tout  $\beta \in [0, \beta_\varepsilon)$  et  $h \geq 0$  petit, on a

$$(1 - \varepsilon) F(0, h) \leq F(\beta, h_c(\beta) + h) \leq F(0, h). \quad (0.2.13)$$

(ii) [64, 6, 147] Soit  $\alpha > 1/2$ . Alors on a quand  $\beta \searrow 0$ ,

$$h_c(\beta) - h_c^a(\beta) \underset{\text{s.v.}}{\asymp}^0 \beta^{\max(\frac{2\alpha}{2\alpha-1}, 2)}. \quad (0.2.14)$$

(Des équivalents précis ont récemment été prouvés pour  $\alpha \in (1/2, 1)$  [50, Theorem 2.4] et  $\alpha > 1$  [24, Theorem 1.6].)

(iii) [90, 43] On suppose toujours  $\alpha > 1/2$ . Alors il existe  $\beta_1 > 0$  tel que pour tout  $\beta \in (0, \beta_1)$ , on a quand  $h \searrow 0$ ,

$$F(\beta, h_c(\beta) + h) \leq C_\beta h^2, \quad (0.2.15)$$

pour un certain  $C_\beta > 0$ .

Commentons ces résultats. Dans le cas (i) on voit que le point critique du modèle quenched correspond à celui du modèle annealed, et dans son voisinage le comportement asymptotique de l'énergie libre est le même que pour le modèle annealed/homogène (en particulier on retrouve un exposant critique  $1/\alpha$ ) : on a donc montré que le désordre est *non-pertinent* dans le cas  $\alpha < 1/2$ . Pour  $\alpha > 1/2$ , (ii) montre qu'il y a un décalage, ou *shift*, du point critique entre les modèles annealed et quenched. Quant à (iii), cela montre que l'exposant critique est au moins  $2 > 1/\min(\alpha, 1)$ , donc la transition de phase est plus régulière que dans le modèle homogène, ce qu'on appelle un lissage ou *smoothing*. Finalement on a prouvé que le désordre est *pertinent* dans le cas  $\alpha > 1/2$ , c'est-à-dire qu'on a un *shift* du point critique et un *smoothing* de la transition de phase. En particulier, ces résultats sont bien conformes au critère de Harris [98].

Mentionnons rapidement le cas “marginal”  $\alpha = 1/2$ , étudié entre autres dans [89, 20]. Il a été démontré que, en considérant  $\tau, \tau'$  deux copies indépendantes du processus de renouvellement de loi  $\mathbf{P}$ , le désordre est pertinent (en terme de shift du point critique) quand l'intersection  $\tau \cap \tau'$  est infinie  $\mathbf{P}$ -p.s. ; sinon elle est finie  $\mathbf{P}$ -p.s. et le désordre est non pertinent (pas de shift). Il est également connu qu'il n'y a pas d'effet de smoothing dans le second cas (i.e. il satisfait Théorème 0.2.3.(i)), mais la question d'en prouver un quand  $\tau \cap \tau'$  est infini  $\mathbf{P}$ -p.s. est encore ouverte aujourd'hui.

Finalement, l'étude de l'influence d'un désordre —c'est-à-dire d'inhomogénéités— sur une transition de phase comme celle d'accrochage n'est pas une tâche aisée ; mais grâce à la contribution de plusieurs mathématiciens ces dernières années et l'usage de techniques de preuve variées, elle a été exhaustivement traitée pour le modèle de pinning —du moins dans le cadre d'un champ de désordre i.i.d.. Nous allons à présent étudier une version “généralisée” (dans un certain sens) du modèle de pinning, et déterminer si elle est sujette aux mêmes résultats.

## 0.2.2 Le modèle de Poland-Scheraga généralisé

Comme mentionné précédemment, le modèle de pinning englobe un certain nombre de modèles de polymères différents ; celui qui nous intéresse en particulier est le *modèle de Poland-Scheraga de dénaturation de l'ADN*. Introduit dans les années 70 [135] pour étudier le-dit phénomène de dénaturation —c'est-à-dire la séparation des deux brins d'ADN formant un chromosome, voir [33, 32]—, le modèle de Poland-Scheraga (ou *PS*) a pu être reformulé dans la littérature mathématique comme un modèle de pinning en notant par un processus de renouvellement les indices des paires de bases étant accrochées entre elles (voir la similarité entre Fig. 0.11 et Fig. 0.12a). Il est également possible de prendre en compte l'inhomogénéité de l'ADN dans ce cadre, en observant que les paires accrochées peuvent être de nature A-T ou C-G, qui font intervenir des énergies d'interaction différentes.

Cependant, remarquons que ce modèle impose aux deux brins d'ADN d'être de même longueur et de s'accrocher parfaitement symétriquement. Ces hypothèses ont été relâchées dans la littérature biophysique, voir [78, 77, 128], définissant le modèle dit de Poland-Scheraga généralisé, ou *gPS*. Plus récemment dans [87], ce modèle a été formulé dans

un cadre mathématique, en utilisant un processus de renouvellement 2-dimensionnel pour décrire les paires de monomères interagissant, voir Fig. 0.12b.



FIGURE 0.12: Représentation des modèles PS et gPS. Dans le premier, les deux brins d'ADN s'accrochent symétriquement, et leur configuration peut être décrite en étudiant seulement un des deux brins (comme en Fig. 0.11) avec un renouvellement 1-dimensionnel. En autorisant les boucles asymétriques dans le second, on doit considérer la suite des paires de monomères accrochés qu'on décrit par un renouvellement 2-dimensionnel : ici  $\tau = \{(1, 1), (2, 2), (6, 4), \dots\}$ .

Nous allons à présent étudier le modèle gPS, et la question de l'influence d'inhomogénéités, ou du désordre, sur sa transition de phase d'accrochage. Étant très similaire au modèle PS (dont nous ne parlerons presque plus dans la suite), nous nous permettons de réutiliser les mêmes notations, avec la différence que les quantités 2-dimensionnelles sont notées en caractères gras ( $\mathbf{n} \in \mathbb{N}^2, \dots$ ).

### Construction du modèle gPS et choix du désordre

Un processus  $\tau = (\tau)_{i \geq 1}$  est un renouvellement 2-dimensionnel (ou “dans  $\mathbb{N}^2$ ”) si  $\tau_0 = (0, 0)$  et si ses incréments  $(\tau_i - \tau_{i-1})_{i \geq 1}$  sont des variables i.i.d. à valeur dans  $\mathbb{N}^2 \cup \{\infty\}$  : en particulier, ses deux coordonnées sont strictement croissantes. Alors, une paire  $(a, b) \in \tau$  signifie que la  $a$ -ème base du premier brin s'est accrochée à la  $b$ -ème base du second brin.

Comme précédemment nous supposons que les incréments sont finis P-p.s.. Nous faisons également l'hypothèse que la probabilité que les brins forment une boucle  $\tau_i - \tau_{i-1} = (a, b)$  entre deux paires accrochées ne dépend que de la taille de la boucle  $a + b$ , et varie régulièrement en celle-ci. Plus précisément il existe  $\alpha$  (dans cette thèse nous nous restreignons à  $\alpha > 0$ ) et une fonction à variations lentes  $L$  tels que

$$\mathbf{P}(\tau_1 = (a, b)) = K(a + b) := \frac{L(a + b)}{(a + b)^{\alpha+2}}, \quad (a, b) \in \mathbb{N}^2. \quad (0.2.16)$$

Soit  $\omega = (\omega_i)_{i \in \mathbb{N}^2}$  un champ aléatoire de loi notée  $\mathbb{P}$ , représentant les disparités d'énergies d'interaction d'une paire à l'autre. Nous supposons que ces variables sont toutes de même

loi, et que pour un certain  $\beta_0 > 0$  et tout  $\beta \in [0, \beta_0)$ , on a

$$\lambda(\beta) := \log \mathbb{E}[e^{\beta\omega_1}] < +\infty, \quad (0.2.17)$$

(cette hypothèse est satisfaite par toutes les lois bornées, et par un grand nombre de lois non-bornées, comme les gaussiennes ou le produit de deux gaussiennes indépendantes). Le choix précis du champ  $\omega$  (et notamment d'une hypothèse d'indépendance) est discuté ci-dessous.

Soit  $\mathbf{n} = (n_1, n_2) \in \mathbb{N}^2$ , et considérons deux brins de longueurs respectives  $n_1$  et  $n_2$ . Nous pouvons alors définir la mesure polymère quenched et contrainte comme en (0.2.3) pour un paramètre homogène  $h \in \mathbb{R}$  et une intensité de désordre  $\beta \in [0, \beta_0)$ . Nous écrivons directement sa fonction de partition, pour une réalisation  $\omega \in \mathbb{R}^{\mathbb{N}^2}$  fixée,

$$Z_{\mathbf{n},h}^{\beta,q} := \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta\omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right], \quad (0.2.18)$$

où  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket := \llbracket 1, n_1 \rrbracket \times \llbracket 1, n_2 \rrbracket \subset \mathbb{N}^2$ , et nous avons rajouté le terme  $-\lambda(\beta)$  dans l'énergie d'interaction pour renormaliser les résultats suivants (on rappelle aussi que cette quantité dépend de  $\omega$ , mais on l'omet dans la notation). Naturellement, on définit aussi son analogue libre  $Z_{\mathbf{n},h}^{\beta,\text{free}}$  en retirant la contrainte  $\mathbf{1}_{\{n \in \tau\}}$  ci-dessus. (Mentionnons qu'une variante plus générale a été étudiée dans [87, 19] en donnant une entropie spécifique aux brins libres, ce qui ne change ni l'énergie libre ni nos résultats).

Parlons à présent du choix du champ de désordre  $\omega$ . Dans le modèle PS (i.e. de pinning unidimensionnel) nous avons considéré un champ i.i.d., et on pourrait bien sûr faire la même hypothèse ici, comme cela a été fait en [18]. Cependant, bien que cette supposition soit tout à fait adéquate pour étudier l'accrochage d'un polymère à une surface 2-dimensionnelle irrégulière, elle est moins pertinente si nous souhaitons modéliser l'accrochage de deux polymères inhomogènes entre eux : les énergies d'interaction liées aux paires  $(1, 1)$  et  $(1, 2)$  n'ont aucune raison d'être indépendantes puisque la même base du premier brin apparaît dans les deux. C'est pourquoi nous proposons ici d'étudier un champ de désordre défini pour tout  $\mathbf{i} = (i_1, i_2)$  par

$$\omega_{\mathbf{i}} := f(\hat{\omega}_{i_1}, \bar{\omega}_{i_2}), \quad (0.2.19)$$

où  $\hat{\omega} = (\hat{\omega}_i)_{i \geq 1}$ ,  $\bar{\omega} = (\bar{\omega}_i)_{i \geq 1}$  sont deux champs de désordre qui décrivent les compositions respectives du premier et second brin, et  $f$  est une fonction décrivant les interactions entre les bases. Nous supposons que chacun des champs  $\hat{\omega}$  et  $\bar{\omega}$  est composé de variables i.i.d., et qu'ils sont indépendants entre eux. Nous notons  $\hat{\mathbb{P}}$  et  $\bar{\mathbb{P}}$  leur loi respectives, et par abus de notation  $\mathbb{P} = \hat{\mathbb{P}} \otimes \bar{\mathbb{P}}$ .

Insistons sur le fait que le champ  $\omega$  est très corrélé —sur chaque “ligne” et chaque “colonne” de  $\mathbb{N}^2$ . Très souvent en mécanique statistique, l'étude de l'influence d'un champ de désordre non-i.i.d. sur une transition de phase est une question difficile —le modèle PS (pinning unidimensionnel) présenté plus tôt ne fait pas exception, voir par exemple [27, 134], [25, 26] ou [56] où les auteurs étudient différents choix de désordre corrélé pour le modèle PS. Cela motive encore davantage notre choix de modèle et de désordre, car étudier



sa (non)-pertinence nous permettra non seulement de quantifier mathématiquement l'effet des inhomogénéités de l'ADN sur sa transition de dénaturation, mais également d'améliorer notre compréhension générale des systèmes corrélés en mécanique statistique.

### Transition de phase et modèle annealed

Avant d'étudier la (non)-pertinence de ce choix de désordre, nous devons bien sûr, comme pour le modèle PS précédemment, définir l'énergie libre puis quantifier la transition de phase homogène. Notons que la question de l'existence de l'énergie libre du modèle gPS présente une subtilité supplémentaire par rapport au modèle PS : nous devons choisir une direction  $\gamma \in \mathbb{R}_+^*$  pour la limite  $\mathbf{n} \rightarrow \infty$ .

**Proposition 0.2.4** (gPS désordonné, énergie libre). *Soient  $\gamma > 0$ ,  $h \in \mathbb{R}$  et  $\beta \geq 0$ . Alors pour toute suite  $(m(n))_{n \geq 1}$  telle que  $\lim_{n \rightarrow \infty} m(n)/n = \gamma$ , on a les convergences*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{\mathbf{n},h}^{\beta,q} = \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{\mathbf{n},h}^{\beta,\text{free}} =: F_\gamma(\beta, h), \quad (0.2.20)$$

où  $\mathbf{n} := (n, m(n))$ , au sens  $\mathbb{P}$ -p.s. et dans  $L^1(\mathbb{P})$ . De plus,  $(\beta, h) \mapsto F_\gamma(\beta, h + \lambda(\beta))$  est positif et convexe, et croissant en chaque coordonnée; et  $\gamma \mapsto F_\gamma(\beta, h)$  est croissant et continu. En outre, pour tous  $0 < \gamma_1 \leq \gamma_2$ ,

$$F_{\gamma_1}(\beta, h) \leq F_{\gamma_2}(\beta, h) \leq \frac{\gamma_2}{\gamma_1} F_{\gamma_1}(\beta, h). \quad (0.2.21)$$

Cette proposition a déjà été démontrée dans [18, Thm. 1.1 et Prop. 2.1] pour un modèle avec désordre i.i.d.; mais leur preuve s'adapte sans difficulté à notre choix de désordre, car la seule hypothèse nécessaire est qu'à  $\tau$  fixé, les énergies récupérées par la trajectoire  $\tau$  sont i.i.d.. C'est bien le cas sous nos hypothèses, puisque  $\tau$  est  $\mathbf{P}$ -p.s. strictement croissant sur chaque coordonnée, et pour tous  $\mathbf{i}_1, \dots, \mathbf{i}_k \in \mathbb{N}^2$ , les  $\omega_{\mathbf{i}_l}$ ,  $1 \leq l \leq k$  sont indépendants si et seulement si les indices  $\mathbf{i}_l$  sont tous sur des lignes et colonnes différentes de  $\mathbb{N}^2$ .

Notons que (0.2.21) implique que nos résultats ne dépendent pas du choix de  $\gamma > 0$  à un facteur constant près; en particulier, on définit le point critique

$$h_c^q(\beta) = h_c(\beta) := \sup\{h \in \mathbb{R}; F_\gamma(\beta, h) = 0\}, \quad (0.2.22)$$

similairement à (0.2.8), et on remarque qu'il ne dépend pas du choix de  $\gamma > 0$ . On peut donc se restreindre au cas  $\gamma = 1$ , ce que nous ferons notamment dans certaines preuves.

Concernant le modèle homogène ( $\beta = 0$ ), son comportement critique a été étudié dans [87, Theorem. 1.2], et nous le rappelons ici.

**Théorème 0.2.5** (gPS homogène, comportement critique). *Soit  $\gamma > 0$ . Alors  $h_c(0) = 0$ . De plus il existe une fonction à variations lentes  $L_\alpha$  et  $c_{\alpha,\gamma} > 0$  tels que, quand  $h \searrow 0$ ,*

$$F_\gamma(0, h) \sim c_{\alpha,\gamma} L_\alpha(1/h) h^{1/\min(\alpha,1)}, \quad (0.2.23)$$

(si  $\alpha = 0$ , cela signifie  $F_\gamma(0, h) = o(h^p)$  quand  $h \searrow 0$  pour tout  $p > 0$ ).

Remarquons que l'exposant critique  $1/\min(\alpha, 1)$  est le même que pour le modèle PS.

La dernière notion qu'il nous reste à redéfinir est le modèle annealed. Similairement à (0.2.10), on rappelle  $\lambda(\beta) = \log \mathbb{E}[e^{\beta\omega_1}] < +\infty$  et on pose pour tout  $\beta \in [0, \beta_0)$ ,

$$\begin{aligned} Z_{n,h}^{\beta,a} &:= \mathbb{E}[Z_{n,h}^{\beta,q}] = \mathbb{E} \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket 1, n \rrbracket} (\beta\omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] \\ &= \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket 1, n \rrbracket} h \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] = Z_{n,h}^{0,q}. \end{aligned} \quad (0.2.24)$$

où l'introduction de la renormalisation  $-\lambda(\beta)$  dans (0.2.18) a légèrement simplifié le résultat, et comme dans (0.2.10) et la Proposition 0.2.4 nous avons utilisé que les  $\omega_i, i \in \tau$  sont i.i.d.. Encore une fois nous retrouvons le modèle homogène, et les mêmes arguments que pour (0.2.12) nous donnent

$$0 \leq h_c(\beta) \leq \lambda(\beta), \quad (0.2.25)$$

où (0.2.12) a été translaté de  $\lambda(\beta)$  par l'introduction de la renormalisation dans (0.2.18).

### 0.2.3 Résultats sur la (non)-pertinence du désordre

Nous avons à présent tous les outils à notre disposition pour étudier la (non)-pertinence du désordre sur le modèle gPS. Rappelons que le critère de Harris mentionné plus tôt nous incite à comparer l'exposant critique  $1/\alpha$  à  $2/d$ , où  $d$  est la dimension du système désordonné. Dans notre cadre, déterminer cette dimension  $d$  est une question plus complexe qu'il n'y paraît : le champ de désordre  $\omega$  est indexé par  $\mathbb{N}^2$ , mais il est construit avec deux champs unidimensionnels  $\widehat{\omega}, \bar{\omega}$  donc a un degré de liberté limité (pour des brins de taille  $(n, n) \in \mathbb{N}^2$ , on compte  $2n$  variables et non  $n^2$ ). En fait, nous allons démontrer que le critère sur  $\alpha$  de (non)-pertinence du désordre *dépend du choix de loi du désordre*  $\mathbb{P}$ , et peut correspondre au critère de Harris pour chacune des dimensions  $d \in \{1, 2\}$ .

Dans cette thèse nous nous restreignons à deux polymères distribués selon la même loi  $\widehat{\mathbb{P}} = \bar{\mathbb{P}}$  et à une fonction d'interaction  $f$  produit, soit

$$\omega_i := f(\widehat{\omega}_{i_1}, \bar{\omega}_{i_2}) = \widehat{\omega}_{i_1} \times \bar{\omega}_{i_2}, \quad \mathbf{i} = (i_1, i_2) \in \mathbb{N}^2. \quad (0.2.26)$$

Nous verrons que ce choix englobe déjà un certain nombre de subtilités du modèle, et certaines de nos preuves restent valables pour des fonctions  $f$  plus générales. Nous supposons également que les variables  $\widehat{\omega}_1, \bar{\omega}_1$  ne sont pas constantes  $\mathbb{P}$ -p.s. (sinon le modèle est homogène), et l'hypothèse  $\lambda(\beta) < +\infty$  peut être garantie en supposant  $\mathbb{E}[e^{\frac{1}{2}\beta\widehat{\omega}_1^2}] < +\infty$  pour tout  $\beta \in [0, \beta_0)$ , en utilisant  $2xy \leq x^2 + y^2, x, y \in \mathbb{R}$ . Cette hypothèse est notamment vérifiée pour  $\widehat{\omega}_1$  gaussien ou borné. Dans la suite nous notons le  $k$ -ème moment  $m_k := \mathbb{E}[\widehat{\omega}_1^k], k \in \mathbb{N}$  —en particulier  $\mathbb{E}[\omega_1^k] = m_k^2$ .

Enfin, nous définissons ici une loi de désordre qui jouera un rôle particulier dans nos résultats. Pour tout  $x > 0$ , on pose

$$\mathbb{P}_{\pm x}(\widehat{\omega}_1 = x) = \mathbb{P}_{\pm x}(\widehat{\omega}_1 = -x) = 1/2, \quad (0.2.27)$$

(et  $\bar{\omega}_1$  a bien sûr la même loi). En particulier ces distributions  $(\mathbb{P}_{\pm x})_{x>0}$  sont caractérisées par les identités  $\mathbb{E}_{\pm x}[\hat{\omega}_1] = m_1 = 0$ , et  $\text{Var}_{\pm x}(\hat{\omega}_1^2) = m_4 - m_2^2 = 0$  —car cela implique  $\hat{\omega}_1^2 = x^2 \mathbb{P}_{\pm x}$ -p.s..

Nous formulons nos résultats dans le cas  $\gamma = 1$ , en commençant par la non-pertinence du désordre pour certains  $\alpha > 0$ .

**Théorème 0.2.6** (gPS, non-pertinence). *Supposons l'une des deux hypothèses suivantes satisfaite :*

- (a)  $0 < \alpha < 1/2$ ,
- (b)  $0 < \alpha < 1$  et  $\mathbb{P} = \mathbb{P}_{\pm x}$  pour un  $x > 0$ ,

alors il existe  $\beta_1 > 0$  tel que pour tout  $\beta \in [0, \beta_1)$ , on a (i)  $h_c(\beta) = h_c^a(\beta) = 0$ , et (ii) pour  $h \searrow 0$ , on a

$$F(\beta, h) \underset{\text{s.v.}}{\asymp}^0 F(0, h). \quad (0.2.28)$$

La borne supérieure  $F(\beta, h) \leq F(0, h)$  était déjà donnée par l'inégalité de Jensen et la définition du modèle annealed, donc l'affirmation intéressante ici est la borne inférieure. On remarque que pour une loi de désordre  $\mathbb{P}_{\pm x}$ ,  $x > 0$ , ce résultat donne la non-pertinence du désordre pour tout  $\alpha < 1$ ; autrement, on ne la prouve que pour  $\alpha < 1/2$ . Ci-dessous, nous affirmons que pour  $\alpha > 1/2$  et  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ , le point critique est décalé :

**Théorème 0.2.7** (gPS, shift sous  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ ). *Soient  $\alpha > 1/2$  et  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  pour tout  $x > 0$ . Alors il existe des fonctions à variations lentes  $L_1, L_2$  et un  $\beta_1$  tels que, pour tout  $\beta \in (0, \beta_1)$ ,*

$$L_1(1/\beta)\beta^{\max(\frac{2\alpha}{2\alpha-1}, 2)} \geq h_c(\beta) \geq \begin{cases} L_2(1/\beta)\beta^{\max(\frac{2\alpha}{2\alpha-1}, 2)} & \text{si } m_1 \neq 0, \\ L_2(1/\beta)\beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)} & \text{si } m_1 = 0, m_4 > m_2^2. \end{cases} \quad (0.2.29)$$

Rappelons que les hypothèses  $m_1 \neq 0$  ou  $m_4 > m_2^2$  couvrent toutes les lois  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ ,  $x > 0$ . Alors ce théorème implique  $h_c(\beta) > h_c^a(\beta) = 0$  pour tout  $\beta \in (0, \beta_1)$  dans ce cas, et donc le désordre est pertinent (en terme de shift). Remarquons que dans le cas  $m_1 = 0$ , nos bornes ne sont pas optimales : nous discuterons de cela plus bas, et en particulier nous conjecturons que l'ordre exact est celui donné par la borne inférieure, soit  $h_c(\beta) \underset{\text{s.v.}}{\asymp}^0 \beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)}$  quand  $\beta \searrow 0$ . En particulier, nous le prouvons pour  $\hat{\omega}$  qui suit une loi gaussienne centrée.

Dans le cas  $\mathbb{P} = \mathbb{P}_{\pm x}$  pour un  $x > 0$  et  $\alpha > 1$ , nous démontrons également un shift du point critique.

**Théorème 0.2.8** (gPS, shift sous  $\mathbb{P}_{\pm x}$ ). *Soit  $\mathbb{P} = \mathbb{P}_{\pm x}$  pour un  $x > 0$ , et  $\alpha > 1$ . Alors quand  $\beta \searrow 0$ , on a*

$$h_c(\beta) \underset{\text{s.v.}}{\asymp}^0 \beta^{\max(\frac{2\alpha}{\alpha-1}, 4)}. \quad (0.2.30)$$

Finalement, nous avons tous les résultats souhaités sur le shift du point critique, à part pour les valeurs marginales :  $\alpha = 1/2$  si  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  pour tout  $x > 0$ , et  $\alpha = 1$  si  $\mathbb{P} = \mathbb{P}_{\pm x}$  pour un  $x > 0$ .

En ce qui concerne l'exposant critique nous prouvons dans le cas  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ ,  $x > 0$  et  $\alpha > 1/2$  qu'il y a bien un smoothing de la transition. Le cas  $\mathbb{P} = \mathbb{P}_{\pm x}$  est discuté plus bas. Similairement à [90], nous devons faire une hypothèse additionnelle sur la loi  $\mathbb{P}$  (principalement pour des raisons techniques), plus précisément sur l'entropie relative d'une dilatation du désordre.

**Hypothèse 0.2.9** (Entropie relative du désordre dilaté). Soit  $\tilde{\mathbb{P}}_\delta$  la loi de  $(1 + \delta)\hat{\omega}_1$  pour tout  $\delta \in \mathbb{R}$ . Il existe  $c > 0$  et  $\delta_0 > 0$  tels que pour tout  $\delta \in (-\delta_0, \delta_0)$ ,

$$H(\tilde{\mathbb{P}}_\delta | \mathbb{P}) := \tilde{\mathbb{E}}_\delta \left[ \log \frac{d\tilde{\mathbb{P}}_\delta}{d\mathbb{P}} \right] \leq c \delta^2, \quad (0.2.31)$$

(où  $H(\tilde{\mathbb{P}}_\delta | \mathbb{P}) \geq 0$  est bien posé dès que  $\frac{d\tilde{\mathbb{P}}_\delta}{d\mathbb{P}}$  existe).

Notons que cette hypothèse est vérifiée pour un grand nombre de lois non-bornées, par exemple si  $\hat{\omega}, \bar{\omega}$  sont des champs gaussiens; en revanche elle n'est pas vérifiée par les lois bornées, en particulier les  $\mathbb{P}_{\pm x}$ ,  $x > 0$ . On a alors le résultat suivant.

**Théorème 0.2.10** (gPS, smoothing). Soit  $\mathbb{P}$  qui satisfait l'Hypothèse 0.2.9. Alors pour tout  $\beta \in (0, \beta_0)$ , il existe  $c_\beta$  et  $h_\beta > 0$  tels que pour tout  $h \in (0, h_\beta)$ ,

$$F(\beta, h_c(\beta) + h) \leq c_\beta h^2. \quad (0.2.32)$$

Dans le cas où  $\alpha > 1/2$ , on observe que le désordre lisse la transition de phase (car  $2 > 1/\min(\alpha, 1)$ ). Avec le Théorème 0.2.7, cela prouve complètement la pertinence du désordre (shift et smoothing) pour  $\alpha > 1/2$  et  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  satisfaisant l'Hypothèse 0.2.9.

Nous résumons nos résultats dans la Table 1.

Shift de $h_c(\beta)$	$\mathbb{P}_{\pm x}, x > 0$	$\mathbb{P} \neq \mathbb{P}_{\pm x}, m_1 = 0^*$	$\mathbb{P} \neq \mathbb{P}_{\pm x}, m_1 \neq 0$
$0 < \alpha < 1/2$	$0^\dagger$	$0^\dagger$	$0^\dagger$
$1/2 < \alpha < 1$	$0^\dagger$	$\beta^{\frac{4\alpha}{2\alpha-1} \ddagger}$	$\beta^{\frac{2\alpha}{2\alpha-1} \ddagger}$
$1 < \alpha < 2$	$\beta^{\frac{2\alpha}{\alpha-1}}$	$\beta^4 \ddagger$	$\beta^2 \ddagger$
$2 < \alpha$	$\beta^4$	$\beta^4 \ddagger$	$\beta^2 \ddagger$

TABLE 1: Ordre du shift du point critique  $h_c(\beta)$  du modèle gPS, au sens de  $\prec_{s.v.}^0$ . \* Dans cette colonne les bornes supérieures sont seulement conjecturées.  $^\dagger$  On prouve qu'il n'y a ni shift ni smoothing.  $^\ddagger$  On prouve un smoothing sous l'Hypothèse 0.2.9.

## Commentaires

On observe que, contrairement au modèle PS vu précédemment, la (non)-pertinence du désordre dépend non seulement de  $\alpha > 0$  (donc de  $\mathbf{P}$ ), mais également du choix de loi de désordre  $\mathbb{P}$ . En particulier, les distributions  $\mathbb{P}_{\pm x}$ ,  $x > 0$  satisfont le critère de Harris pour

une dimension 2, alors que toutes les autres lois le satisfont pour une dimension 1. Une manière d'anticiper ce résultat est de s'intéresser au champ  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$  des contributions du désordre à la fonction de partition. Les corrélations de ce champ jouent un rôle important dans la preuve du Théorème 0.2.6, en particulier on remarque qu'elles sont nulles sur les colonnes et lignes de  $\mathbb{N}^2$  si et seulement si  $\mathbb{P} = \mathbb{P}_{\pm x}$  pour un  $x > 0$ . Par conséquent, bien qu'il ne soit pas i.i.d., le champ  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$  n'est pas corrélé pour une loi  $\mathbb{P}_{\pm x}$ , mais a de fortes corrélations sur chaque ligne et colonne pour les autres lois. Cela semble expliquer la similarité du critère sur  $\alpha$  de (non)-pertinence du désordre que nous avons prouvé, avec la prédiction de Harris pour un champ de dimension 2 dans le premier cas, et de dimension 1 dans le second.

Il est particulièrement intéressant de faire la comparaison avec [18] où les auteurs ont considéré le modèle gPS muni d'un désordre i.i.d.. Ils démontrent que le désordre est non-pertinent pour  $\alpha < 1$ , et obtiennent le même décalage du point critique que notre Théorème 0.2.8 pour  $\mathbb{P} = \mathbb{P}_{\pm x}, x > 0$ . Ainsi, l'influence de notre choix de désordre non-indépendant, pour ces lois  $\mathbb{P}_{\pm x}$ , est identique à celle du désordre i.i.d..

Nous pouvons également faire la comparaison avec le modèle PS discuté plus tôt, et nos résultats pour  $\mathbb{P} \neq \mathbb{P}_{\pm x}, x > 0$ . Dans les deux cas le désordre est non-pertinent pour  $\alpha < 1/2$ , et si  $\alpha > 1/2$  on retrouve le même ordre du shift du point critique dans les Théorèmes 0.2.3 et 0.2.7 pour  $m_1 \neq 0$ . Cette comparaison est d'autant plus forte que, dans notre cadre et dans le cas marginal  $\alpha = 1/2$  (non énoncé ci-dessus), nous prouvons qu'une hypothèse suffisante pour la non-pertinence du désordre est que  $\tau^{(1)} \cap \tau'^{(1)}$  est fini  $\mathbf{P}$ -p.s., où  $\tau, \tau'$  sont deux copies indépendantes du renouvellement, et  $\tau^{(i)}$  désigne la projection du processus sur sa  $i$ -ème coordonnée ; en particulier, c'est un renouvellement unidimensionnel, et on peut facilement constater qu'il a une queue  $\mathbf{P}(\tau^{(i)} = a) = \tilde{L}(a)a^{1+\alpha}$ . Cela coïncide très visiblement au résultat énoncé sous le Théorème 0.2.3 pour le cas marginal du modèle PS (désordre non-pertinent si  $\tau \cap \tau'$  est fini), et nous avons de bonnes raisons de conjecturer que cette hypothèse est également nécessaire. Additionnellement, pour les lois  $\mathbb{P}_{\pm x}$  et  $\alpha = 1$ , nous obtenons comme hypothèse suffisante que  $\tau \cap \tau'$  est fini  $\mathbf{P}$ -p.s., ce qui coïncide avec [18, Theorem 1.3] pour le modèle gPS avec désordre i.i.d..

## 0.2.4 Autre approche : la limite de faible couplage

Pour compléter notre étude de la (non)-pertinence du désordre, nous proposons ici une autre approche de ce problème, inspirée de [3, 44] : étudier la limite d'échelle du modèle. Il s'agit d'un travail en cours, nous présentons ici nos premiers résultats et conjectures.

Cette méthode a d'abord été proposée pour étudier un polymère en environnement aléatoire dans [3], puis a été adaptée à d'autres modèles dont le pinning dans [44]. L'idée à l'origine de cette approche est de choisir une intensité de désordre  $\beta = \beta_n$  qui dépend de  $n$  et décroît vers 0, pour mettre en évidence un régime entre celui du modèle homogène ( $\beta = 0$ ) et celui désordonné ( $\beta > 0$  constant), appelé *désordre intermédiaire*. En appliquant au modèle gPS les observations faites dans les articles sus-cités (voir aussi [49]), nous pouvons conjecturer que si le désordre est pertinent (au sens vu plus tôt, dit "de Harris"), alors il est

possible de choisir des suites  $\beta_n, h_n$  décroissant vers le point critique homogène  $(0, 0)$ , telles que la fonction de partition  $Z_{n, h_n}^{\beta_n, q}$  admet une limite non-triviale —en particulier qui porte toujours l'aléa du champ  $\omega$ . En revanche si le désordre est non-pertinent, nous conjecturons que toute limite d'échelle intermédiaire de la fonction de partition est triviale. Nous allons adapter cette approche à notre cadre, plus spécifiquement dans le cas  $1/2 < \alpha < 1$  : nous avons observé que, selon le choix de loi  $\mathbb{P}$ , le désordre peut être pertinent ou non-pertinent pour ces valeurs de  $\alpha$  (on rappelle la Table 1) ; donc nous espérons construire une limite d'échelle qui sera constante pour les lois  $\mathbb{P}_{\pm x}, x > 0$ , et aléatoire sinon.

La méthode employée dans [3, 44] est un développement en *chaos polynomial* de la fonction de partition, et la convergence vers un *chaos de Wiener*. Pour  $\beta \geq 0$  et une réalisation de  $\omega$ , nous définissons  $\zeta_i = \zeta_i(\beta) := e^{\beta\omega_i - \lambda(\beta)} - 1, \mathbf{i} \in \mathbb{N}^2$  ; alors pour  $\mathbf{t} = (t_1, t_2) \in (\mathbb{R}_+^*)^2$  et  $n$  tels que  $n\mathbf{t} \in \mathbb{N}^2$ , on peut réécrire la fonction de partition libre,

$$\begin{aligned} Z_{n, h}^{\beta, \text{free}} &= \mathbf{E} \left[ \prod_{\mathbf{1} \preccurlyeq \mathbf{i} \preccurlyeq n\mathbf{t}} \left( 1 + (e^h \zeta_i + e^h - 1) \mathbf{1}_{\{\mathbf{i} \in \tau\}} \right) \right] \\ &= 1 + \sum_{k=1}^{\min(nt_1, nt_2)} \sum_{\mathbf{0} = \mathbf{i}_0 \prec \mathbf{i}_1 \prec \dots \prec \mathbf{i}_{k-1} \prec \mathbf{i}_k \preccurlyeq n\mathbf{t}} \prod_{l=1}^k \left( (e^h \zeta_{\mathbf{i}_l} + e^h - 1) u(\mathbf{i}_l - \mathbf{i}_{l-1}) \right), \end{aligned} \quad (0.2.33)$$

où on a posé la *fonction de renouvellement*  $u(\mathbf{i}) := \mathbf{P}(\mathbf{i} \in \tau)$  pour tout  $\mathbf{i} \in \mathbb{N}^2$ , et on définit les relations d'ordre  $\prec, \preccurlyeq$  par

$$\begin{aligned} \mathbf{i} \prec \mathbf{j} &\iff i_1 < j_1 \quad \text{et} \quad i_2 < j_2, \\ \mathbf{i} \preccurlyeq \mathbf{j} &\iff i_1 \leq j_1 \quad \text{et} \quad i_2 \leq j_2, \end{aligned} \quad (0.2.34)$$

pour tous  $\mathbf{i} = (i_1, i_2), \mathbf{j} = (j_1, j_2) \in \mathbb{N}^2$  (et on rappelle que  $\tau$  est  $\mathbf{P}$ -p.s. strictement croissant sur chaque coordonnée). Nous avons choisi la fonction de partition libre  $Z_{n, h}^{\beta, \text{free}}$  par commodité, mais ses variantes quenched  $Z_{n, h}^{\beta, q}$  et conditionnée  $Z_{n, h}^{\beta, \text{cond}}$  peuvent bien sûr être réécrites de manière similaire.

Commençons par le terme  $k = 1$  de cette somme. Quand  $h \rightarrow 0$ , nous pouvons le réécrire,

$$(1 + o(1)) \sum_{\mathbf{i} \in \llbracket \mathbf{1}, n\mathbf{t} \rrbracket} \zeta_i u(\mathbf{i}) + (1 + o(1)) h \sum_{\mathbf{i} \in \llbracket \mathbf{1}, n\mathbf{t} \rrbracket} u(\mathbf{i}), \quad (0.2.35)$$

$\llbracket \mathbf{1}, n\mathbf{t} \rrbracket := \{\mathbf{i} \in \mathbb{N}^2, \mathbf{1} \preccurlyeq \mathbf{i} \preccurlyeq n\mathbf{t}\}$ , et nous souhaitons étudier sa convergence pour une suite  $(\beta_n, h_n) \rightarrow (0, 0)$  que nous choisirons plus loin. Le comportement asymptotique de la fonction de renouvellement  $u$  est déjà connu : on rappelle que  $L$  est défini dans (0.2.16), et on cite le résultat principal de [154].

**Proposition 0.2.11** (Limite de la fonction de renouvellement, [154]). *Soit  $0 < \alpha < 1$ . Il existe une fonction  $\varphi : (\mathbb{R}_+^*)^2 \rightarrow \mathbb{R}_+$ , continue, telle que pour tout  $\mathbf{s} \in (\mathbb{R}_+^*)^2$ ,*

$$\lim_{n \rightarrow \infty} n^{2-\alpha} L(n) \mathbf{P}(\lfloor n\mathbf{s} \rfloor \in \tau) = \varphi(\mathbf{s}). \quad (0.2.36)$$

Si on écrit  $\mathbf{s} = r e^{i\theta}$  sous forme polaire, alors  $\varphi(\mathbf{s}) = r^{\alpha-2} a(\theta)$  où  $a : [0, \pi/2] \rightarrow \mathbb{R}_+$  est continue et vérifie  $a(0) = a(\pi/2) = 0$ .

En associant ce résultat avec la convergence d'une somme de Riemann, on obtient

$$\frac{L(n)}{n^\alpha} \sum_{\mathbf{i} \in \llbracket 1, nt \rrbracket} u(\mathbf{i}) \xrightarrow{n \rightarrow +\infty} \int_{[0, t]} \varphi(\mathbf{s}) d\mathbf{s}, \quad (0.2.37)$$

où  $[0, t] := \{\mathbf{s} \in \mathbb{R}^2; \mathbf{0} \preceq \mathbf{s} \preceq \mathbf{t}\}$ , et cette intégrale est finie car  $\varphi(\mathbf{s}) \leq c \|\mathbf{s}\|^{\alpha-2}$  et on a supposé  $\alpha < 1$ . En appliquant ce résultat à (0.2.35), on remarque que le second terme converge si on choisit  $h = h_n := \hat{h} L(n) n^{-\alpha}$  et  $\hat{h} \in \mathbb{R}$ .

Étudions à présent la convergence du premier terme dans (0.2.35). Dans [3, 44], les auteurs choisissent la suite  $(\beta_n)_{n \geq 1}$  telle qu'il converge vers une intégrale stochastique de la fonction  $\varphi$  contre un bruit blanc. Ce résultat n'est malheureusement pas adaptable à notre cas de façon directe, car le champ  $(\zeta_i)_{i \in \mathbb{N}^2}$  n'est pas i.i.d. —tout comme  $(\omega_i)_{i \in \mathbb{N}^2}$ , il est corrélé sur chaque ligne et colonne de  $\mathbb{N}^2$ . Cependant, nous pouvons obtenir un résultat similaire en démontrant d'une part que le champ  $(\zeta_i)_{i \in \mathbb{N}^2}$  converge, avec un scaling approprié, vers un *champ continu aléatoire limite*  $\mathcal{M}$  —qui n'est pas un bruit blanc, mais possède la même structure de corrélation par lignes et colonnes que  $\zeta$ —, et d'autre part que le premier terme de (0.2.35) converge vers *l'intégrale de la fonction  $\varphi$  contre ce champ  $\mathcal{M}$* , où nous définirons l'intégrale d'une fonction contre un champ 2-dimensionnel rigoureusement plus loin.

Comme nous nous attendons à obtenir une limite qui dépende du choix de  $\mathbb{P}$ , nous notons  $\mathfrak{P}$  l'ensemble des mesures de probabilité  $\mathbb{P}$  satisfaisant toutes les hypothèses faites en (0.2.19) et (0.2.26), et nous posons

$$\begin{aligned} \mathfrak{P}_0 &:= \{\mathbb{P}_{\pm x}; x > 0\}, \\ \mathfrak{P}_1 &:= \{\mathbb{P} \in \mathfrak{P}; m_1 \neq 0\}, \\ \mathfrak{P}_2 &:= \{\mathbb{P} \in \mathfrak{P}; m_1 = 0, m_4 > m_2^2\}, \end{aligned} \quad (0.2.38)$$

qui forment une partition de  $\mathfrak{P}$ . Nous prenons une suite  $(\beta_n)_{n \geq 1}$ , nous rappelons  $\zeta_i = e^{\beta_n \omega_i - \lambda(\beta)}$  —1,  $\mathbf{i} \in \mathbb{N}^2$  (notons que  $\zeta_i$  est centré), et nous posons pour tous  $\mathbf{s} \in (\mathbb{R}_+^*)^2$  et  $n \in \mathbb{N}$ ,

$$M_n(\mathbf{s}) := \sum_{\mathbf{i} \in \llbracket 1, [ns] \rrbracket} \zeta_i. \quad (0.2.39)$$

Enfin, nous posons  $\mathcal{M}$  le champ gaussien sur  $(\mathbb{R}_+^*)^2$  centré et de matrice de covariance

$$K(\mathbf{u}, \mathbf{v}) := (u_1 \wedge v_1)(u_2 \wedge v_2)(u_1 \vee v_1 + u_2 \vee v_2), \quad (0.2.40)$$

pour tous  $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in (\mathbb{R}_+^*)^2$ , où  $\wedge$  (resp.  $\vee$ ) désigne le minimum (resp. maximum). Alors nous prouvons le théorème suivant.

**Théorème 0.2.12.** *Soit  $\mathbf{t} \in (\mathbb{R}_+^*)^2$ .*

(i) *Si  $\mathbb{P} \in \mathfrak{P}_1$ , et si  $\beta_n \rightarrow 0$  et  $n^{1/2}\beta_n \rightarrow +\infty$  quand  $n \rightarrow \infty$ , alors*

$$\left( \frac{1}{n^{3/2}\beta_n} M_n(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} \xrightarrow{(d)} \left( m_1 \sqrt{m_2 - m_1^2} \mathcal{M}(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} . \quad (0.2.41)$$

(ii) Si  $\mathbb{P} \in \mathfrak{P}_2$ , et si  $\beta_n \rightarrow 0$  et  $n^{1/4}\beta_n \rightarrow +\infty$  quand  $n \rightarrow \infty$ , alors

$$\left( \frac{1}{n^{3/2}\beta_n^2} M_n(\mathbf{s}) \right)_{\mathbf{s} \in [0,t]} \xrightarrow{(d)} \left( \frac{m_2}{2} \sqrt{m_4 - m_2^2} \mathcal{M}(\mathbf{s}) \right)_{\mathbf{s} \in [0,t]}. \quad (0.2.42)$$

Les convergences ci-dessus ont lieu pour la topologie de la norme uniforme sur  $[0, t]$ . Dans les résultats (i) et (ii) ci-dessus, nous affirmons que la limite du champ discret  $M_n$ ,  $n \geq 1$  est un champ aléatoire non-trivial, non-isotropique, et très fortement corrélé sur chaque ligne et chaque colonne (voir Fig. 0.13).

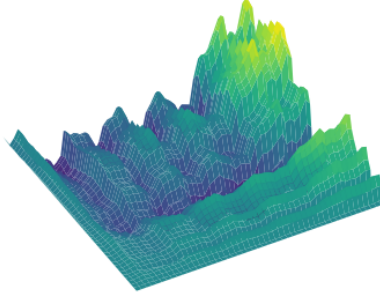


FIGURE 0.13: Simulation du champ gaussien  $\mathcal{M}$ , limite en loi du champ (discret) de désordre  $(\zeta_i)_{i \in \mathbb{N}^2}$  du modèle gPS.

Ce théorème nous permet de démontrer la convergence suivante en distribution et dans  $L^2(\mathbb{P})$ ,

$$\frac{L(n)}{c_r(\beta_n)^r n^{\alpha - \frac{1}{2}}} \sum_{i \in \llbracket 1, nt \rrbracket} \zeta_i u(\mathbf{i}) \xrightarrow{n \rightarrow +\infty} \int_{[0,t]} \varphi(\mathbf{s}) \mathcal{M}(d\mathbf{s}), \quad (0.2.43)$$

pour  $\mathbb{P} \in \mathfrak{P}_r$ ,  $r \in \{1, 2\}$ , et où

$$c_1 := m_1 \sqrt{m_2 - m_1^2}, \quad \text{et} \quad c_2 := \frac{m_2}{2} \sqrt{m_4 - m_2^2}. \quad (0.2.44)$$

Ce résultat n'est pas immédiat, notamment il nous demande de définir l'intégrale d'une fonction contre le champ 2-dimensionnel et fortement corrélé  $\mathcal{M}$ . Nous présentons les idées principales de la construction de cette intégrale ci-dessous. Avec la convergence (0.2.43), il apparaît naturellement que nous devons choisir  $\beta_n := \hat{\beta} (n^{\frac{1}{2} - \alpha} L(n))^{\frac{1}{r}}$  avec  $r \in \{1, 2\}$  et  $\hat{\beta} \in [0, +\infty)$  pour faire converger le premier terme dans (0.2.35). Remarquons aussi que notre restriction à  $\alpha \in (1/2, 1)$  nous assure que les hypothèses du Théorème 0.2.12 sont satisfaites par ce choix de suite  $(\beta_n)_{n \geq 1}$ .

## 0.2.5 Intégration contre le champ $\mathcal{M}$

Pour définir l'intégrale contre le champ  $\mathcal{M}$ , nous nous inspirons de la construction de l'intégrale d'Itô et de l'isométrie du même nom. Considérons par exemple un cas simple, unidimensionnel, comme une fonction non-aléatoire  $f : [0, t] \rightarrow \mathbb{R}$  et un processus  $(X)_{s \in [0,t]}$



convenable (disons un mouvement brownien) : une manière de définir l'intégrale de  $f$  contre  $X$  est de déterminer une mesure  $\mu_X$  sur  $[0, t)$  telle que

$$\int_{[0,t)} f^2 d\mu_X = \mathbb{E} \left[ \left( \int_{[0,t)} f dX \right)^2 \right], \quad (0.2.45)$$

(dans le cas du mouvement brownien, il s'agit de la mesure de Lebesgue). Plus précisément, pour  $[u, v) \subset [0, t)$  on pose  $\int_{[u,v)} 1_{[u,v)} dX := X([u, v)) := X(v) - X(u)$ , et la propriété d'isométrie (0.2.45) permet d'étendre linéairement la définition à un ensemble de fonctions "intégrables" (i.e. telles que  $\int_0^t f^2 d\mu_X < \infty$ ). Cependant, cette construction ne fonctionne pas si  $X$  est un processus plus complexe, par exemple dont les incréments sont corrélés : en effet, pour  $[u, v), [u', v') \subset [0, t)$  disjoints,

$$\begin{aligned} \mathbb{E} \left[ \left( X([u, v)) + X([u', v')) \right)^2 \right] &= \mathbb{E} \left[ X([u, v))^2 \right] + \mathbb{E} \left[ X([u', v'))^2 \right] \\ &\quad + \mathbb{E} \left[ X([u, v)) X([u', v')) \right]. \end{aligned}$$

Si  $X$  est une martingale alors le dernier terme est nul, donc on peut construire une mesure  $\mu_X$  donnant une isométrie comme en (0.2.45). Mais cela ne peut pas être garanti pour un processus  $X$  quelconque, et dans ce cas on ne peut pas définir d'intégrale stochastique par cette méthode.

Mentionnons que l'intégration stochastique a été construite pour les processus multidimensionnels, voir [41] entre autres ; mais toujours dans le cadre de martingales (à plusieurs indices) ou processus similaires, ce qui ne convient pas pour le champ  $\mathcal{M}$  apparaissant dans le modèle gPS. Il existe également d'autres généralisations de l'intégrale stochastique, mais à notre connaissance elles s'intéressent principalement à des processus stationnaires (comme le mouvement brownien fractionnaire). Dans cette thèse nous proposons une autre approche, qui permet de rendre compte de la structure de corrélation du champ  $\mathcal{M}$  dans l'intégrale : plutôt que de définir une mesure sur  $[0, t) \subset \mathbb{R}^2$  (ou  $[0, t)$  en dimension 1), nous construisons une mesure  $\nu_{\mathcal{M}}$  sur  $[\mathbf{0}, \mathbf{t})^2 \subset \mathbb{R}^4$  (ou  $[0, t)^2$ ), telle que pour des fonctions  $f, g$  "intégrables", on a

$$\int_{[0,t)^2} f(\mathbf{u})g(\mathbf{v}) d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) = \mathbb{E} \left[ \left( \int_{[0,t)} f d\mathcal{M} \right) \left( \int_{[0,t)} g d\mathcal{M} \right) \right]. \quad (0.2.46)$$

Plus précisément, nous définissons d'abord les intégrales contre  $\mathcal{M}$  de fonction simples (c'est-à-dire les indicatrices de rectangles  $[u, v) \subset [0, t)$ ), et nous étendons cette construction à un ensemble de fonctions "intégrables" par bilinéarité de (0.2.46). La mesure  $\nu_{\mathcal{M}}$  sera appelée la *mesure de covariance* du champ  $\mathcal{M}$ .

Avant d'énoncer nos résultats, introduisons quelques notations. Soit  $\mathbf{t} \in (\mathbb{R}_+)^2$ , et posons

$$\mathcal{S}_{\mathbf{t}} := \{ [u, v) ; \mathbf{0} \preceq u \preceq v \preceq \mathbf{t} \} \cup \{ \emptyset \},$$

l'ensemble des rectangles inclus dans  $[0, \mathbf{t})$ , fermés en bas et à gauche, ouverts en haut et à droite. Dans cette thèse nous nous restreignons à  $[0, \mathbf{t})$  un rectangle borné, mais nos résultats

se généralisent à  $\mathbb{R}^2$  (nous donnons plus de détails dans le Chapitre 4, voir Remark 4.1.8). Notons que  $\mathcal{S}_t$  est un semi-anneau d'ensembles, c'est-à-dire stable par intersection finie, et pour tous  $A, B \in \mathcal{S}_t$ ,  $A \setminus B$  peut s'écrire comme une union finie d'éléments de  $\mathcal{S}_t$ . De plus,  $\sigma(\mathcal{S}_t) = \text{Bor}([0, t])$ .

On appelle  $L^2(\mathbb{P})$ -*champ aléatoire sur  $\mathcal{S}_t$*  (ou sur  $[0, t]$ ) une famille  $X$  de variables aléatoires dans  $L^2(\mathbb{P})$  indicée par les rectangles :  $X(A) \in L^2(\mathbb{P})$ ,  $A \in \mathcal{S}_t$ . Si  $X$  est une fonction de  $[0, t]$  dans  $L^2(\mathbb{P})$ , alors on peut construire un champ  $(X(A))_{A \in \mathcal{S}_t}$  en considérant les *incrément*s de  $X$  sur chaque rectangle : pour  $[\mathbf{u}, \mathbf{v}] \in \mathcal{S}_t$ ,

$$X([\mathbf{u}, \mathbf{v}]) := X(v_1, v_2) - X(u_1, v_2) - X(v_1, u_2) + X(u_1, u_2),$$

(notons que  $X$  doit également être définie sur les bords droit et supérieur de  $[0, t]$ ). Toute fonction sur  $[0, t]$  permet de définir un champ sur  $\mathcal{S}_t$ , mais la réciproque n'est pas vraie (par exemple le bruit blanc n'est pas défini ponctuellement). Comme nos principaux résultats sont valables pour des champs  $X$  généraux, nous ne distinguerons pas les notations entre champ et incrément, mais nous précisons explicitement si un champ est construit à partir des incréments d'une fonction définie ponctuellement.

Reprenons l'équation (0.2.46), que nous cherchons à obtenir en construisant une mesure de covariance et une intégrale stochastique appropriées. Pour un champ aléatoire  $X$  (défini ponctuellement ou non), il est naturel de poser  $\int_{[\mathbf{0}, \mathbf{t}]} 1_{[\mathbf{u}, \mathbf{v}]} dX := X([\mathbf{u}, \mathbf{v}])$  pour tout rectangle  $[\mathbf{u}, \mathbf{v}] \in \mathcal{S}_t$ . En particulier, la mesure de covariance de  $X$ , si elle existe, doit vérifier  $\nu_X(A \times B) = \mathbb{E}[X(A)X(B)]$  pour tous  $A, B \in \mathcal{S}_t$ . Dans la proposition suivante, nous montrons qu'une telle mesure existe pour le champ  $\mathcal{M}$ .

**Proposition 0.2.13.** *Soit  $\mathcal{M}$  un champ gaussien sur  $(\mathbb{R}_+)^2$ , centré et de matrice de covariance  $K(\cdot, \cdot)$  définie par (0.2.40), et soit  $\mathbf{t} \in (\mathbb{R}_+)^2$ . Il existe une unique mesure (positive, finie)  $\nu_{\mathcal{M}}$  sur  $\text{Bor}([0, \mathbf{t}]^2)$  telle que pour tous  $A, B \in \mathcal{S}_t$ ,  $\nu_{\mathcal{M}}(A \times B) = \mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)]$ . De plus, pour toutes fonctions  $g, h : [0, t] \rightarrow \mathbb{R}$  mesurables, positives ou bornées, on a*

$$\int_{[0, \mathbf{t}]^2} g(\mathbf{u})h(\mathbf{v}) d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) = \int_{[0, \mathbf{t}]} g(\mathbf{u}) \left( \int_0^{t_1} h(x, u_2) dx + \int_0^{t_2} h(u_1, y) dy \right) d\mathbf{u}. \quad (0.2.47)$$

L'expression de  $\nu_{\mathcal{M}}$  donnée par (0.2.47) met clairement en évidence la structure de corrélation par lignes et par colonnes du champ gaussien  $\mathcal{M}$ . De plus, dès qu'un champ  $X$  admet une mesure de covariance sur  $\text{Bor}([0, \mathbf{t}]^2)$ , positive, finie, alors l'intégrale contre le champ  $X$  est bien définie.

**Théorème 0.2.14.** *Soit  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  un champ aléatoire sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , et supposons qu'il existe une mesure positive, finie  $\nu$  sur  $\text{Bor}([0, \mathbf{t}]^2)$  telle que pour tous  $A, B \in \mathcal{S}_t$ ,*

$$\mathbb{E}[X(A)X(B)] = \nu(A \times B). \quad (0.2.48)$$

Soit  $L_X^2$  l'ensemble des fonctions mesurables  $f : [0, \mathbf{t}] \rightarrow \mathbb{R}$  qui satisfont

$$\int_{[0, \mathbf{t}]^2} |f(\mathbf{u})f(\mathbf{v})| d\nu(\mathbf{u}, \mathbf{v}) < +\infty.$$

Alors  $L_X^2$  est un espace vectoriel, et contient  $L^\infty([0, \mathbf{t}])$ . De plus, si on pose pour tout  $A \in \mathcal{S}_t$ ,

$$1_A \diamond X := X(A) \in L^2(\mathbb{P}),$$

alors l'application  $f \mapsto f \diamond X$  peut être prolongée linéairement sur  $L_X^2$  de sorte que pour tous  $g, h \in L_X^2$ ,

$$\mathbb{E}[(g \diamond X)(h \diamond X)] = \int_{[0, \mathbf{t}]^2} g(\mathbf{u})h(\mathbf{v}) d\nu(\mathbf{u}, \mathbf{v}). \quad (0.2.49)$$

La variable aléatoire  $g \diamond X \in L^2(\mathbb{P})$  sera appelée *l'intégrale de  $g$  contre  $X$* , et sera parfois notée

$$\int g dX := g \diamond X.$$

L'affirmation que l'ensemble  $L^\infty([0, \mathbf{t}])$  des fonctions mesurables bornées est contenu dans  $L_X^2$  est évidente, car  $\nu([0, \mathbf{t}]^2) < \infty$  par hypothèse; mais elle n'en est pas moins importante, puisqu'elle prouve que le prolongement construit par le Théorème 0.2.14 n'est pas vide.

Les propriétés de cette intégrale sont détaillées dans le Chapitre 4; en particulier, nous démontrons dans la Proposition 4.1.9 que la fonction  $\varphi$  définie en (0.2.36) est bien intégrable contre  $\mathcal{M}$ , soit  $\varphi \in L_{\mathcal{M}}^2$ . Nous définissons également les intégrales de rang supérieur  $k > 1$ , qui serviront pour décrire les limites de tous les termes du développement polynomial de la fonction de partition. Finalement, nous conjecturons le résultat suivant.

**Conjecture 0.2.15.** Soit  $\tau$  vérifiant (0.2.16) avec  $\alpha \in (1/2, 1)$ . Soit  $\mathbb{P} \in \mathfrak{P}_r$ , pour un  $r \in \{1, 2\}$ , et soient  $(\beta_n)_{n \geq 1}, (h_n)_{n \geq 1}$  deux suites vérifiant

$$\beta_n \left( n^{\alpha - \frac{1}{2}} L(n)^{-1} \right)^{\frac{1}{r}} \xrightarrow{n \rightarrow \infty} \hat{\beta} \in [0, +\infty), \quad \text{et} \quad h_n n^\alpha L(n)^{-1} \xrightarrow{n \rightarrow \infty} \hat{h} \in \mathbb{R}. \quad (0.2.50)$$

Alors pour tout  $\mathbf{t} \succ \mathbf{0}$ , nous avons la convergence

$$Z_{[\mathbf{nt}], h_n}^{\beta_n, \omega, \mathbf{q}} \xrightarrow[n \rightarrow +\infty]{(d)} Z_{\mathbf{t}, \hat{h}}^{\hat{\beta}, \mathcal{M}, \mathbf{q}}, \quad (0.2.51)$$

où la variable aléatoire  $Z_{\mathbf{t}, \hat{h}}^{\hat{\beta}, \mathcal{M}, \mathbf{q}}$  est définie par le développement en chaos

$$Z_{\mathbf{t}, \hat{h}}^{\hat{\beta}, \mathcal{M}, \mathbf{q}} := 1 + \sum_{k=1}^{+\infty} \frac{1}{k!} \int \cdots \int_{[0, \mathbf{t}]^k} \psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k) \prod_{j=1}^k \left( c_r \hat{\beta}^r d\mathcal{M}(\mathbf{s}_j) + \hat{h} ds_j \right), \quad (0.2.52)$$

dans lequel  $c_r > 0$  est défini par (0.2.44), et la fonction  $\psi_{\mathbf{t}}^{\mathbf{q}}$  est symétrique, pour toute suite  $\mathbf{0} =: \mathbf{s}_0 \prec \mathbf{s}_1 \prec \cdots \prec \mathbf{s}_k \prec \mathbf{t}$  elle vaut

$$\psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \varphi(\mathbf{s}_1) \varphi(\mathbf{s}_2 - \mathbf{s}_1) \cdots \varphi(\mathbf{t} - \mathbf{s}_k) \quad (0.2.53)$$

et  $\psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k) = 0$  si  $\mathbf{s}_1, \dots, \mathbf{s}_k$  ne peuvent pas être ordonnés pour  $\prec$ . De plus, la série dans (0.2.52) converge au sens  $L^2$ .

Enfin, les fonctions de partitions libre<sup>free</sup> et conditionnées<sup>cond</sup> vérifient le même résultat, en remplaçant  $\psi_t^q$  respectivement par  $\psi_t^{\text{free}}$  symétrique, valant

$$\psi_t^{\text{free}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \varphi(\mathbf{s}_1)\varphi(\mathbf{s}_2 - \mathbf{s}_1) \cdots \varphi(\mathbf{s}_k - \mathbf{s}_{k-1}), \quad (0.2.54)$$

pour  $\mathbf{0} =: \mathbf{s}_0 \prec \mathbf{s}_1 \prec \cdots \prec \mathbf{s}_k \prec \mathbf{t}$  qui peuvent être ordonnés et 0 sinon; et pour la seconde par

$$\psi_t^{\text{cond}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \frac{\psi_t^q(\mathbf{s}_1, \dots, \mathbf{s}_k)}{\varphi(\mathbf{t})}. \quad (0.2.55)$$

Il s'agit d'un travail en cours, présenté en détail dans le Chapitre 4. Dans l'état actuel, nous avons prouvé que la série (0.2.52) est bien posée (en particulier toutes les intégrales contre  $\mathcal{M}$  sont bien définies) et converge dans  $L^2(\mathbb{P})$ , ainsi que la convergence du premier terme du développement polynomial de  $Z_{[nt],h_n}^{\beta_n, \omega, q}$  vers son analogue continu, soit (0.2.43). Il nous reste à démontrer la même convergence pour les termes d'ordre  $k \in \mathbb{N}$  quelconque; alors la convergence (0.2.51) sera une conséquence directe.

## Commentaires

Avec ce résultat, nous avons bien démontré l'existence d'un régime de désordre intermédiaire et d'une limite d'échelle non triviale pour le modèle gPS dans le cas  $1/2 < \alpha < 1$  et  $\mathbb{P} \in \mathfrak{P}_1 \cup \mathfrak{P}_2$ , là où nous avons déjà montré que le désordre est pertinent au sens de Harris —c'est-à-dire qu'il affecte le point critique et l'exposant critique pour toute intensité  $\beta > 0$  fixée : on rappelle la Table 1. Un fait remarquable est que la définition du champ  $\mathcal{M}$ , et par extension la loi de la limite d'échelle, ne dépendent pas de  $\mathbb{P} \in \mathfrak{P}_1 \cup \mathfrak{P}_2$ , et seule la renormalisation est différente à travers le choix de la suite  $(\beta_n)_{n \geq 1}$ . Remarquons également que l'introduction de l'exposant  $r \in \{1, 2\}$  correspond à la différence d'ordre entre les deux bornes inférieures prouvées dans le Théorème 0.2.7. En outre, notre preuve de la borne supérieure du shift repose sur une estimation du second moment de la fonction de partition, que nous obtenons de manière précise si  $\mathbb{P} \in \mathfrak{P}_1$  ou pour une loi gaussienne  $\hat{\omega}_1 \sim \mathcal{N}(0, \sigma^2)$ ,  $\sigma^2 > 0$ , mais pas pour une loi  $\mathbb{P} \in \mathfrak{P}_2$  générale à cause de la structure de corrélation du champ  $\omega$ . Avec la Conjecture 0.2.15, nous obtenons une meilleure borne sur le second moment de  $Z_{n,h}^{\beta,q}$ , laquelle est suffisante pour obtenir une borne supérieure de l'ordre de  $\beta^{4\alpha/(2\alpha-1)}$  sur le shift de  $h_c(\beta)$  pour n'importe quelle loi  $\mathbb{P} \in \mathfrak{P}_2$ , et donc termine la preuve de notre conjecture dans la Table 1 pour  $\alpha < 1$  (cela sera discuté en détail dans la Section 3.3.3).

Concernant le Théorème 0.2.12, nous pouvons anticiper l'ordre de la renormalisation nécessaire à la convergence en calculant la variance de  $M_n(\mathbf{1})$ , laquelle est une somme de termes  $\mathbb{E}[\zeta_i \zeta_j]$ ,  $\mathbf{i}, \mathbf{j} \in \mathbb{N}^2$ . On remarque que les seuls termes contribuant sont les  $\mathbf{i} = \mathbf{j}$ , et les  $\mathbf{i}$  "alignés" avec  $\mathbf{j}$  : c'est-à-dire sur la même ligne ou colonne de  $\mathbb{N}^2$ , ce que nous noterons  $\mathbf{i} \leftrightarrow \mathbf{j}$ ; en revanche si  $\mathbf{i} \not\leftrightarrow \mathbf{j}$ , alors  $\mathbb{E}[\zeta_i \zeta_j] = 0$  car  $\zeta_i$  et  $\zeta_j$  sont indépendants et centrés. Remarquons que les couples  $(\mathbf{i}, \mathbf{j})$  satisfaisant la première condition sont au nombre de  $n^2$ , et pour la seconde  $n^2(2n-1)$ . Ainsi, quand  $n \rightarrow \infty$  et  $\beta_n \rightarrow 0$ , on a

$$\mathbb{E}\left[\left(M_n(\mathbf{1})\right)^2\right] = (1 + o(1))\left(n^2 \mathbb{E}[\zeta_1^2] + 2n^3 \mathbb{E}[\zeta_{(1,1)} \zeta_{(1,2)}]\right), \quad (0.2.56)$$

et avec des développements limités (ou un calcul exact dans le cas  $\mathbb{P} \in \mathfrak{P}_0$ ), nous obtenons (voir Lemma 4.2.1 pour les détails)

$$\mathbb{E} [\zeta_i \zeta_j] = \begin{cases} 0 & \text{si } i \leftrightarrow j, \\ (m_2^2 - m_1^4) \beta_n^2 + o(\beta_n^2) & \text{si } i = j, \\ m_1^2(m_2 - m_1^2) \beta_n^2 + o(\beta_n^2) & \text{si } i \leftrightarrow j, i \neq j \text{ et } \mathbb{P} \in \mathfrak{P}_1, \\ \frac{m_2^2}{4}(m_4 - m_2^2) \beta_n^4 + o(\beta_n^4) & \text{si } i \leftrightarrow j, i \neq j \text{ et } \mathbb{P} \in \mathfrak{P}_2, \\ 0 & \text{si } i \leftrightarrow j, i \neq j \text{ et } \mathbb{P} \in \mathfrak{P}_0. \end{cases} \quad (0.2.57)$$

Notons que ce résultat est tout à fait comparable au calcul des corrélations du champ  $(e^{\beta \omega_i})_{i \in \mathbb{N}^2}$  mentionné dans la section précédente. On remarque donc que dans (0.2.56), si  $\mathbb{P} \in \mathfrak{P}_1 \cup \mathfrak{P}_2$  et si  $(\beta_n)_{n \geq 1}$  ne décroît pas trop vite vers 0, alors la contribution des indices  $i = j$  à la variance est négligeable comparée aux indices  $i \leftrightarrow j$  (nous verrons même dans (4.2.8) que, en remplaçant  $M_n(\mathbf{1})^2$  par  $M_n(\mathbf{u})M_n(\mathbf{v})$  dans ce calcul, nous obtenons la fonction de corrélation  $K$  définie dans (0.2.40)). Nous voyons également dans cette heuristique que la seule différence entre  $\mathfrak{P}_1$  et  $\mathfrak{P}_2$  est dans la renormalisation requise, donc dans le choix de  $(\beta_n)_{n \geq 1}$ ; mais la limite reste la même.

En revanche pour  $\mathbb{P} \in \mathfrak{P}_0$ , le second terme disparaît dans (0.2.56), et on n'observe plus la même limite du champ  $M_n$ . Pour aller plus loin, il faudrait renormaliser le champ  $M_n$  par  $n^{-1} \beta_n^{-1}$  pour espérer obtenir une convergence, avec une suite  $(\beta_n)_{n \geq 1}$  satisfaisant  $\beta_n \rightarrow 0$  et  $n \beta_n \rightarrow +\infty$ . Mais si nous utilisons ce résultat pour obtenir une convergence similaire à (0.2.43) du cas  $\mathbb{P} \in \mathfrak{P}_1 \cup \mathfrak{P}_2$ , nous observons que le choix de  $\beta_n$  imposé par (0.2.35) diverge pour  $\alpha < 1$ . Finalement, cela implique que notre développement en chaos polynomial ne permet pas de trouver une limite d'échelle non nulle pour  $\mathbb{P} \in \mathfrak{P}_0$  et  $\alpha < 1$ , ce qui coïncide parfaitement avec le régime de non-pertinence du désordre au sens de Harris prouvé dans le Théorème 0.2.6 (on rappelle également la Table 1).

## 0.2.6 Synthèse

Le modèle gPS est une variante plus complexe du modèle de pinning (ou PS). En nous inspirant des outils développés pour le modèle de pinning, nous avons étudié l'influence de l'introduction d'inhomogénéités dans le modèle gPS sur sa transition d'accrochage. Nous avons répondu assez largement à cette question avec des quantifications du shift du point critique et du smoothing de la transition de phase, en remarquant que ces valeurs dépendent du choix de loi de désordre  $\mathbb{P}$  dans la partition  $\mathfrak{P}_0 \cup \mathfrak{P}_1 \cup \mathfrak{P}_2$ . Nous avons également complété ces résultats avec une approche de limite de faible couplage dans le régime  $1/2 < \alpha < 1$ , mettant encore en évidence les variations de comportement du modèle selon le choix de  $\mathbb{P}$ .

Ces résultats répondent à notre problématique initiale, à savoir quantifier l'influence des inhomogénéités sur la transition de phase d'accrochage. En outre, ils nous donnent des intuitions supplémentaires pour étudier le comportement des systèmes de mécanique statistique soumis à un désordre corrélé. Néanmoins, de très nombreuses questions restent ouvertes sur ce modèle :

- *Valeurs marginales.* Dans notre étude nous avons exclu les valeurs marginales,  $\alpha = 1$  si  $\mathbb{P} \in \mathfrak{P}_0$  et  $\alpha = 1/2$  sinon. Comme mentionné plus haut, nous obtenons des hypothèses suffisantes sur  $\tau$  pour prouver la non-pertinence dans les deux cas marginaux, et nous anticipons que ces hypothèses sont également nécessaires (similairement à [18] pour un désordre i.i.d.). Ce n'est qu'une conjecture, et la prouver demanderait certainement un travail technique important. Mentionnons également la valeur marginale  $\alpha = 0$  pour laquelle, dans l'esprit de [9], nous anticipons que le désordre est *fortement non-pertinent* (i.e. le point critique n'est décalé pour aucun  $\beta \geq 0$ ). Une difficulté notable pour les cas  $\alpha = 1$  et  $\alpha = 0$  est que les estimées sur les processus de renouvellement bi-dimensionnels sont respectivement très complexes (voir [17]) et inexistantes (il n'existe pas de domaine d'attraction  $\alpha$ -stable pour  $\alpha = 0$ ).
- *Intensité de désordre élevée.* Nous n'avons prouvé nos résultats que pour des intensités de désordre  $\beta$  faibles. Bien que ce soit suffisant pour étudier la (non)-pertinence du désordre au sens de Harris, l'étude de valeurs plus élevées serait intéressante, mais n'est pas évidente. Un premier résultat accessible serait de montrer pour  $\alpha < 1/2$  et  $\beta \rightarrow \beta_0$  l'existence d'un régime *fortement désordonné*, dans l'idée de [149, Theorem 2.1, Corollary 3.2]. Pour les valeurs intermédiaires de  $\beta$  en revanche, il n'est même pas clair que  $\beta \mapsto h_c(\beta)$  est monotone (rappelons que nous avons introduit la renormalisation  $-\lambda(\beta)$  dans la fonction de partition); les arguments de [88] ne s'adaptent pas au modèle gPS désordonné non-i.i.d., donc leur étude demande un travail supplémentaire.
- *Autres transitions.* Il est prouvé dans [87] que le modèle gPS est sujet à d'autres transitions de phase, dites de *condensation*. Nous n'en avons pas parlé dans cette thèse, mais étudier l'influence du désordre sur ces transitions est une question intéressante. En particulier, des résultats récents [86] sur des pinning unidimensionnels peuvent laisser penser que les transitions de condensation, contrairement à celle d'accrochage, disparaissent dans le modèle désordonné.
- *Choix du désordre.* Concernant notre choix de désordre  $\omega$ , nous conjecturons que nos résultats s'étendent à une fonction  $f$  plus générale (on rappelle (0.2.26)), car l'élément déterminant pour la majeure partie de nos résultats est le calcul des corrélations du champ  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$  —plus précisément si elles sont nulles, d'ordre  $\beta^2, \beta^4 \dots$ . En revanche nous avons fait l'hypothèse que les suites  $\hat{\omega}, \bar{\omega}$  sont indépendantes, ce qui n'est pas du tout vrai pour les brins d'ADN que nous souhaitons modéliser (une base A sur un brin fait face à une base T etc). Nous pourrions donc choisir d'étudier un désordre de la forme

$$\omega_{\mathbf{i}} := f(\hat{\omega}_{i_1}, \hat{\omega}_{i_2}), \quad \mathbf{i} = (i_1, i_2) \in \mathbb{N}^2, \quad (0.2.58)$$

où  $(\hat{\omega}_i)_{i \geq 1}$  est une suite de variables aléatoires : ce modèle a notamment fait l'objet d'études numériques, voir [78, 77]. Cependant ce choix de désordre est bien plus complexe à étudier dans un cadre mathématique —même sous l'hypothèse que  $\hat{\omega}$  est un champ i.i.d.— car les arguments de la Proposition 0.2.4 ne s'appliquent plus, par conséquent l'énergie libre quenched n'est pas clairement définie. Une autre difficulté

est que le modèle annealed ne coïncide plus avec le modèle homogène, et son étude semble difficile.

- *Autres limites d'échelle.* Enfin nous pouvons étudier la limite d'échelle de désordre intermédiaire du modèle gPS pour  $\alpha > 1$ . Dans le développement (0.2.56) et selon le scaling  $(\beta_n)_{n \geq 1}$ , il n'est pas inenvisageable que le premier terme du développement contribue lui aussi à la limite pour certains  $\alpha$  quand  $\mathbb{P} \in \mathfrak{P}_2$ . Le champ limite  $\mathcal{M}$  est-il affecté? Et pour  $\mathbb{P} \in \mathfrak{P}_0$ , le calcul de la covariance nous laisse conjecturer que son champ limite est, pour  $\alpha > 1$  et un scaling approprié, un produit de mouvements browniens indépendants,  $(\mathcal{B}^1(s_1)\mathcal{B}^2(s_2))_{(s_1, s_2) \in (\mathbb{R}_+^*)^2}$ . Construire ces limites, par exemple dans le régime  $1 < \alpha < 2$ , serait intéressant pour compléter notre étude. Toutefois, adapter le Théorème 0.2.12 n'est pas la seule difficulté supplémentaire dans le cas  $1 < \alpha < 2$ . En effet, dans ce régime le processus de renouvellement  $\tau$ , une fois renormalisé, n'explore pas le quadrant  $(\mathbb{R}_+^*)^2$  aussi librement que pour  $\alpha < 1$  : il va se localiser autour de la diagonale avec des fluctuations d'ordre  $n^{1/\alpha}$  (voir [17]). Cela a certainement des conséquences non-triviales sur la convergence (0.2.43), sur la construction de l'intégrale à la limite et sur le résultat attendu (Conjecture 0.2.15). Il n'est pas impossible que le cas  $\alpha > 2$  soit plus facile à traiter (où les fluctuations du renouvellement autour de la diagonale sont d'ordre  $n^{1/2}$ ).

## Organisation de la thèse

Cette thèse est séparée en deux parties, une pour chaque approche proposée. Dans la Partie I “The Interacting Partially Directed Self-Avoiding Walk”, nous étudions le modèle du même nom en commençant par le Chapitre 1 consacré à la variante interagissant avec un mur dur. Nous y démontrons les résultats énoncés dans Théorème 0.1.2, Proposition 0.1.3, Théorème 0.1.4 et Corollaire 0.1.5 concernant la transition de surface. Dans le Chapitre 2, nous revenons au modèle sans mur et démontrons les Théorèmes 0.1.6 et 0.1.7. Comme la construction de ces deux variantes de l'IPDSAW sont différentes, les deux chapitres peuvent être lus indépendamment même s'ils présentent des similarités.

Dans la Partie II “The generalized Poland-Scheraga model”, nous nous penchons sur le modèle gPS désordonné. Le Chapitre 3 démontre tous nos résultats sur la question de la (non)-pertinence du désordre, du Théorème 0.2.6 au 0.2.10. Dans le Chapitre 4, nous démontrons d'abord le Théorème 0.2.12 de convergence du champ de désordre corrélé, puis nous construisons rigoureusement la mesure de covariance de  $\mathcal{M}$  et l'intégrale contre un champ (Proposition 0.2.13 et Théorème 0.2.14), et présentons nos résultats intermédiaires nous menant à la Conjecture 0.2.15.

Les Parties I et II sont complètement indépendantes. Les Chapitres 1 et 2 sont également indépendants, chacun définissant et étudiant une variante différente du modèle IPDSAW (avec et sans mur). Les motivations du Chapitre 4 sont fortement reliées aux résultats du Chapitre 3, donc nous ne les détaillons pas à nouveau dans le Chapitre 4 ; en revanche, nous y redéfinissons le modèle gPS afin qu'il soit indépendant mathématiquement du chapitre

précédent.

Enfin, nous présentons dans le Chapitre “Conclusion and perspectives” Une petite synthèse de nos résultats, et proposons des pistes de futures recherches qui complètent celles déjà présentées plus haut pour chaque approche.





# Presentation and outline of the thesis

We provide here a short introduction and some motivations to the work displayed in this document, reproducing the first few paragraphs of the previous Chapter “Introduction et vue d’ensemble”; as well as an outline of the thesis.

## Presentation of the thesis

The word “polymers” refers to a large family of molecules, among which lie proteins, DNA strands, plastics and many other materials. They are a ubiquitous object of interest in physics, chemistry and biophysics literature, because they play a role in numerous phenomena that can be observed or modeled. For the past decades, mathematicians have also taken hold of those models and have investigated them from their standpoint. One family of such models aims to represent the *pinning* (or similarly *wetting*) of polymers along an object in the environment, like a cell membrane, a wall or a line of defects in the recipient. Not only do they depict quite faithfully a *phase transition* phenomenon which is a major topic of interest in physics and chemistry literature, they are also approachable (or even solvable) in a purely mathematical setting and have been extensively studied in [84, 92, 100] among many other references (see the bibliography therein). We will present here one example of a wetting model, the *partially directed, self-avoiding walk in  $\mathbb{Z}^2$  interacting with a hard wall*, and display the so-called phenomenon of *wetting phase transition* in this framework. Then we will extend our analysis to a more general setting called *the pinning model*, and we will use our observations as a starting point for this thesis, questioning what elements in a model may affect the aforementioned phase transition and to what extent.

## An example of wetting of a polymer: PDSAW model interacting with a hard wall

A *polymer* is a very large molecule composed of many repeating elementary structures called *monomers* and arranged in a chain fashion. One of the most famous examples is the DNA strand, whose composing monomers are usually denoted A, T, C and G bases. A mathematical object well suited to depict those chain-shaped molecules is the *walk on a network*, each site visited by the walk corresponding to the location of a monomer, and the whole trajectory describing the full geometry of the polymer in its environment. For

instance, let us consider a walk of length  $L \in \mathbb{N}$  in  $\mathbb{Z}^2$ , visiting  $L + 1$  sites of  $\mathbb{Z}^2$  each being the location of one monomer. Since real molecules cannot overlap in space, it is natural to limit ourselves to *self-avoiding walks* in  $\mathbb{Z}^2$ , that is any site can only be visited once by the walk —we will call them SAW for short. Of course we can alter this first representation by adding other geometrical constraints on trajectories, like restraining them to be *partially directed* in a direction, let us say the right, i.e. the trajectory cannot take steps to the left; those are called Partially Directed, Self-Avoiding Walks (PDSAW). We may also consider the presence of obstacles or impediments in the environment, such as a hard wall which the polymer cannot cross (see Fig. 0.14).

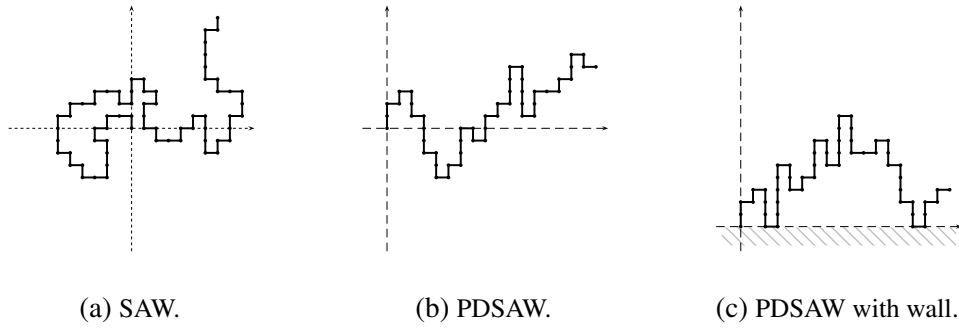


Figure 0.14: Depiction of polymers with walks on  $\mathbb{Z}^2$ . The first one is self-avoiding (SAW), the second one is additionally partially directed to the right (PDSAW), and the last one is constrained to remain above a hard wall.

The main focus of a wetting model is to investigate the behavior of a polymer in the surroundings of an object (such as a cell membrane, defects at the surface of a recipient or a wall) with which it may interact on contact. A natural question in such a setting is determining the *location* of the polymer: does it remain close to the object or wander away? How far? How many times does the polymer touch the object? To answer those questions, let us look at the representation from Fig. 0.14c and use it as a groundwork for further modeling (in particular this representation will also be the starting point of Chapter 1).

Let  $L \in \mathbb{N}$ , and let us define  $\Omega_L^{\text{wet}}$  the set of trajectories of length  $L$  described by Fig. 0.14c, that is they are self-avoiding, partially directed to the right and have to remain above the horizontal line. That is,

$$\Omega_L^{\text{wet}} := \left\{ (x_i)_{i=0}^L \in (\mathbb{Z}^2)^{L+1} \left| \begin{array}{l} x_0 = (0, 0), x_i \neq x_j \quad \forall 0 \leq i < j \leq L \\ x_i - x_{i-1} \in \{\uparrow, \downarrow, \rightarrow\} \quad \forall 1 \leq i \leq L, \\ x_i \in \mathbb{Z} \times \mathbb{N} \quad \forall 1 \leq i \leq L \end{array} \right. \right\}, \quad (0.3.1)$$

where  $\uparrow = (0, 1)$ ,  $\downarrow = (0, -1)$  and  $\rightarrow = (1, 0)$ .

We mentioned moments ago that the polymer may *interact* with the wall on contact; that is, whenever a segment of the polymer  $[x_{i-1}, x_i]$ ,  $1 \leq i \leq L$  lays along the wall:  $[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}$ . Each one of those contact interactions involves an *interaction energy*

$\delta \in \mathbb{R}$ , where we use the convention that  $\delta > 0$  corresponds to an energetic *reward* granted to the polymer, whereas  $\delta < 0$  is a penalty or energetic *cost* due to maintaining the contact (notice that our sign convention is opposed to the most common one in physics literature, however it lightens a little upcoming formulae). Let  $x \in \Omega_L^{\text{wet}}$ , and let  $H_{L,\delta}^{\text{wet}}(x)$  denote the total of interaction energies given to trajectory  $x$ , that is

$$H_{L,\delta}^{\text{wet}}(x) := \sum_{i=1}^L \delta \mathbf{1}_{\{[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}}. \quad (0.3.2)$$

With our sign convention, the Boltzmann law claims that the higher the energy reward  $H_{L,\delta}^{\text{wet}}(x)$  to a trajectory  $x$ , the more likely the trajectory is for the polymer. More precisely, the trajectory of the polymer may be seen as a random variable on  $\Omega_L^{\text{wet}}$ , with distribution

$$\mathbf{P}_{L,\delta}^{\text{wet}}(x) := \frac{e^{H_{L,\delta}^{\text{wet}}(x)}}{Z_{L,\delta}^{\text{wet}}}, \quad x \in \Omega_L^{\text{wet}}, \quad (0.3.3)$$

which is called *Gibbs measure* or *polymer measure*, and where the normalizing term

$$Z_{L,\delta}^{\text{wet}} := \sum_{x \in \Omega_L^{\text{wet}}} e^{H_{L,\delta}^{\text{wet}}(x)}, \quad (0.3.4)$$

is called the *partition function*.

## The wetting phase transition

Recall that we want to investigate the location of the polymer interacting with a wall, that is the localization of trajectories with distribution  $\mathbf{P}_{L,\delta}^{\text{wet}}$ . To do so, a useful tool in statistical mechanics is the exponential growth rate of  $Z_{L,\delta}^{\text{wet}}$ , called the *free energy* of the system,

$$\mathbf{F}^{\text{wet}}(\delta) := \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\delta}^{\text{wet}}. \quad (0.3.5)$$

In this introduction we assume it is well-defined. We may also prove that it has many properties, notably it is convex, non-decreasing, continuous and non-negative on  $\mathbb{R}$ . Moreover when  $\delta$  goes to  $-\infty$ ,  $\mathbf{F}^{\text{wet}}(\delta)$  becomes locally constant (since  $Z_{L,\delta}^{\text{wet}}$  barely varies in large  $-\delta$ ), whereas when  $\delta$  goes to  $+\infty$ ,  $\mathbf{F}^{\text{wet}}(\delta)$  also diverges to  $+\infty$ . Those observations imply that the function  $\delta \mapsto \mathbf{F}^{\text{wet}}(\delta)$  undergoes a drastic change of behavior at some  $\delta_c^{\text{wet}} \in \mathbb{R}$ , defined such as

- $\delta \mapsto \mathbf{F}^{\text{wet}}(\delta)$  is constant on  $(-\infty, \delta_c^{\text{wet}})$ ,
- $\delta \mapsto \mathbf{F}^{\text{wet}}(\delta)$  is increasing on  $(\delta_c^{\text{wet}}, +\infty)$ .

Such a point where the free energy loses regularity is called a *critical point* (see Fig. 0.15b below). The existence of such a point, and the change of behavior of  $\mathbf{F}$  from one side to the other, have substantial implications once we prove that  $\mathbf{F}$  is closely related to properties of

typical trajectories under  $\mathbf{P}_{L,\delta}^{\text{wet}}$ . Specifically, consider a polymer of length  $L \in \mathbb{N}$  and let  $N_L^{\text{wet}}(\delta)$  denote the average fraction of the polymer in contact with the wall,

$$N_L^{\text{wet}}(\delta) := \frac{1}{L} \mathbf{E}_{L,\delta}^{\text{wet}} \left[ \sum_{i=1}^L \mathbf{1}_{\{[x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}} \right]. \quad (0.3.6)$$

Then, with a straightforward computation, we may prove that for  $\delta \neq \delta_c^{\text{wet}}$ ,

$$N_L^{\text{wet}}(\delta) \xrightarrow{L \rightarrow \infty} (\mathbf{F}^{\text{wet}})'(\delta). \quad (0.3.7)$$

To keep this introduction uncluttered, we do not talk about the case  $\delta = \delta_c^{\text{wet}}$ . Recollecting the aforementioned properties of  $\mathbf{F}$  (see Fig. 0.15b), we deduce from (0.3.7) that, depending on the value of  $\delta$ , two very different behavior may appear under the measure  $\mathbf{P}_{L,\delta}^{\text{wet}}$ :

- If  $\delta < \delta_c^{\text{wet}}$ ; then for typical trajectories under  $\mathbf{P}_{L,\delta}^{\text{wet}}$ , the fraction of the polymer making contact with the wall goes to 0 as  $L$  becomes large, thus very few monomers are located along the wall. This regime  $\mathcal{D} := (-\infty, \delta_c^{\text{wet}})$  is called the *desorbed* or *delocalized phase*.
- If  $\delta > \delta_c^{\text{wet}}$ ; then typical trajectories under  $\mathbf{P}_{L,\delta}^{\text{wet}}$  have a positive fraction of the polymer making contact with the wall, thus a large amount of monomers are located along the wall. This regime  $\mathcal{A} := (\delta_c^{\text{wet}}, +\infty)$  is called the *adsorbed* or *localized phase*.

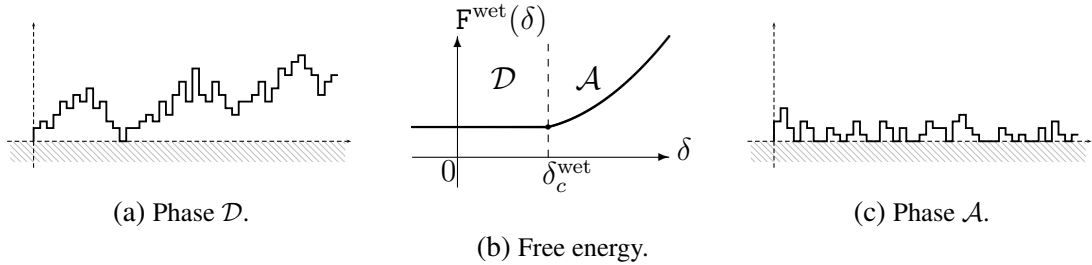


Figure 0.15: Phase diagram of the PDSAW interacting with a wall, and typical trajectories in each phase  $\mathcal{D}$  and  $\mathcal{A}$ . In  $\mathcal{D}$ , the polymer scarcely interacts with the wall and wanders away. In  $\mathcal{A}$ , the polymer interacts a lot with the wall, hence it remains close to it.

Such a change of behavior at the critical point  $\delta_c^{\text{wet}}$  —that is where  $\mathbf{F}^{\text{wet}}$  loses its regularity— is called a *phase transition*. In this setting, the polymer transitions from a delocalized regime  $\mathcal{D}$  where it scarcely interacts with the wall, hence wandering away from it; to a localized regime  $\mathcal{A}$  where it interacts a lot with the wall hence remains close to it (see Fig. 0.15). This particular phase transition is called the *adsorption* or *pinning* transition, and it will be the primary focus of most of this thesis.

## A more general setting: pinning models

The model constructed above may be called a *wetting model*, which refers to a large family of models designed to investigate similar phenomena (see e.g. [73]). Over the few past decades they have been rephrased in a setting even more general, which is commonly called *pinning models* (see [84, Chapter 1]), dropping the description of trajectories with walks in networks to the benefit of *renewal processes*; the main idea being that one does not have to describe the full trajectory and geometry of the polymer in space, but it is enough to mark down which monomers are interacting with the wall.

A renewal process  $\tau = (\tau_i)_{i \geq 1}$  is a sequence of random variables such as its increments  $(\tau_i - \tau_{i-1})$ ,  $i \geq 1$ , are i.i.d. in  $\mathbb{N} \cup \{+\infty\}$  (with the convention  $\tau_0 = 0$ ). We may abusively note this process as a subset  $\tau \subset \mathbb{N}$ . An important observation is that, letting  $(S_n)_{n \geq 1}$  be some generic random walk in a network  $\mathbb{Z}^d$ ,  $d \geq 1$ , and denoting  $\tau$  its sequence of return times to a subset  $A \subset \mathbb{Z}^d$  (e.g. the line  $\mathbb{Z} \times \{0\}^{d-1}$ ), then  $\tau$  is (usually) a renewal process; moreover, this claim (usually) still holds when additionally constraining the walk, for instance to be partially directed or to remain outside of some subset of  $\mathbb{Z}^d$ . This implies that this renewal structure is quite universal, and moreover the process  $\tau$  fully describes the contacts between the walk and the chosen subset  $A \subset \mathbb{Z}^d$ . Thereby we can construct a very similar model to the previous one, by indexing the monomers in  $\mathbb{N}$ , denoting the sequence of monomers in contact with the object with a renewal process  $\tau \subset \mathbb{N}$ , and yielding an energy  $\delta \in \mathbb{R}$  to each renewal point. For a polymer of length  $L \in \mathbb{N}$  and a sequence of contact indices  $\tau$ , the associated total energy is

$$H_{L,\delta}^{\text{pin}}(\tau) := \sum_{i=1}^L \delta \mathbf{1}_{\{i \in \tau\}},$$

thus we define the polymer measure  $\mathbf{P}_{L,\delta}^{\text{pin}}$  such as a given sequence of contact indices  $\tau$  has probability proportional to  $e^{H_{L,\delta}^{\text{pin}}(\tau)} \mathbf{P}(\tau)$ , with  $\mathbf{P}$  the initial law of the renewal sequence  $\tau$ . Similarly to the previous construction of a PDSAW near a wall, the normalizing constant  $Z_{L,\delta}^{\text{pin}}$  is called the partition function, and its exponential growth rate  $F^{\text{pin}}(\delta)$  as  $L \rightarrow \infty$  is called the free energy.

Models constructed this way (with any renewal process  $\tau$  in  $\mathbb{N}$ ) are called *pinning models*. The PDSAW model interacting with a wall presented before could be rephrased in those terms, by writing the sequence of contact indices  $\tau(x) = \{i \in \mathbb{N}, [x_{i-1}, x_i] \subset \mathbb{R} \times \{0\}\}$ ,  $x \in \Omega_L^{\text{wet}}$  as a renewal process (we do not write the details here).

The first main result regarding pinning models is that, under minimal assumptions (i.e. the tail of  $\tau_1$  must be “regularly varying”, see Appendix: Notation), they undergo the same phase transition that we displayed above for the PDSAW near a wall: there is a phase  $(-\infty, \delta_c^{\text{pin}})$  where  $F^{\text{pin}}$  is constant, and a phase  $(\delta_c^{\text{pin}}, +\infty)$  where it is increasing. Some details such as the value of the critical point  $\delta_c^{\text{pin}}$  or estimates of  $F$  around  $\delta_c^{\text{pin}}$  (which we call *critical behavior*) may change from one choice of renewal distribution  $\mathbf{P}$  to another; but they are explicitly known (see [84, Theorem 2.1]), and the (de)-localization phenomenon which we claimed in (0.3.7) (for the PDSAW near a wall) still holds true: in the first phase

the number of contacts with the object —i.e. the number of renewal points  $|\tau \cap [1, L]|$ — under  $\mathbf{P}_{L,\delta}^{\text{pin}}$  is negligible compared to the size  $L$  of the full polymer; and in the second phase that number of contacts (at the first order) is given by  $(F^{\text{pin}})'(\delta)L$  with  $(F^{\text{pin}})'(\delta) > 0$  (see [84, Theorem 2.4]).

## Perturbations of the pinning phase transition and outline of the thesis

The pinning phase transition is thus a widespread phenomenon in models of a polymer interacting with another object, and its main features are quite extensively known in mathematics literature today. Thereby, a very natural question is to investigate how sturdy this phenomenon is when the model is altered by additional elements, and to what extent it may be affected. What happens if the polymer may also interact with a second object? What happens if it may additionally interact with its surroundings by other means than contacts? What happens if all contact interaction are not the same, and vary according to inhomogeneities in the object or in the composition of the very polymer itself? Under such assumptions, does the pinning phase transition still hold? Are its critical point and critical behavior the same?

Many similar questions may be asked, each requiring a specific analysis. In this thesis, we propose to investigate those questions under two approaches, each adding a different deviation from the models we introduced above, and each yielding very different effects on the phase transition.

In Part I, we focus on the PDSAW model interacting with a wall, in which we additionally assume that the polymer interacts at each *self-touching*: that is, whenever two monomers, which are not consecutive in the polymer, lay one next to another in space. This defines a model with two quite different interactions, each of them able to greatly affect the geometry and localization of the polymer. More precisely, we define a *collapsed* regime of the polymer, where it folds over on itself into a compact globule in order to maximize self-touchings. In Chapter 1, we prove (in a slightly simplified version of this model) that the collapsed polymer still undergoes a pinning transition, across which the outer layer of the globule may either pin to the wall or wander away from it. However, this transition only affects the surface of the globule and not the whole system, hence we prove that it is not a proper phase transition, and we call this phenomenon a *surface transition*. In Chapter 2 we remove the wall (defining what is called the “IPDSAW” model) and study it sharply, obtaining an exact equivalent of its partition function and some geometric estimates, which help understanding the behavior of the collapsed IPDSAW, and additionally support our simplification and results from Chapter 1.

In Part II, we place ourselves in the general formalism of pinning models (that is with renewal processes), and we assume that all contact interaction energies are not the same. This approach can be used to study inhomogeneous polymers such as DNA strands; and with that in mind, we “generalize” the pinning model to investigate the pinning of two polymers one to another, defining the *generalized Poland-Scheraga* model (gPS), which is designed to emulate the denaturation phenomenon of DNA. In Chapter 3 we study the

effect of inhomogeneities —also called *disorder*— on the pinning (or “naturation”) phase transition, quantifying how and when it affects the critical point and behavior. Then in Chapter 4, we approach this question under the angle of *intermediate disorder* and scaling limits. We derive a scaling limit of the gPS model when the intensity of disorder decays to zero, and discuss the link between the existence of such limits and the effect of disorder on phase transitions.

Parts I and II are fully independent. Chapters 1 and 2 are also independent, since each introduces and studies a different version of the IPDSAW model (with and without wall respectively). The motivations to Chapter 4 depend on results from Chapter 3, so we do not state them in details again; however, we redefine the gPS model in Chapter 4, so it is mathematically self-contained.







## **The Interacting Partially Directed Self-Avoiding Walk**



---

# Surface transition in the Interacting Partially Directed Self-Avoiding Walk adsorbed along a hard wall

## 1.1 Introduction

In the present paper, we investigate a model for a  $1 + 1$  dimensional polymer dipped in a poor solvent and simultaneously adsorbed along a horizontal hard wall. Although the model has attracted a continuous attention in the physics literature starting in the 90's (see e.g. [1], [75], [76]) until more recently (see e.g. [123], [137] or [133]), it had, up to our knowledge, not been considered so far in the mathematical literature. This model interpolates between two families of polymer models that have been entirely solved in the last 20 years, i.e., the wetting of a  $1 + 1$ -dimensional random walk adsorbed along a hard wall (see e.g [100], [84] and [92]) and the collapse transition of the 2-dimensional Interacting Partially-Directed Self-avoiding Walk (IPDSAW) (see [54] for a review).

The coupling parameters of the model are  $\beta \in [0, \infty)$  the repulsion intensity between the monomers and the solvent around them —or, equivalently, the attraction intensity in between monomers— and  $\delta \in [0, \infty)$  the interaction intensity between the monomers and the hard wall. We will discuss in detail the phase diagram of the model in Section 1.2.2 below, but let us mention already that the phase diagram is divided into two main phases:

- $\mathcal{E}$ : an Extended phase inside which a typical trajectory has a macroscopic horizontal extension, comparable to its total length,

---

This chapter is joint work with Nicolas Pétrélis, and is currently in reviewing process [121].

- $\mathcal{C}$ : a Collapsed phase inside which the vertical width and the horizontal extension of a typical configuration are comparable, and of smaller order than its total size.

It turns out that  $\mathcal{E}$  can be divided into two sub-phases. A critical curve is indeed conjectured to partition  $\mathcal{E}$  into a Desorbed-Extended phase ( $\mathcal{DE}$ ) inside which the polymer wanders away from the hard wall and an Adsorbed-Extended phase ( $\mathcal{AE}$ ) inside which the polymer is localized along the wall (see e.g. [75, Figure 2]). The situation is more subtle in the Collapsed phase where typical configurations look roughly like a globule (see Fig. 1.2 (B)). The number of contacts between this globule and the hard-wall changes drastically inside  $\mathcal{C}$  along some other critical curve which triggers what physicists call a surface transition, that is a loss of analyticity of the second order term of the exponential development of the partition function, whereas the leading order term (i.e., the free energy) remains linear.

The aim of our paper is to investigate the collapsed phase and in particular the surface transition mentioned above. To that aim, we will introduce in Section 1.3 a simplified version of our model called the *one-bead model*. In a few words (see Section 1.3.1 for more details) every trajectory considered in our model can be decomposed into a family of sub-trajectories called beads. Those beads are typically of finite size in  $\mathcal{E}$  but are much larger inside  $\mathcal{C}$ . We can even safely conjecture that inside  $\mathcal{C}$ , a typical trajectory is made of a unique macroscopic bead (this is proven e.g. for the 2-dimensional IPDSAW in [121]). For this reason, we will restrict the set of allowed paths to those forming only one bead. This restricted version of the model turns out to be more tractable and should share many features with its non-restricted counterpart.

Let us give a short outline of the paper. In Section 1.2 below, we begin with a rigorous definition of the model and then we provide a qualitative description of its phase diagram. With Theorem 1.2.2 we identify rigorously the Collapsed phase ( $\mathcal{C}$ ) and the Extended phase ( $\mathcal{E}$ ). Section 1.3 is dedicated to the definition of the single-bead version of the model. Theorem 1.3.2, which is the most important result of the paper, is stated in Section 1.3.3 and allows us to characterize the surface transition with the help of sharp asymptotic developments of the partition function inside  $\mathcal{C}$ . We prove Theorem 1.3.2 (along with Corollary 1.3.3), in Section 1.4. We delay the proofs of Theorem 1.2.2 and Proposition 1.3.1 to Section 1.5 for they are quite standard (apart from the random walk representation introduced in Section 1.4). We then collect the proofs of technical estimates in Section 1.6. Appendix 1.A provides well-known results on the wetting model, and Appendix 1.B displays a (conditional) FKG inequality on random walks with distribution  $\mathbf{P}_\beta$  (defined in (1.2.8) below).

## 1.2 Description of the model and phase diagram

### 1.2.1 The model

For a polymer of length  $L \in \mathbb{N}$ , the set of its allowed configurations is denoted by  $\Omega_L^+$  and consists of those trajectories of a  $1 + 1$ -dimensional self-avoiding random walk on  $\mathbb{Z}^2$  taking unitary steps up, down and to the right and constrained to remain above the horizontal

axis  $y = 0$ . An alternative representation of such trajectories can be given by decomposing them according to their number of horizontal/rightward steps, and the length and orientation of the vertical stretches in between, i.e.,

$$\Omega_L^+ := \bigcup_{N \geq 1} \mathcal{L}_{N,L}^+ := \bigcup_{N \geq 1} \left\{ (\ell_i)_{i=1}^N \in \mathbb{Z}^N ; \sum_{i=1}^N |\ell_i| = L - N, \sum_{i=1}^k \ell_i \geq 0, \forall k \leq N \right\}. \quad (1.2.1)$$

Henceforth, we will only use this latter representation and we note that each vertical stretch is followed by a horizontal step—in particular we assume that all trajectories end with a horizontal step. For every  $\ell \in \Omega_L^+$ , we denote by  $N_\ell$  its horizontal extension (i.e., its number of horizontal steps) so that  $\ell \in \mathcal{L}_{N_\ell, L}^+$ .

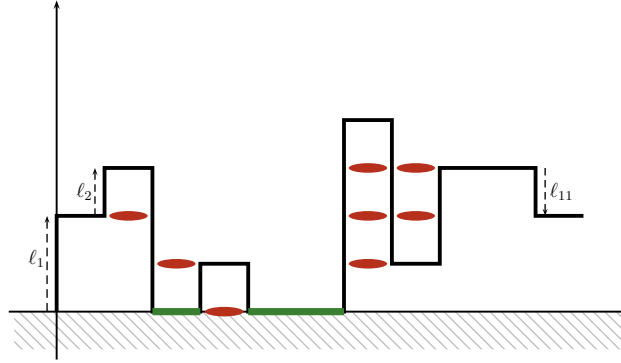


Figure 1.1: Representation of a polymer of length  $L = 29$ , with horizontal extension  $N = 11$ . Each self-touching (in red) is rewarded with an energy  $\beta$ , and each contact with the wall (in green) is rewarded with  $\delta$ .

With each configuration we associate a Hamiltonian, which takes into account that monomers are both attracting each other and adsorbed along the  $x$ -axis. To be more specific, a given  $\ell \in \Omega_L^+$  is assigned an energetic reward  $\beta \geq 0$  for every *self-touching* (i.e., a pair of neighboring sites visited non-consecutively by  $\ell$ ) and an energetic reward  $\delta \geq 0$  for every contact with the  $x$ -axis (see Fig. 1.1). Thus, define

$$H(\ell) := \beta \sum_{i=0}^{N_\ell} \ell_i \tilde{\wedge} \ell_{i+1} + \delta \sum_{k=1}^{N_\ell} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}} \quad (1.2.2)$$

where the operator  $\tilde{\wedge}$  is defined for any  $x, y \in \mathbb{Z}$  by  $x \tilde{\wedge} y := \min\{|x|, |y|\} \mathbf{1}_{\{xy \leq 0\}}$ , and where we set  $\ell_0 = \ell_{N_\ell+1} = 0$  for notational convenience. At this stage we introduce the polymer measure  $\mathbf{P}_L^{\beta, \delta}$ , a probability on  $\Omega_L^+$  defined as

$$\mathbf{P}_L^{\beta, \delta}(\ell) = \frac{e^{H(\ell)}}{Z_{L, \beta, \delta}^+}, \quad \ell \in \Omega_L^+, \quad (1.2.3)$$

where  $Z_{L,\beta,\delta}^+$  is a normalization term called *partition function* of the system. The *free energy* provides the exponential growth rate of  $Z_{L,\beta,\delta}^+$  in  $L$ . It is defined as

$$f(\beta, \delta) = \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\beta,\delta}^+, \quad (1.2.4)$$

(we will show that  $f$  is well-defined in the proof of Theorem 1.2.2).

**Remark 1.2.1.** The present model may be seen as an advanced version of IPDSAW, that was introduced in [155]. For the latter model, there is no hard-wall preventing the polymer to enter the lower half-plane and also no wetting interaction with the hard-wall. As a consequence, the allowed configurations of IPDSAW are obtained by relaxing the constraint  $\sum_{i=1}^k \ell_i \geq 0$  in (1.2.1) and the Hamiltonian by removing the term  $\delta \sum_{k=1}^N \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}}$  in (1.2.2).

## 1.2.2 Phase diagram

Since the coupling parameters  $\beta$  and  $\delta$  are both non-negative, the phase diagram is drawn on the first quadrant  $\mathcal{Q} := [0, \infty)^2$ . Similarly to what is observed for IPDSAW (see e.g. [130] or [54]), the phase diagram can be divided into a collapsed phase  $\mathcal{C}$  inside which typical trajectories undergo a self-touching saturation and an extended phase  $\mathcal{E}$  inside which the horizontal extension of a typical trajectory is comparable to its total size. To be more specific, inside  $\mathcal{C}$ , we expect that a typical trajectory of length  $L$  (i.e.,  $\ell$  sampled from  $\mathbf{P}_L^{\beta,\delta}$ ) satisfies  $H(\ell) = \beta L + o(L)$ . For this reason its horizontal extension must be small (i.e.,  $N_\ell = o(L)$ ) and its vertical stretches should be long with alternating signs. In  $\mathcal{E}$ , in turn, a typical trajectory is expected to be composed of  $O(L)$  vertical stretches of finite length (which is also what is expected at  $\beta = 0$ ).

Let us now briefly explain (with three simple observations) why the free energy  $f(\beta, \delta)$  is equal to  $\beta$  in  $\mathcal{C}$ . First,  $f(\beta, \delta) \geq \beta$  for every  $(\delta, \beta) \in \mathcal{Q}$ . This inequality derives from restricting the computation of  $Z_{L,\beta,\delta}^+$  to a unique trajectory  $\tilde{\ell} \in \mathcal{L}_{\sqrt{L},L}$  given by

$$\tilde{\ell}_i = (-1)^{i-1}(\sqrt{L} - 1) \quad \text{for every } i \in \{1, \dots, \sqrt{L}\} \quad (1.2.5)$$

(we assume  $\sqrt{L} \in \mathbb{N}$  for conciseness) so that  $H(\tilde{\ell}) = \beta(\sqrt{L} - 1)^2 + \delta(\sqrt{L} - 1)$ . Second, those trajectories in  $\Omega_L$  that are performing a self-touching saturation are not many and therefore they do not carry any entropy. Third, we mentioned above that such saturated trajectories are made of  $o(L)$  stretches and a trajectory may touch the hard-wall at most once per vertical stretch (recall Fig. 1.1), hence their interactions with the wall cannot be numerous enough to contribute to the free energy. These three points are sufficient to understand why the free energy equals  $\beta$  in  $\mathcal{C}$  and thus, it is natural to define the *excess* free energy of the system as

$$\tilde{f}(\beta, \delta) = f(\beta, \delta) - \beta, \quad (1.2.6)$$

which allows us to define the *extended* and the *collapsed* phases as

$$\begin{aligned}\mathcal{C} &:= \{(\beta, \delta) \in \mathcal{Q} : \tilde{f}(\beta, \delta) = 0\}, \\ \mathcal{E} &:= \mathcal{Q} \setminus \mathcal{C} = \{(\beta, \delta) \in \mathcal{Q} : \tilde{f}(\beta, \delta) > 0\}.\end{aligned}\quad (1.2.7)$$

Before stating Theorem 1.2.2 below, we need to settle some notations. We define for any  $\beta > 0$  the following probability distribution on  $\mathbb{Z}$ :

$$\mathbf{P}_\beta(\cdot = k) = \frac{e^{-\frac{\beta}{2}|k|}}{c_\beta}, \quad c_\beta := \sum_{k \in \mathbb{Z}} e^{-\frac{\beta}{2}|k|} = \frac{1 + e^{-\beta/2}}{1 - e^{-\beta/2}}. \quad (1.2.8)$$

We consider a one-dimensional random walk  $X := (X_i)_{i \geq 0}$  starting from the origin and such that  $(X_{i+1} - X_i)_{i \geq 0}$  is an i.i.d. sequence of random variables with law  $\mathbf{P}_\beta$ . Then, we let  $h_\beta(\delta)$  be the free energy of the wetting model that consists of the random walk  $X$  constrained to remain non-negative and to finish on the  $x$ -axis, and pinned at the origin by an energetic factor  $\delta$ , i.e.,

$$h_\beta(\delta) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{i=1}^N 1_{\{X_i=0\}}} 1_{\{X \in B_N^{0,+}\}} \right], \quad (1.2.9)$$

where  $B_N^{0,+}$  is the set of non-negative trajectories of length  $N$  ending at 0 —more generally, define for all  $y \geq 0$ ,

$$B_n^{y,+} := \left\{ (X_k)_{k=1}^n \in \mathbb{Z}^n ; X_n = y, X_k \geq 0 \forall 1 \leq k \leq n \right\}. \quad (1.2.10)$$

An explicit formula for  $h_\beta(\delta)$  is given in Appendix 1.A (see (1.A.4)). We also define  $\Gamma_\beta := c_\beta e^{-\beta}$  which is decreasing in  $\beta$ , and  $\beta_c > 0$  the unique solution of the equation  $\Gamma_\beta = 1$ .

**Theorem 1.2.2.** *The boundary between the collapsed and the extended phase can be characterized explicitly, i.e.,*

$$\mathcal{C} = \{(\beta, \delta) \in \mathcal{Q} : \beta \geq \beta_c, \delta \leq \delta_c(\beta)\},$$

where, for every  $\beta \geq \beta_c$ , the quantity  $\delta_c(\beta)$  is the unique solution in  $\delta$  of

$$\log \Gamma_\beta + h_\beta(\delta) = 0, \quad (1.2.11)$$

which yields the following analytic expression,

$$\delta_c(\beta) = \log \left( \frac{\sinh(\beta) + \sqrt{\sinh(\beta)^2 + 1 - e^\beta}}{1 - e^{-\beta}} \right). \quad (1.2.12)$$

**Remark 1.2.3.** Note that our formula for the critical curve in (1.2.12) was already conjectured in [76, equation 19] or [104, equation 21] (both expressions coincide provided we set  $\kappa = e^{\delta_c(\beta)}$  and  $\tau = e^\beta$ ). In [76] and [104], the heuristics supporting this formula are based on some additional assumption and on a computation of the grand canonical partition function with the help of a transfer matrix.



## Discussion

Let us further explain the phenomenon behind the existence of a surface transition inside  $\mathcal{C}$  that physicists have conjectured (see e.g. [123, Fig. 2]). As mentioned above, a typical trajectory in the collapsed phase looks like a globule delimited by a lower envelope and an upper envelope (see their rigorous definition in (1.3.16)). For  $L \in \mathbb{N}$ ,  $\beta > \beta_c$  and  $\delta < \delta_c(\beta)$ , we will prove in Section 1.4.2 that the lower envelope of a trajectory sampled from  $\mathbf{P}_L^{\beta, \delta}$  behaves roughly as a random walk of length  $O(\sqrt{L})$ , constrained to remain non-negative, and pinned at the  $x$ -axis (hard wall) with intensity  $\delta$ . This leaves us with a wetting model whose critical point  $\tilde{\delta}_c(\beta)$  can be explicitly computed, see (1.3.6) (and it satisfies  $\tilde{\delta}_c(\beta) < \delta_c(\beta)$ ). Thus, when  $\delta \leq \tilde{\delta}_c(\beta)$  the lower-envelope touches the hard-wall only  $o(\sqrt{L})$  times, whereas when  $\delta > \tilde{\delta}_c(\beta)$  it remains localized along the hard wall and touches it  $O(\sqrt{L})$  times. As a consequence, this wetting transition of the lower envelope is not encoded in the excess free energy  $\tilde{f}(\beta, \delta)$  (which remains equal to 0 in  $\mathcal{C}$ ) simply because the number of contacts between the polymer (or equivalently its lower envelope) and the hard-wall is at most  $O(\sqrt{L})$ . To be more specific, we will see with Theorem 1.3.2 below that, in the exponential growth rate of  $Z_{L, \beta, \delta}^+$ ,  $\delta \in (\tilde{\delta}_c(\beta), \delta_c(\beta))$  only contributes to the second order term. We will further discuss the behavior of a typical lower envelope in  $\mathcal{C}$  after stating Theorem 1.3.2.

In order to display a qualitative picture of the phase diagram (see Fig. 1.2), let us end this discussion with a few words about the extended regime  $\mathcal{E}$ , where a typical trajectory of length  $L$  is expected to have an horizontal extension of order  $L$ . Physicists (see [76, Fig.2] or [104]) have conjectured that another critical curve  $\hat{\delta}_c : [0, \beta_c] \rightarrow \mathbb{R}^+$  divides  $\mathcal{E}$  into a Desorbed-Extended phase denoted by  $\mathcal{DE}$  and an Adsorbed-Extended phase denoted by  $\mathcal{AE}$  but there is so far no guess for what the value of  $\hat{\delta}_c(\beta)$  could be.

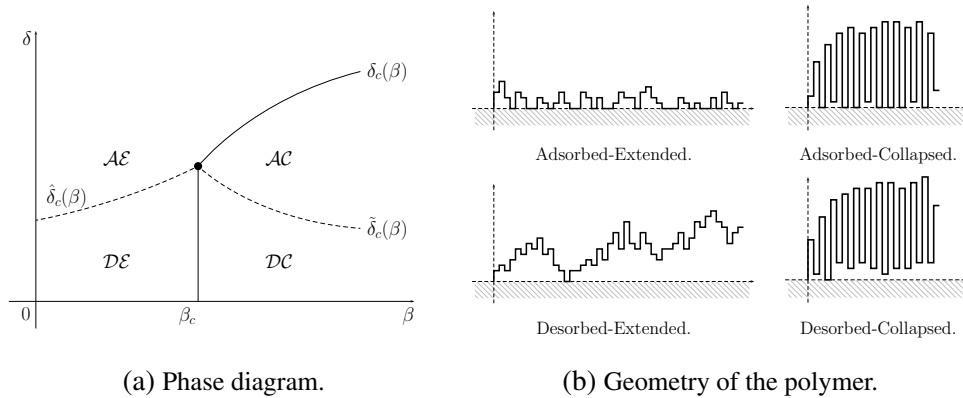


Figure 1.2: Qualitative picture of the phase diagram and the geometry of the polymer in each phase. In this paper, we rigorously determine the critical line  $\delta_c(\beta)$  between the extended and collapsed phases, and we characterize the surface transition critical line  $\tilde{\delta}_c(\beta)$ . No analytic expression has been conjectured for the critical line  $\hat{\delta}_c(\beta)$  yet.

## 1.3 Inside the collapsed phase: restriction to the single-bead model. Asymptotics of the partition functions

### 1.3.1 Bead decomposition of a trajectory

A trajectory of  $\Omega_L^+$  can be decomposed into a collection of sub-trajectories called beads. A bead is a succession of non-zero vertical stretches with alternating signs, which ends when two consecutive stretches have the same orientation, or when a stretch has length zero. To be more specific, we consider  $\ell \in \Omega_L^+$  and we recall that  $N_\ell$  is its horizontal extension and that by convention  $\ell_0 = \ell_{N_\ell+1} = 0$ . Then, we set  $x_0 = 0$  and for  $j \in \mathbb{N}$  such that  $x_{j-1} < N_\ell$  we set

$$x_j = \inf\{i \geq x_{j-1} + 1 : l_i \tilde{\wedge} l_{i+1} = 0\}$$

so that  $x_j$  is the index of the last vertical stretch composing the  $j$ -th bead of  $\ell$ . Finally, we let  $n(\ell)$  be the number of beads in  $\ell$ , in particular it satisfies  $x_{n(\ell)} = N_\ell$ . Thus, we can decompose any trajectory  $\ell \in \Omega_L$  into a succession of beads denoted by  $\mathcal{B}_j$  with  $j \in \{1, \dots, n(\ell)\}$  as follows

$$\ell = \odot_{j=1}^{n(\ell)} \mathcal{B}_j := \odot_{j=1}^{n(\ell)} (\ell_{x_{j-1}+1}, \dots, \ell_{x_j}), \quad (1.3.1)$$

where  $\odot$  denotes the concatenation.

A key issue concerning the collapsed phase of our model consists in showing that a typical trajectory contains a unique macroscopic bead outside which very few monomers are laying. Such result was derived for IPDSAW in [52] and recently improved in [121]. Its proof requires some sharp asymptotics of the partition function restricted to beads. Those estimates are a lot more technical in the model with a hard wall, since the shape of a bead strongly depends on its position with respect to the wall (i.e. its starting and ending heights). In the present paper, we consider the model restricted to describe a single-bead which begins and ends at the wall, see Section 1.3.2 below for a definition. Not only this restriction makes the model more tractable, we also conjecture that the unrestricted model, inside its collapsed phase, describes a unique macroscopic bead of the same nature as the one studied in this paper, i.e. beginning and ending in the vicinity of the wall.

### 1.3.2 Single-bead restriction of the model

Let  $L \in \mathbb{N}$ , and define  $\Omega_L^{\circ,+}$  the subset of  $\Omega_L^+$  gathering those trajectories  $\ell$  constrained to form only one “bead” —all its stretches are of non-zero length and alternate orientations— and to come back to the wall with its last stretch (in particular its horizontal extension  $N_\ell$  must be even). That is,

$$\Omega_L^{\circ,+} := \bigcup_{N \geq 1} \mathcal{L}_{N,L}^{\circ,+} := \bigcup_{N \geq 1} \left\{ \begin{array}{l} (\ell_i)_{i=1}^{2N} \in \mathbb{Z}^{2N} ; \sum_{i=1}^{2N} |\ell_i| = L - 2N, \sum_{i=1}^{2N} \ell_i = 0, \\ \sum_{i=1}^k \ell_i \geq 0, \forall k \leq 2N, \ell_i \ell_{i+1} < 0, \forall 1 \leq i < 2N \end{array} \right\}. \quad (1.3.2)$$

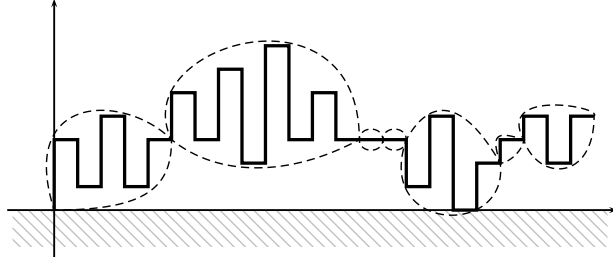


Figure 1.3: Decomposition of a trajectory into a succession of seven beads. For all  $\ell \in \Omega_L^+$  and  $1 \leq i \leq N_\ell$ , a new bead starts at the  $i$ -th vertical stretch if and only if  $\ell_{i-1}\ell_i \geq 0$  (where we recall  $\ell_0 = \ell_{N_\ell+1} = 0$ ).

The partition function restricted to such trajectories becomes:

$$Z_{L,\beta,\delta}^{\circ,+} := \sum_{N=1}^{L/2} \sum_{\substack{\ell \in \mathcal{L}_{N,L}^{\circ,+} \\ \ell_0 = \ell_{2N+1} = 0}} e^{H(\ell)}, \quad (1.3.3)$$

(recall that  $\ell_0 = \ell_{2N+1} = 0$  for notational convenience).

### 1.3.3 Asymptotics of the single-bead partition function inside the collapsed phase

The single-bead model undergoes the same phase transition as the full model, albeit its critical curve differs slightly. Let us define  $f^\circ$  and  $\tilde{f}^\circ$  respectively the free energy and excess free energy of the single-bead model,

$$f^\circ(\beta, \delta) := \lim_{L \rightarrow \infty} \log Z_{L,\beta,\delta}^{\circ,+} \quad \text{and} \quad \tilde{f}^\circ(\beta, \delta) := f^\circ(\beta, \delta) - \beta.$$

Since  $\tilde{\ell}$  forms a single bead (recall (1.2.5)), it follows that  $\tilde{f}^\circ(\beta, \delta) \geq 0$  for all  $(\beta, \delta) \in \mathcal{Q}$ .

**Proposition 1.3.1.** *For the single-bead model, we have*

$$\mathcal{C}_{\text{bead}} := \{(\beta, \delta) \in \mathcal{Q} : \tilde{f}^\circ(\beta, \delta) = 0\} = \{(\beta, \delta) \in \mathcal{Q} : \beta \geq \beta_c, \delta \leq \delta_c^\circ(\beta)\}, \quad (1.3.4)$$

where, for every  $\beta \geq \beta_c$ , the quantity  $\delta_c^\circ(\beta)$  is the unique solution in  $\delta$  of

$$2 \log \Gamma_\beta + h_\beta(\delta) = 0. \quad (1.3.5)$$

Notice that an analytic expression of  $\delta_c^\circ(\beta)$  can be derived from (1.3.5) and (1.A.4), similarly to (1.2.12) in Theorem 1.2.2. Let us now focus on the collapsed phase of the single-bead model. The surface transition occurs along a curve denoted by  $\tilde{\delta}_c : [\beta_c, \infty) \mapsto \mathbb{R}^+$

where  $\tilde{\delta}_c(\beta)$  turns out to be the critical point of the wetting model introduced in (1.2.9), that is for every  $\beta \geq \beta_c$ ,

$$\tilde{\delta}_c(\beta) := \inf\{\delta \geq 0: h_\beta(\delta) > 0\} = -\log(1 - e^{-\beta/2}). \quad (1.3.6)$$

The second identity in (1.3.6) is proven in Proposition 1.A.1. Definitions (1.3.5) and (1.3.6) ensure us that  $\tilde{\delta}_c(\beta) \leq \delta_c(\beta) \leq \delta_c^\circ(\beta)$  for  $\beta \geq \beta_c$ : thus, the curve  $\tilde{\delta}_c : [\beta_c, \infty) \mapsto \mathbb{R}^+$  lies in both  $\mathcal{C}$  and  $\mathcal{C}_{\text{bead}}$ . We will see in Theorem 1.3.2 below that the second order term in the exponential development of  $Z_{L,\beta,\delta}^{\circ,+}$ —which we may also call the *surface order*—loses its analyticity along that curve.

At this stage, we divide the collapsed phase  $\mathcal{C}_{\text{bead}}$  into a *desorbed collapsed* phase  $\mathcal{DC}$  and an *adsorbed collapsed* phase  $\mathcal{AC}$  defined as

$$\begin{aligned} \mathcal{DC} &:= \{(\beta, \delta) \in \mathcal{Q}: \beta \geq \beta_c, \delta \leq \tilde{\delta}_c(\beta)\}, \\ \mathcal{AC} &:= \{(\beta, \delta) \in \mathcal{Q}: \beta \geq \beta_c, \tilde{\delta}_c(\beta) \leq \delta \leq \delta_c^\circ(\beta)\}, \end{aligned} \quad (1.3.7)$$

where we dropped the subscript “bead” to lighten notations. To fully state Theorem 1.3.2, we need to introduce some definitions. Let  $\mathcal{L}$  be the logarithmic moment generating function of the distribution  $\mathbf{P}_\beta$  (recall (1.2.8)), that is for any  $|h| < \beta/2$ ,

$$\mathcal{L}(h) := \log \mathbf{E}_\beta[e^{hX_1}], \quad (1.3.8)$$

and for every  $\mathbf{h} = (h_0, h_1) \in \mathcal{D}_\beta := \{(h_0, h_1) \in \mathbb{R}^2, |h_1| < \beta/2, |h_0 + h_1| < \beta/2\}$ , define

$$\mathcal{L}_\Lambda(\mathbf{h}) := \int_0^1 \mathcal{L}(h_0x + h_1)dx, \quad (1.3.9)$$

which is convex on  $\mathcal{D}_\beta$ . In [52, Lemma 5.3], it is proven that  $\mathbf{h} \in \mathcal{D}_\beta \mapsto \nabla \mathcal{L}_\Lambda(\mathbf{h}) = (\partial_{h_0} \mathcal{L}_\Lambda, \partial_{h_1} \mathcal{L}_\Lambda)(\mathbf{h})$  is a  $\mathcal{C}^1$ -diffeomorphism from  $\mathcal{D}_\beta$  to  $\mathbb{R}^2$ , so let  $\tilde{\mathbf{h}} : \mathbb{R}^2 \rightarrow \mathcal{D}_\beta$  be its inverse. With those notations at hand, define

$$\varphi_{(\beta,\delta)}(a) := a \left( 2 \log \Gamma_\beta + h_\beta(\delta) - \frac{1}{2a^2} \tilde{h}_0 \left( \frac{1}{2a^2}, 0 \right) + \mathcal{L}_\Lambda \left( \tilde{\mathbf{h}} \left( \frac{1}{2a^2}, 0 \right) \right) \right), \quad a \in (0, \infty). \quad (1.3.10)$$

We will prove that  $\varphi_{(\beta,\delta)}$  is negative and strictly concave on  $(0, \infty)$  and reaches its maximum at some  $\tilde{a} = \tilde{a}(\beta, \delta) \in (0, \infty)$ . Then, set  $\Phi(\beta, \delta) := \varphi_{(\beta,\delta)}(\tilde{a}(\beta, \delta))$ , let  $\sigma_\beta^2$  be the variance of  $X_1$  under  $\mathbf{P}_\beta$ , and recall (1.3.5) and (1.3.6) for the definitions of  $\delta_c^\circ(\beta)$  and  $\tilde{\delta}_c(\beta)$  respectively. Finally, for any two sequences  $(a_L)_{L \geq 1}$ ,  $(b_L)_{L \geq 1}$ , we write  $a_L \asymp b_L$  if there exists  $C > 0$  and  $n_0 \in \mathbb{N}$  such that  $\frac{1}{C}a_n \leq b_n \leq Ca_n$  for all  $n \geq n_0$  (see also (5.0.1)).

**Theorem 1.3.2.** *Let  $\beta > \beta_c$ .*

(i) *For  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ , then*

$$Z_{L,\beta,\delta}^{\circ,+} \asymp \frac{1}{L^{3/4}} e^{\beta L + \Phi(\beta,\delta)\sqrt{L}}, \quad (1.3.11)$$

(ii) for  $\delta \in (0, \tilde{\delta}_c(\beta)]$  and  $\varepsilon > 0$ , there exist  $C > 0$  and  $L_0 \in \mathbb{N}$  such that for  $L \geq L_0$ ,

$$e^{\beta L + \Phi(\beta, 0)\sqrt{L} + (\Psi(\beta) - \varepsilon)L^{1/6}} \leq Z_{L, \beta, \delta}^{\circ, +} \leq \frac{C}{L^{3/4}} e^{\beta L + \Phi(\beta, 0)\sqrt{L}}, \quad (1.3.12)$$

(iii) for  $\delta = 0$ , then

$$Z_{L, \beta, \delta}^{\circ, +} = e^{\beta L + \Phi(\beta, 0)\sqrt{L} + \Psi(\beta)L^{1/6}(1+o(1))}, \quad \text{as } L \rightarrow \infty, \quad (1.3.13)$$

where

$$\Psi(\beta) := -|a_1| \left[ \frac{\tilde{a}(\beta, 0) \sigma_\beta^2}{2} \tilde{h}_0 \left( \frac{1}{2\tilde{a}^2(\beta, 0)}, 0 \right)^2 \right]^{1/3}$$

and  $a_1$  denotes the first zero (in absolute value) of the Airy function.

These estimate allow us to derive some properties of typical trajectories under the polymer measure, most notably regarding the number of contacts with the hard wall in the collapsed phase.

**Corollary 1.3.3.** *Let  $\beta > \beta_c$ .*

(i) *The function  $\delta \mapsto \Phi(\beta, \delta)$  is  $\mathcal{C}^1$  on  $(\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ . For  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$  and for any  $\varepsilon > 0$  we have*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L, \beta, \delta}^{\circ, +} \left( \sum_{k=1}^{N_\ell} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}} \in \left[ \partial_\delta \Phi(\beta, \delta) - \varepsilon, \partial_\delta \Phi(\beta, \delta) + \varepsilon \right] \sqrt{L} \right) = 1. \quad (1.3.14)$$

(ii) *For  $\delta \in [0, \tilde{\delta}_c(\beta))$ , there exist some  $K > 0$  (which only depends on  $\beta$ ) such that*

$$\lim_{L \rightarrow \infty} \mathbf{P}_{L, \beta, \delta}^{\circ, +} \left( \sum_{k=1}^{N_\ell} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}} \leq KL^{1/6} \right) = 1. \quad (1.3.15)$$

**Remark 1.3.4.** • The surface transition occurring along the curve  $\{(\beta, \tilde{\delta}_c(\beta)), \beta > \beta_c\}$  is proven by Theorem 1.3.2 (i)-(ii) (and confirmed by Corollary 1.3.3). Indeed,  $\delta > \tilde{\delta}_c(\beta)$  implies that  $h_\beta(\delta) > h_\beta(0) = 0$ , hence  $\varphi_{(\beta, \delta)}(a) > \varphi_{(\beta, 0)}(a)$  for  $a \in (0, \infty)$ , and  $\Phi(\beta, \delta) > \Phi(\beta, 0)$  (whereas  $h_\beta(\delta) = 0$  for all  $\delta \leq \tilde{\delta}_c(\beta)$ ).

• In Theorem 1.3.2 (ii), we conjecture that the upper bound is not optimal, and that (iii) should apply at least for every  $\delta < \tilde{\delta}_c(\beta)$ . Similarly, for Corollary 1.3.3 (ii), we expect the typical number of contacts with the hard wall to be of smaller order than  $L^{1/6}$ .

• In Theorem 1.3.2 (iii), obtaining  $\Psi(\beta)$  requires to compute the Laplace transform of the area enclosed by a Brownian meander of length  $T$ . The first zero of the Airy function  $a_1$  appears in the leading order of such Laplace transform as  $T \rightarrow \infty$  (see Section 1.4.4).

## Discussion

Let us give some insights into Theorem 1.3.2. For any trajectory  $\ell \in \Omega_L^{\circ,+}$  forming a single bead, we define its *lower envelope*  $I := (I_i)_{i=0}^{N_\ell/2}$  and *upper envelope*  $S := (S_i)_{i=0}^{N_\ell/2+1}$  as follow:

$$S_k = \sum_{i=1}^{2k-1} \ell_i, \quad k \in \{1, \dots, \frac{N_\ell}{2} + 1\}$$

$$I_k = \sum_{i=1}^{2k} \ell_i, \quad k \in \{1, \dots, \frac{N_\ell}{2}\}$$

and  $S_0 = I_0 = 0$ . The single-bead constraint ensures that  $S$  (resp.  $I$ ) describes the topmost (resp. bottommost) layer of the polymer (see Fig. 1.4).

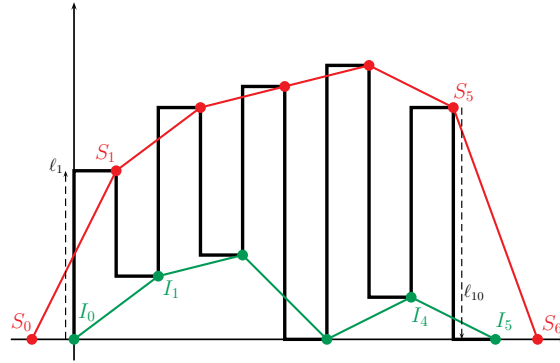


Figure 1.4: Representation of a single-bead trajectory, with its upper envelope (in red) and its lower envelope (in green) respectively described by the sequences  $S$  and  $I$ .

In Section 1.4.1 below, we will show that, under the polymer measure, the envelopes  $I$  and  $S$  of a given single-bead configuration may be sampled as trajectories of non-negative random walk bridges that are coupled via geometric constraints. In Section 1.4.2, we will break that geometric coupling and integrate over  $S$  so that  $I$  can be investigated on its own (see (1.4.12)). However, this comes with a cost (see Proposition 1.4.2 below), the law of  $I$  being perturbed by:

1. a wetting term  $\delta \sum_{k=1}^N 1_{\{I_k=0\}}$  which comes from the fact that the polymer (or equivalently its lower envelope) is adsorbed along the hard wall,
2. a pre-wetting term proportional to  $-A_N(I)/N$  where  $A_N(I)$  is the area below  $I$  (see (1.4.6)). The latter penalization comes from the fact that the upper envelope  $S$  must sweep an abnormally large area, i.e.,  $A_N(S) = A_N(I) + qN^2$  with  $q > 0$ . Therefore,  $S$ , which is already in a large deviation regime, pushes down the lower envelope  $I$  so as to keep  $A_N(I)$  small.

The influence of a pre-wetting term on a  $1 + 1$ -dimensional random walk constrained to remain positive has already been studied in [102] and [106] (see also [107] for a review).

Among physical motivations is e.g. the study of a liquid-gas interface when a thermodynamically stable gas is in contact with a substrate (hard-wall) that has a strong preference for the liquid phase, with the temperature decreasing to the liquid/gas critical point. From this point of view, the present paper displays a new example of a physical object (i.e., the lower envelope of a collapsed homopolymer interacting with a hard wall) associated with a model for which pre-wetting appears naturally.

As stated in Theorem 1.3.2, the pre-wetting term does not have an influence on the lower-envelope inside  $\mathcal{AC}$  since the pinning term is strong enough to keep the lower envelope at finite distance from the hard-wall. Inside  $\mathcal{DC}$ , in turn the pre-wetting term should dominate and we expect that [102, Theorem 1.2] also apply here, implying that  $I$  has fluctuations of order  $L^{1/6}$  (its length being  $\sqrt{L}$ ).

### Open problems

From a mathematical point of view, there are still many issues that remain to be settled concerning the present model. Let us list a few here.

1. Consider the unrestricted model (defined in (1.2.1)–(1.2.4)) and prove that inside its collapsed phase  $\mathcal{C}$  it undergoes the same surface transition as the single-bead model. We conjecture that it takes place along the very same critical curve  $\beta \mapsto \tilde{\delta}_c(\beta)$  than in the single-bead model (even though the phases  $\mathcal{C}$  and  $\mathcal{C}_{\text{bead}}$  differ slightly). However, this is actually a very intricate matter to investigate, since the bead decomposition requires us to consider the heights of starting and ending points of each bead in addition to their length. Moreover, our estimates for the single-bead model rely noticeably on the observation that the area enclosed by the lower envelope is not too large ( $A_N(I) = O(N^{3/2})$ ). This observation fails for beads away from the wall, implying in particular that their variational problem (1.3.10)—hence the surface term in the logarithmic development of their partition function— will differ.
2. For the unrestricted model again, prove that a typical trajectory is made of a unique macroscopic bead, and give a bound on the number of monomers that may lay outside this bead. A similar result was proven in [52, Theorem C] and [121] for the model without wall, and comforts us in the idea that, in the Collapsed phase, the single-bead model encapsulates quite accurately the behavior of the unrestricted model.
3. Consider a random walk  $(X_i)_{i=0}^N$  (recall (1.2.8)) constrained to remain non-negative and perturbed by both a wetting and a pre-wetting terms of parameter  $\delta$  and  $\gamma$  respectively, i.e.,

$$\mathbf{E}_\beta \left[ e^{\delta \sum_{i=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} e^{-\frac{\gamma}{N} A_N(X)} \right]. \quad (1.3.16)$$

Display sharp asymptotics for the partition function in (1.3.16) when  $\delta$  is not larger than the critical point of the pure wetting model (i.e.,  $\delta \leq \tilde{\delta}_c(\beta)$ ). Such estimates would be the key to improve Theorem 1.3.2 and Corollary 1.3.3 in  $\mathcal{DC}$ .

4. Provide a characterization of the critical curve dividing  $\mathcal{E}$  into an Adsorbed-Extended phase and a Desorbed-Extended phase. In the Extended regime, the bead decomposi-

tion and upper/lower envelopes description are far less efficient, so a whole different approach may be needed.

## 1.4 Proof of Theorem 1.3.2

We divide the proof of Theorem 1.3.2 into 6 steps. First, we adapt the random-walk representation of IPDSAW initially introduced in [130] to the present one-bead model. We derive a probabilistic representation of the partition function by rewriting, for every  $N \leq L/2$ , the contribution to the partition function of those trajectories made of  $2N$  stretches, in terms of two auxiliary random walks  $S$  and  $I$ . One particularity comes from the fact that  $I$  and  $S$  are coupled since the area enclosed in-between  $S$  and  $I$  is imposed by the length of the polymer, and another one comes from the one-bead constraint which implies that they cannot cross trajectories: hence  $S$  (resp.  $I$ ) will play the role of the *upper envelope* (resp. *lower envelope*) of the polymer. The second step consists in breaking the geometric coupling between  $I$  and  $S$ . This is achieved by integrating over  $S$ , and transforming the area constraint into an exponential perturbation of the law of  $I$ . In the third and fourth steps we estimate the partition function of the lower envelope  $I$ , in  $\mathcal{AC}$  and  $\mathcal{DC}$  respectively. Finally the fifth step proves that inside the collapsed phase, the horizontal extension of a typical trajectory is of order  $\sqrt{L}$ , and the sixth step collects all those estimates to prove Theorem 1.3.2.

### 1.4.1 Step 1: random-walk representation

The understanding of IPDSAW (recall Remark 1.2.1) has recently been improved (see [54] for a review). The key tool was a new probabilistic representation of the partition function based on an auxiliary random walk conditioned to enclose a prescribed area. It turns out that this representation is of substantial help for the present model as well but under a different form. This is the object of the present step.

Let us first provide a probabilistic description of the one-bead partition function  $Z_{L,\beta,\delta}^{\circ,+}$ . We recall (1.2.2) and (1.3.3) and we observe that

$$x \tilde{\wedge} y = \frac{|x| + |y| - |x + y|}{2}, \quad x, y \in \mathbb{Z}. \quad (1.4.1)$$

Hence the one-bead partition function can be written as

$$Z_{L,\beta,\delta}^{\circ,+} = e^{\beta L} \sum_{N=1}^{L/2} e^{-2\beta N} \sum_{\substack{\ell \in \mathcal{L}_{N,L}^{\circ,+} \\ \ell_0 = \ell_{2N+1} = 0}} e^{-\frac{\beta}{2} \sum_{i=0}^{2N} |\ell_i + \ell_{i+1}|} e^{\delta \sum_{k=1}^{2N} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}}}. \quad (1.4.2)$$

At this stage, recall the definition of  $\mathbf{P}_\beta$  in (1.2.8). We consider two independent random walks  $S := (S_i)_{i \geq 0}$  and  $I := (I_i)_{i \geq 0}$  starting from 0 and such that  $(S_{i+1} - S_i)_{i \geq 0}$  and



$(I_{i+1} - I_i)_{i \geq 0}$  are i.i.d. sequences of random variables of law  $\mathbf{P}_\beta$ . We notice that for every  $\ell \in \mathcal{L}_{N,L}^{\circ,+}$  (with  $\ell_0 = \ell_{2N+1} = 0$ ) the first factor in the second sum in (1.4.2) satisfies

$$e^{-\frac{\beta}{2} \sum_{i=0}^{2N} |\ell_i + \ell_{i+1}|} = c_\beta^{2N+1} \mathbf{P}_\beta \left( S_k = \sum_{i=0}^{2k-1} \ell_i, \forall k \leq N+1 \right) \mathbf{P}_\beta \left( I_k = \sum_{i=0}^{2k} \ell_i, \forall k \leq N \right). \quad (1.4.3)$$

Let us introduce a one-to-one correspondence between trajectories  $(\ell_i)_{i=0}^{2N+1} \in \mathbb{Z}^{2N+2}$  with  $\ell_0 = \ell_{2N+1} = 0$ , and  $(S_i)_{i=0}^{N+1} \in \mathbb{Z}^{N+2}$ ,  $(I_i)_{i=0}^N \in \mathbb{Z}^{N+1}$  with  $S_0 = I_0 = 0$  and  $S_{N+1} = I_N$ , by letting  $S_k = \sum_{i=0}^{2k-1} \ell_i, \forall 1 \leq k \leq N+1$  and  $I_k = \sum_{i=0}^{2k} \ell_i, \forall 1 \leq k \leq N$ . Then, constraints on  $\ell \in \mathcal{L}_{N,L}^{\circ,+}$  can be transcribed to  $S$  and  $I$  (recall (1.3.2)). Indeed,  $\sum_{i=1}^{2N} \ell_i = 0$  is equivalent to  $S_{N+1} = I_N = 0$ , and  $\sum_{i=1}^k \ell_i \geq 0, \forall k \leq 2N$  is equivalent to  $S_i \geq 0, \forall i \leq N+1$  and  $I_i \geq 0, \forall i \leq N$ . Besides, we can write

$$\sum_{i=1}^{2N} |\ell_i| = \sum_{k=1}^N |I_k - S_k| + \sum_{k=1}^{N+1} |S_k - I_{k-1}| =: G(S, I), \quad (1.4.4)$$

which is the “geometric area” between  $S$  and  $I$ , and it is constrained to be equal to  $L - 2N$  in  $\mathcal{L}_{N,L}^{\circ,+}$ . Finally,  $\ell_i \ell_{i+1} < 0, \forall 1 \leq i < 2N$  is equivalent to  $S \succ I$  where we define

$$\{S \succ I\} := \{S_k > \max(I_k, I_{k-1}), \forall 1 \leq k \leq N\}. \quad (1.4.5)$$

In particular this means that in the one-bead model,  $S$  remains above  $I$  (thereby we respectively call  $S$  and  $I$  the *upper* and *lower envelopes* of the polymer—recall Fig. 1.4), and it implies that we can rewrite the geometric area  $G(S, I)$ : defining the *signed area* as

$$A_n(X) := \sum_{i=0}^n X_n, \quad (1.4.6)$$

for any  $n \in \mathbb{N}$  and  $(X_k)_{k=0}^n \in \mathbb{Z}^{n+1}$ , we then have that  $\{S \succ I\}$  and  $S_{N+1} = I_N = 0$  imply  $G(S, I) = 2(A_{N+1}(S) - A_N(I))$ .

Going back to (1.4.2) and plugging (1.4.3) in, those observations prove that we can rewrite the partition function as follows. Let  $\mathbf{P}_{\beta,x}$  denote the law of a random walk on  $\mathbb{Z}$  starting from  $x$  with increments distributed as  $\mathbf{P}_\beta$  (henceforward we will omit the  $x$ -dependence in  $\mathbf{P}_{\beta,x}$  when  $x$  is clear from context). Recall that  $\Gamma_\beta = c_\beta e^{-\beta}$ , and  $B_N^{\circ,+}$  is defined in (1.2.10).

**Proposition 1.4.1.** *For  $\beta > 0, \delta \geq 0$  and  $L \geq 1$ ,*

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+} := \frac{1}{c_\beta e^{\beta L}} Z_{L,\beta,\delta}^{\circ,+} = \sum_{N=1}^{L/2} (\Gamma_\beta)^{2N} D_{N,q_N}^{\circ}, \quad (1.4.7)$$

with  $q_N^L := (L - 2N)/2N^2$ , and for  $q \in (0, \infty) \cap \frac{\mathbb{N}}{2N^2}$ ,

$$D_{N,q}^{\circ} := \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{S \in B_{N+1}^{\circ,+}, I \in B_N^{\circ,+}\}} \mathbf{1}_{\{S \succ I\}} \mathbf{1}_{\{A_{N+1}(S) = A_N(I) + qN^2\}} \right]. \quad (1.4.8)$$

where  $S, I$  are two independent random walks distributed as  $\mathbf{P}_{\beta,0}$ .

Our aim is to provide sharp estimates on the partition function  $Z_{L,\beta,\delta}^{\circ,+}$ . To that purpose, we need to estimate  $D_{N,q}^{\circ}$  uniformly in  $q \in [q_1, q_2] \cap \frac{1}{2N^2} \mathbb{N}$  for any  $0 < q_1 < q_2$  fixed.

### 1.4.2 Step 2: integrating over $S$ in $D_{N,q}^\circ$

With Proposition 1.4.2 below, we show that the geometric constraint imposed to  $I$  and  $S$  can be relaxed provided we introduce an exponential perturbation to the law of  $I$ , thus breaking the coupling between  $I$  and  $S$ .

Define  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  by

$$g(q, p) := \tilde{\mathbf{h}}^{q,p} \cdot (q, p) - \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q,p}), \quad (1.4.9)$$

where  $\mathcal{D}_\beta \subset \mathbb{R}^2$ ,  $\tilde{\mathbf{h}}^{q,p} : \mathbb{R}^2 \rightarrow \mathcal{D}_\beta$  and  $\mathcal{L}_\Lambda : \mathcal{D}_\beta \rightarrow \mathbb{R}$  are all defined in (1.3.9) —from now on we will note  $\tilde{\mathbf{h}}^{q,p} := \tilde{\mathbf{h}}(\cdot, \cdot)$  to tighten notations. We will see that  $g$  is the rate function of a large deviation principle associated to the couple  $\Lambda_N := (A_N(X)/N, X_N)$  under  $\mathbf{P}_\beta$  (see the proof of Proposition 1.4.3). Using [101, Lemma V.4], it follows that  $g$  is convex. Moreover  $\tilde{\mathbf{h}}$  and  $\mathcal{L}_\Lambda$  are  $\mathcal{C}^1$  functions and

$$\nabla g(q, p) := (\partial_q g, \partial_p g)(q, p) = \tilde{\mathbf{h}}^{q,p}, \quad (1.4.10)$$

in particular  $g$  is  $\mathcal{C}^2$ . Finally  $\nabla g(0, 0) = \tilde{\mathbf{h}}^{0,0} = (0, 0)$ , so  $g(q, p) \geq g(0, 0) = 0$  for all  $(q, p) \in \mathbb{R}^2$ .

**Proposition 1.4.2.** *Fix some  $q_2 > q_1 > 0$ . Then*

$$D_{N,q}^\circ \asymp \frac{e^{-Ng(q,0)}}{N^2} E_{N,q}^\circ, \quad (1.4.11)$$

uniformly in  $q \in [q_1, q_2] \cap \frac{1}{2N^2}\mathbb{N}$ , where

$$E_{N,q}^\circ := \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} e^{-\partial_q g(q,0) \frac{A_N(I)}{N}} \right]. \quad (1.4.12)$$

#### Outline of the proof of Proposition 1.4.2

Obtaining an upper bound on  $D_{N,q}^\circ$  is rather straightforward: we will simply remove the events  $\{S \succ I\}$  and  $\{S_i \geq 0, \forall 0 \leq i \leq N+1\}$  from the definition of  $D_{N,q}^\circ$  and then apply large deviation estimate for the signed area  $A_{N+1}(S)$  of the upper envelope. The proof of the lower bound is more involved and will consist in

- (a) getting rid of the event  $\{S \succ I\}$  by first introducing a stronger constraint, i.e.,  $S$  (resp.  $I$ ) remains above (resp. below) some deterministic curves.
- (b) conditioning on  $I$  that remains below the deterministic curve introduced in (a) and integrating over  $S$ . Then, with the help of large deviation estimates for random walks enclosing very large area, proving that the events  $\{S \in B_{N+1}^{0,+}\}$  and  $\{A_{N+1}(S) = qN^2 + A_N(I)\}$  have a probability  $\frac{1}{N^2} e^{-Ng(q,0)} e^{-\partial_q g(q,0) \frac{A_N(I)}{N}}$  and that the cost for  $S$  to remain above the deterministic curve introduced in (a) is bounded from below by a constant,
- (c) getting rid of the additional constraint on  $I$  introduced in (a) with an FKG inequality, by observing that the random walk  $I$  whose law is penalized exponentially by  $-\frac{1}{N}A_N(I)$  typically remains below its unpenalized counterpart.

**Proof of the lower bound in (1.4.11)**
**Integrating on the upper envelope  $S$** 

To begin with, we constrain the upper and lower envelopes to remain respectively above and below some fixed curves. In particular this allows us to handle the condition  $S \succ I$ . Define  $f(t) := \gamma(t \wedge (1 - t))$ ,  $t \in [0, 1]$ , and

$$\begin{aligned} \tilde{f}_S(k) &:= (N + 1)f\left(\frac{k}{N+1}\right) + K_S, & \forall 1 \leq k \leq N, \\ \tilde{f}_I(k) &:= Nf\left(\frac{k}{N}\right) + K_I, & \forall 1 \leq k \leq N - 1, \end{aligned} \quad (1.4.13)$$

where  $\gamma, K_S, K_I > 0$  are constants. If we constrain  $S_k$ ,  $1 \leq k \leq N$  (resp.  $I_k$ ,  $1 \leq k \leq N - 1$ ) to remain above  $\tilde{f}_S(k)$  (resp. below  $\tilde{f}_I(k)$ ), and provided that  $K_S - K_I \geq 1 + \gamma$ , then we have

$$\{S \succ I\} \supset \left\{ \begin{array}{l} S_k \geq \tilde{f}_S(k), \forall 1 \leq k \leq N; \\ I_k \leq \tilde{f}_I(k), \forall 1 \leq k \leq N - 1 \end{array} \right\}. \quad (1.4.14)$$

Geometrically,  $\tilde{f}_S(k)$  is a piecewise linear curve above the wall of order  $N$  when  $k = N/2$ , and constant at its ends. The constants  $K_S, K_I$  are only there for technical purposes and are mostly irrelevant when  $N$  is large (see Figure 1.5). More precisely, we will fix  $\gamma$  such that Proposition 1.4.3 below applies (for any  $K_S > 0$ ), then we fix  $K_I$  when applying Lemma 1.4.5, and we finally fix  $K_S > K_I$  such that (1.4.14) holds.

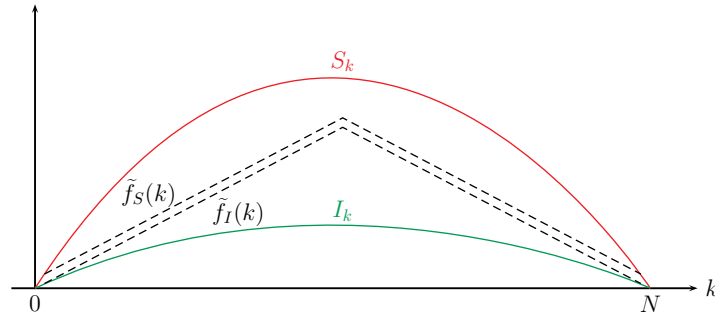


Figure 1.5: Representation of the constraints on  $S, I$ , for  $N$  large. The piecewise linear curves  $\tilde{f}_S$  and  $\tilde{f}_I$  are above the wall, reaching a height of order  $N$  in the middle and bounded at both ends.  $S$  is constrained to remain above  $\tilde{f}_S$ , and  $I$  below  $\tilde{f}_I$ .

We also add the constraint  $\{A_N(I) \leq C_A N^{3/2}\}$ , where  $C_A > 0$  is some constant which will also be fixed when applying Lemma 1.4.5 below. Let us condition  $D_{N,q}^\circ$  over the trajectory  $I$  (recall (1.4.8)), so that we separate the constraints on  $I$  and  $S$  and we obtain the lower bound

$$D_{N,q}^\circ \geq \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \mathbf{1}_{\{A_N(I) \leq C_A N^{3/2}\}} \mathbf{1}_{\{I_k \leq \tilde{f}_I(k), \forall 1 \leq k < N\}} \widehat{D}_{N,q}^\circ(I) \right], \quad (1.4.15)$$

with:

$$\widehat{D}_{N,q}^\circ(I) := \mathbf{E}_{\beta,0} \left[ \mathbf{1}_{\{S_{N+1}=0\}} \mathbf{1}_{\{A_{N+1}(S)=A_N(I)+qN^2\}} \mathbf{1}_{\{S_k \geq \tilde{f}(k), \forall 1 \leq k < N+1\}} \middle| I \right]. \quad (1.4.16)$$

Let us fix  $I = (I_k)_{k=1}^N$  which satisfies all constraints in (1.4.15). Proposition 1.4.3 below allows us to estimate  $\widehat{D}_{N,q}^\circ(I)$  up to uniform constant factors. We postpone the proof to Section 1.6.3.

**Proposition 1.4.3.** *Let  $\varepsilon > 0$  and  $q_1 < q_2, p_1 < p_2 \in \mathbb{R}$ . There exists  $\gamma > 0$  such that for all  $K_S > 0$ , there exist  $C_2 > C_1 > 0$  and  $N_0 \in \mathbb{N}$  such that for every  $N \geq N_0$ , one has*

$$\begin{aligned} \frac{C_1}{N^2} e^{-Ng(q,p)} &\leq \mathbf{E}_{\beta,0} \left[ \mathbf{1}_{\{X_N=pN, A_N(X)=qN^2\}} \mathbf{1}_{\{X_k \geq \tilde{f}(k), \forall 1 \leq k < N\}} \right] \\ &\leq \mathbf{E}_{\beta,0} \left[ \mathbf{1}_{\{X_N=pN, A_N(X)=qN^2\}} \right] \leq \frac{C_2}{N^2} e^{-Ng(q,p)}, \end{aligned} \quad (1.4.17)$$

where we define  $\tilde{f}(k) := Nf(\frac{k}{N}) + kp + K_S$ , and these bounds hold uniformly in  $q \in [q_1, q_2] \cap \frac{\mathbb{Z}}{2N^2}$ ,  $p \in [p_1, p_2] \cap \frac{\mathbb{Z}}{N}$  satisfying  $p \leq 2q - \varepsilon$ .

**Remark 1.4.4.** (i) Although we will only use Proposition 1.4.3 with  $p = 0$ , we prove it in the general case ( $p \neq 0$ ) because such estimates will be useful when studying the model without the single-bead restriction.

(ii) The assumption  $q \geq \frac{p+\varepsilon}{2}$  ensures that the trajectory has to enclose a large area, so it is prompted to draw a concave shape above the straight line from  $(0, 0)$  to  $(N, pN)$ ; hence it remains above the curve  $\tilde{f}$  provided that  $\gamma$  is sufficiently small. Notice that the assumption  $q = \frac{p}{2}$  would be satisfied by a linear trajectory from  $(0, 0)$  to  $(N, pN)$ , and a lower value of  $q$  would force the trajectory to be convex (below the linear trajectory).

(iii) To prove Proposition 1.4.3, we first recall from [52] an estimate of the expectation without the constraint  $\{X_k \geq \tilde{f}(k), \forall 1 \leq k < N\}$ , then we prove that this constraint only costs up to a constant factor. In particular this strongly reinforces [52, Prop. 2.5], since it yields that there exists  $C_3 > 0$  such that uniformly in  $N \geq N_0$  and in  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$  with  $q_1 > 0$ ,

$$\mathbf{P}_{\beta,0}(X_i > 0, 0 < i < N \mid A_N = qN^2, X_N = 0) \geq C_3.$$

Recall (1.4.16), and let  $p' = 0$ ,  $q' := (qN^2 + A_N(I))/(N+1)^2$ . Under our assumptions, there is a compact subset  $[q'_1, q'_2] \subset (0, \infty)$  and  $N_0 \in \mathbb{N}$  such that  $q' \in [q'_1, q'_2]$  for all  $N \geq N_0$ ,  $q \in [q_1, q_2]$  and  $I$  satisfying all constraints from (1.4.15) (recall that  $A_N(I) = O(N^{3/2})$ ). Moreover we have  $0 = p' < q'_1/2$ , hence we can apply Proposition 1.4.3 to  $\widehat{D}_{N,q}^\circ(I)$ , and we obtain the uniform bound for  $N \geq N_0$ ,

$$\widehat{D}_{N,q}^\circ(I) \geq \frac{C_3}{N^2} e^{-Ng(q',0)}. \quad (1.4.18)$$

Now recall that  $g$  is  $\mathcal{C}^2$  and convex, so there exists  $C_4 > 0$  such that

$$g(q', 0) \leq g(q, 0) + \partial_q g(q, 0)(q' - q) + C_4(q' - q)^2,$$

uniformly in  $q \in [q_1, q_2]$  and  $q' \in [q'_1, q'_2]$ . Recall that  $A_N(I) = O(N^{3/2})$ , and notice  $q' - q = \frac{A_N(I)}{N^2} + O(1/N)$ . Thereby (1.4.18) becomes

$$\widehat{D}_{N,q}^\circ(I) \geq \frac{C_5}{N^2} e^{-Ng(q,0) - \partial_q g(q,0) \frac{A_N(I)}{N}}, \quad (1.4.19)$$

where  $C_5$  is uniform in  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$  and  $I$ . Plugging (1.4.19) into (1.4.15) we obtain

$$\begin{aligned} D_{N,q}^\circ &\geq \frac{C_5}{N^2} e^{-Ng(q,0)} \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \mathbf{1}_{\{A_N(I) \leq C_A N^{3/2}\}} \right. \\ &\quad \left. \times \mathbf{1}_{\{I_k \leq \tilde{f}_I(k), \forall 1 \leq k < N\}} e^{-\partial_q g(q,0) \frac{A_N(I)}{N}} \right]. \end{aligned} \quad (1.4.20)$$

Recall that  $g$  is non-negative, and let us point out that  $\partial_q g(q, 0) = \tilde{h}_0^{q,0} > 0$  for all  $q > 0$  (recall (1.4.10), and see [52, Rem. 5.5] or Section 1.6.3).

### Relaxing some constraints on the lower envelope $I$

Our goal now is to drop the events  $\{I_k \leq \tilde{f}_I(k), \forall 1 \leq k < N\}$  and  $\{A_N(I) \leq C_A N^{3/2}\}$  in the r.h.s. in (1.4.20) by paying up to a constant factor uniform in  $N \geq N_0$  and  $q \in [q_1, q_2]$ . To that aim we use an FKG inequality. Indeed, notice that the functions

$$\begin{aligned} x = (x_k)_{k=1}^N &\longmapsto e^{\delta \sum_{k=1}^N \mathbf{1}_{\{x_k=0\}}} e^{-\partial_q g(q,0) \frac{A_N(x)}{N}}, \\ \text{and} \quad x &\longmapsto \mathbf{1}_{\{x_k \leq \tilde{f}_I(k), \forall 1 \leq k < N\}} \mathbf{1}_{\{A_N(x) \leq C_A N^{3/2}\}}, \end{aligned}$$

are both bounded and *non-increasing* on  $B_N^{0,+}$ , where we say that  $f : B_N^{0,+} \rightarrow \mathbb{R}$  is non-increasing if for all  $x, y \in B_N^{0,+}$  such that  $x_k \leq y_k, \forall 1 \leq k \leq N$ , one has  $f(x) \geq f(y)$ . Thereby the FKG inequality claimed in Proposition 1.B.1 yields that

$$\begin{aligned} &\mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{X_k=0\}}} \mathbf{1}_{\{X_k \leq \tilde{f}_I(k), \forall 1 \leq k < N; A_N(X) \leq C_A N^{3/2}\}} e^{-\partial_q g(q,0) \frac{A_N(X)}{N}} \middle| X \in B_N^{0,+} \right] \\ &\geq \mathbf{P}_{\beta,0} \left( X_k \leq \tilde{f}_I(k), \forall 1 \leq k < N; A_N(X) \leq C_A N^{3/2} \middle| X \in B_N^{0,+} \right) \\ &\quad \times \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{X_k=0\}}} e^{-\partial_q g(q,0) \frac{A_N(X)}{N}} \middle| X \in B_N^{0,+} \right]. \end{aligned} \quad (1.4.21)$$

Finally, the first factor of the r.h.s. is bounded from below by some constant (close to 1) by the following lemma, which is proven in Section 1.6.1.

**Lemma 1.4.5.** *Let  $\varepsilon, \gamma > 0$ . There exist  $C_A, K_I > 0$  and  $N_0 \in \mathbb{N}$  such that*

$$\mathbf{P}_{\beta,0} \left( X_k \leq \tilde{f}_I(k), \forall 1 \leq k < N; A_N(X) \leq C_A N^{3/2} \middle| X \in B_N^{0,+} \right) \geq 1 - \varepsilon, \quad (1.4.22)$$

for all  $N \geq N_0$ .

By recollecting (1.4.20) and (1.4.21), and provided that  $C_A, K_I$  are sufficiently large to apply Lemma 1.4.5, we conclude that for every  $N \geq N_0$  and every  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$  we have

$$D_{N,q}^\circ \geq \frac{C_6}{N^2} e^{-Ng(q,0)} \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} e^{-\partial_q g(q,0) \frac{A_N(I)}{N}} \right], \quad (1.4.23)$$

which completes the proof of the lower bound.

### Proof of the upper bound in (1.4.11)

We recall (1.4.8) and we bound  $D_{N,q}^\circ$  from above by relaxing partially the constraints on  $S$  -more precisely the constraints  $\{S \succ I\}$  and  $\{S_i \geq 0, \forall 1 \leq i \leq N\}$ . Therefore,

$$\begin{aligned} D_{N,q}^\circ &\leq \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \mathbf{1}_{\{S_{N+1}=0\}} \mathbf{1}_{\{A_{N+1}(S)=A_N(I)+qN^2\}} \right], \\ &= \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \widetilde{D}_{N,q}^\circ(I) \right]. \end{aligned} \quad (1.4.24)$$

with

$$\widetilde{D}_{N,q}^\circ(I) := \mathbf{P}_{\beta,0} \left[ S_{N+1} = 0, A_{N+1}(S) = A_N(I) + qN^2 \mid I \right].$$

In order to get uniform bounds on  $\widetilde{D}_{N,q}^\circ(I)$  with Proposition 1.4.3, we need to drop those  $I$  sweeping a too large area. To that aim, for any  $c > 0$  we rewrite (1.4.24) as  $\widetilde{D}_{N,q}^\circ \leq R_{N,q}^1(c) + R_{N,q}^2(c)$  with

$$\begin{aligned} R_{N,q}^1(c) &:= \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}, A_N(I) \leq cN^2\}} \widetilde{D}_{N,q}^\circ(I) \right], \\ R_{N,q}^2(c) &:= \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}, A_N(I) > cN^2\}} \widetilde{D}_{N,q}^\circ(I) \right], \end{aligned}$$

and we write the very crude bound

$$\begin{aligned} R_{N,q}^2(c) &\leq e^{\delta N} \mathbf{P}_{\beta,0} \left[ A_N(I) > cN^2 \right] \leq e^{\delta N} \mathbf{P}_{\beta,0} \left[ \max_{1 \leq i \leq N} I_i > cN \right] \\ &\leq e^{\delta N} \mathbf{P}_{\beta,0} \left[ \sum_{i=1}^N |I_i - I_{i-1}| > cN \right]. \end{aligned} \quad (1.4.25)$$

Then, we set  $M := 1 + \max\{g(q,0), q \in [q_1, q_2]\}$  and we use the fact that  $(I_i - I_{i-1})_{i \geq 1}$  are i.i.d. with finite small exponential moments to conclude (via a Markov exponential inequality) that there exists a  $c_M > 0$  such that for every  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$ ,

$$R_{N,q}^2(c_M) \leq e^{-MN}. \quad (1.4.26)$$

Let us now consider  $R_{N,q}^1(c_M)$  and use Proposition 1.4.3 to assert that there exists  $C_7 > 0$  and  $N_0 \in \mathbb{N}$  (depending on  $M$ ) such that for every  $N \geq N_0$ ,  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$  and  $I$  satisfying  $A_N(I) \leq c_M N^2$  we have

$$\widetilde{D}_{N,q}^\circ(I) \leq \frac{C_7}{N^2} e^{-Ng(q',0)},$$

with  $q' = (qN^2 + A_N(I))/(N+1)^2$ . Recall that  $g$  is convex, therefore we can bound  $g(q', 0)$  from below with

$$g(q', 0) \geq g(q, 0) + \partial_q g(q, 0)(q' - q).$$

Moreover  $q' - q = \frac{A_N(I)}{N^2} + O(1/N)$ , thus,

$$\begin{aligned} R_{N,q}^1(c_M) &\leq \frac{C_8}{N^2} e^{-g(q,0)N} \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I \in B_N^{0,+}, A_N(I) \leq c_M N^2\}} e^{-\partial_q g(q,0) \frac{A_N(I)}{N}} \right] \\ &\leq \frac{C_8}{N^2} e^{-g(q,0)N} E_{N,q}^\circ, \end{aligned} \quad (1.4.27)$$

where we recall (1.4.12).

Our proof will be complete once we show that for  $N$  large enough and for every  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$ , the r.h.s. in (1.4.26) (i.e.,  $e^{-MN}$ ) is not larger than the r.h.s. in (1.4.27). To that aim, we write the lower bound

$$E_{N,q}^\circ \geq \mathbf{P}_{\beta,0}(A_N(I) \leq C_A N^{3/2} \mid I \in B_N^{0,+}) \mathbf{P}_{\beta,0}(I \in B_N^{0,+}) e^{-\max\{\partial_q g(q,0), q \in [q_1, q_2]\} C_A \sqrt{N}}, \quad (1.4.28)$$

for any  $C_A > 0$ , where we constrained the walk to have area at most  $C_A N^{3/2}$ . A direct consequence of Lemma 1.4.5 is that there exist  $C_A, C_9 > 0$  such that for  $N$  large enough,

$$\mathbf{P}_{\beta,0}(A_N(I) \leq C_A N^{3/2} \mid I \in B_N^{0,+}) \geq C_9,$$

(more generally this follows from an invariance principle on non-negative random walk bridges, see (1.6.2) below). The second factor in (1.4.28) is bounded from below by  $C_{10} N^{-3/2}$  for some  $C_{10}$  (see (1.6.1)). Recalling that  $q \mapsto \partial_q g(q, 0)$  is continuous on  $[a_1, a_2]$ , we conclude that there exists  $C_{11}, c_1 > 0$  such that for  $N$  large enough and every  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{2N^2}$ , we have

$$E_{N,q}^\circ \geq \frac{C_{11}}{N^{3/2}} e^{-c_1 \sqrt{N}}. \quad (1.4.29)$$

Recalling that  $M = 1 + \max\{g(q, 0), q \in [q_1, q_2]\} > 0$ , it suffices to combine (1.4.26) with (1.4.27) and (1.4.29) to complete the proof of the upper bound in (1.4.11).

### 1.4.3 Step 3 : area-penalized wetting model for $\delta > \tilde{\delta}_c(\beta)$

In this section we give estimates on  $E_{N,q}^\circ$  when  $\delta > \tilde{\delta}_c(\beta)$  —in particular in the phase  $\mathcal{AC}$ . Notice that  $E_{N,q}^\circ$  is the partition function of a wetting model with an additional pre-wetting term. Let  $Z_{\text{wet},N}^{\beta,\delta}$  be the standard wetting partition function (without pre-wetting, see (1.A.1)). Such wetting models have already been studied extensively by mathematicians —we provide well-known results on them in Proposition 1.A.1. Recall that  $\tilde{\delta}_c(\beta) = \inf\{\delta > 0, h_\beta(\delta) > 0\}$  is the critical point of the wetting model  $Z_{\text{wet},N}^{\beta,\delta}$ ,  $h_\beta(\delta)$  is its free energy, and  $\asymp$  is defined in (5.0.1).

**Proposition 1.4.6.** *Let  $q_2 > q_1 > 0$ , and assume  $\delta > \tilde{\delta}_c(\beta)$ . Then  $E_{N,q}^\circ \asymp e^{h_\beta(\delta)N}$  uniformly in  $q \in [q_1, q_2]$ .*

*Proof.* The upper bound is a straightforward consequence of the inequality  $E_{N,q}^\circ \leq Z_{\text{wet},N}^{\beta,\delta}$  and Proposition 1.A.1 (ii). For the lower bound, we first bound  $\partial_q g(q, 0)$  by some  $c > 0$  for all  $q \in [q_1, q_2]$  (recall that it is a continuous and positive function), then decompose trajectories in  $E_{N,q}^\circ$  into excursions:

$$E_{N,q}^\circ \geq \sum_{r=1}^N \sum_{\substack{1 \leq t_1, \dots, t_r \leq N, \\ t_1 + \dots + t_r = N}} \prod_{i=1}^r K_\beta(t_i) e^\delta \mathbf{E}_{\beta,0} \left[ e^{-c \frac{A_{t_i}(I)}{N}} \mid \tau = t_i, I_{t_i} = 0 \right], \quad (1.4.30)$$

where we define  $\tau = \inf\{t \geq 1; I_t \leq 0\}$  and  $K_\beta(t) = \mathbf{P}_{\beta,0}(\tau = t, I_t = 0)$ ,  $t \geq 1$  as in Appendix 1.A. Using Jensen's inequality, we can claim that

$$\mathbf{E}_{\beta,0} \left[ e^{-c \frac{A_t(I)}{N}} \mid \tau = t, I_t = 0 \right] \geq e^{-c \frac{1}{N} \mathbf{E}_{\beta,0}[A_t(I) \mid \tau=t, I_t=0]}, \quad t \in \mathbb{N}. \quad (1.4.31)$$

Moreover, by using the inequality  $\mathbf{E}_{\beta,0}(|A_t(I)|) \leq \mathbf{E}_{\beta,0}(|I_1|) \sum_{j=1}^t j$  combined with estimates on  $K_\beta$  (see (1.A.2)) we can state that there exists a  $\bar{c} > 0$  such that

$$\mathbf{E}_{\beta,0} [A_t(I) \mid \tau = t, I_t = 0] \leq \mathbf{E}_{\beta,0}(|A_t(I)|) K_\beta(t)^{-1} \leq \bar{c} t^{7/2}, \quad t \in \mathbb{N}. \quad (1.4.32)$$

Using (1.4.30–1.4.32) we obtain that there exists a  $c_1 > 0$  such that,

$$E_{N,q}^\circ \geq \sum_{r=1}^N \sum_{\substack{1 \leq t_1, \dots, t_r \leq N, \\ t_1 + \dots + t_r = N}} \prod_{i=1}^r K_\beta(t_i) e^\delta e^{-c_1 \frac{t_i^{7/2}}{N}}. \quad (1.4.33)$$

Since  $\delta > \tilde{\delta}_c(\beta)$ , equation (1.A.3) guarantees that

$$Q_\infty(t) := K_\beta(t) e^\delta e^{-h_\beta(\delta)t} \quad t \in \mathbb{N}, \quad (1.4.34)$$

is a probability distribution on  $\mathbb{N}$ . We also define for all  $N \in \mathbb{N}$ ,

$$Q_N(t) := K_\beta(t) e^\delta e^{-h_N t} e^{-\frac{c_1}{N} t^{7/2}} \quad t \in \mathbb{N}, \quad (1.4.35)$$

where  $h_N$  is the unique solution of

$$\sum_{t \geq 1} K_\beta(t) e^\delta e^{-h_N t} e^{-\frac{c_1}{N} t^{7/2}} = 1. \quad (1.4.36)$$

Notice that  $h_N \leq h_\beta(\delta)$  and  $h_N \rightarrow h_\beta(\delta)$  as  $N \rightarrow \infty$ . Plugging these into (1.4.33), we have

$$E_{N,q}^\circ \geq e^{h_N N} Q_N(N \in \tau), \quad (1.4.37)$$

where  $\tau$  is a renewal process in  $\mathbb{N}$  whose inter-arrival distribution is given by  $Q_N$ .

**Lemma 1.4.7.** *There is  $C_{12} > 0$  such that for all  $N \in \mathbb{N}$ ,*

$$0 \leq h_\beta(\delta) - h_N \leq \frac{C_{12}}{N}. \quad (1.4.38)$$



*Proof.* We already stated that  $h_N \leq h_\beta(\delta)$  for all  $N \in \mathbb{N}$ . Using that  $Q_\infty$  and  $Q_N$  are probability distributions on  $\mathbb{N}$ , we write

$$\sum_{t \geq 1} K_\beta(t) e^\delta e^{-h_\beta(\delta)t} \left( e^{(h_\beta(\delta) - h_N)t} - 1 \right) = \sum_{t \geq 1} K_\beta(t) e^\delta e^{-h_N t} \left( 1 - e^{-\frac{c_1}{N} t^{7/2}} \right) \geq 0.$$

We bound the left hand side from below using (1.4.34) and

$$e^{(h_\beta(\delta) - h_N)t} - 1 \geq (h_\beta(\delta) - h_N)t \geq h_\beta(\delta) - h_N,$$

for all  $t \geq 1$ , so

$$h_\beta(\delta) - h_N \leq \sum_{t \geq 1} K_\beta(t) e^\delta e^{-h_N t} \left( 1 - e^{-\frac{c_1}{N} t^{7/2}} \right).$$

Since  $\lim_{N \rightarrow \infty} h_N = h_\beta(\delta)$ , we have some  $N_0 \in \mathbb{N}$  such that  $h_N \geq h_\beta(\delta)/2$  for all  $N \geq N_0$ . Moreover  $1 - e^{-\frac{c_1}{N} t^{7/2}} \leq \frac{c_1}{N} t^{7/2}$ , so for all  $N \geq N_0$ ,

$$h_\beta(\delta) - h_N \leq \frac{c_1}{N} \sum_{t \geq 1} K_\beta(t) e^\delta e^{-(h_\beta(\delta)/2)t} t^{7/2},$$

and that sum is finite, which concludes the proof.  $\square$

Applying this lemma to (1.4.37), we obtain

$$E_{N,q}^\circ \geq C_{13} e^{h_\beta(\delta)N} Q_N(N \in \tau) \quad (1.4.39)$$

and therefore the proof will be complete once we show that  $\liminf_N Q_N(N \in \tau) > 0$ . Applying the main theorem from [129], there is  $M$  a random variable with distribution  $\text{Geom}(b)$ ,  $b \in (0, 1)$ , and a sequence  $Z_i = C' + \zeta_i$ ,  $i \geq 1$  where  $C' > 0$  and  $(\zeta_i)_{i \in \mathbb{N}}$  are i.i.d. variables with distribution  $\text{Exp}(a)$ ,  $a > 0$ , such that  $M$  and  $(Z_i)_{i \in \mathbb{N}}$  are independent, and

$$\left| Q_N(N \in \tau) - \frac{1}{\mu_N} \right| \leq C_{14} \mathbf{P} \left( \sum_{i=1}^M Z_i > n \right), \quad (1.4.40)$$

for all  $n \in \mathbb{N}$  and  $N \in \mathbb{N}$ , where  $\mu_N := \mathbf{E}_{Q_N}[\tau_1] \in (0, \infty)$ . The key feature here is that  $C', C_{14} > 0$ ,  $a > 0$  and  $b \in (0, 1)$  can be taken uniformly in  $N \in \mathbb{N}$ . We do not write the details here as it suffices to check the proof in [129]—more precisely one has to use that  $Q_N(t) \leq C_{15} Q_\infty(t)$  uniformly in  $N, t \in \mathbb{N}$  (this follows from Lemma 1.4.7) in [129, (2.9)], and  $Q_N(t) \geq C_{16} Q_\infty(t)$  uniformly in  $N \in \mathbb{N}$  and finitely many  $t \in \mathbb{N}$  in [129, (2.13)].

In particular when  $n = N$ , the right hand side in (1.4.40) decays to 0 as  $N \rightarrow \infty$ , so  $Q_N(N \in \tau) - \frac{1}{\mu_N}$  converges to 0 as  $N \rightarrow \infty$ . Finally, we note that  $\lim_{N \rightarrow \infty} \mu_N = \mu_\infty := \mathbf{E}_{Q_\infty}[\tau_1] \in (0, \infty)$  (by dominated convergence theorem). Thus,  $Q_N(N \in \tau) \geq \frac{1}{2\mu_\infty}$  for  $N$  sufficiently large, and this concludes the proof of Proposition 1.4.6.  $\square$

#### 1.4.4 Step 4 : area-penalized wetting model for $\delta \leq \tilde{\delta}_c(\beta)$

Estimates of  $E_{N,q}^\circ$  in the phase  $\mathcal{DC}$  are more involved than in  $\mathcal{AC}$ . Actually we only manage to find a sharp asymptotic of  $E_{N,q}^\circ$  when  $\delta = 0$ . To lighten upcoming notations, let us define for all  $\gamma > 0$ ,

$$E_N(\gamma) := \mathbf{E}_{\beta,0} \left[ e^{-\gamma \frac{A_N(I)}{N}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \right]. \quad (1.4.41)$$

Recall that  $\sigma_\beta^2$  is the variance of  $I_1$  under  $\mathbf{P}_\beta$ .

For any  $T > 0$  and  $y \geq 0$ , let  $(\mathcal{M}_{s,T}^y)_{s \in [0,T]}$  denote the Brownian meander on  $[0, T]$  starting from  $y$ —henceforth we will denote its law with  $\mathbb{P}$  to distinguish it from  $\mathbf{P}_\beta$ . When  $y > 0$ , it has the same law as the Brownian motion starting from  $y$  and conditioned to remain positive on  $[0, T]$ . We will omit the superscript  $y$  when  $y = 0$ , and the subscript  $T$  when  $T = 1$ . Define for all  $\gamma > 0$ ,

$$J(\gamma) := \lim_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{E} \left[ e^{-\gamma \int_0^T \mathcal{M}_{s,T} ds} \right] = -2^{-1/3} |a_1| \gamma^{2/3}, \quad (1.4.42)$$

where  $a_1$  denotes the smallest zero (in absolute value) of the Airy function. More precisely one notices  $\mathbb{E} \left[ e^{-\gamma \int_0^T \mathcal{M}_{s,T} ds} \right] = \mathbb{E} \left[ e^{-\gamma T^{3/2} \int_0^1 \mathcal{M}_s ds} \right]$ , and the latter expectation has been computed analytically in [145] (see also [108, (209)]), which leads to the second identity. We claim the following.

**Proposition 1.4.8.** *One has*

$$\frac{1}{N^{1/3}} \log E_N(\gamma) \xrightarrow{N \rightarrow \infty} J(\sigma_\beta \gamma), \quad (1.4.43)$$

locally uniformly in  $\gamma \in (0, \infty)$ .

*Proof.* First, we claim that it suffices to prove the pointwise convergence. Indeed, one notices that  $\gamma \mapsto J(\sigma_\beta \gamma)$  is continuous, and  $\gamma \mapsto \frac{1}{N^{1/3}} \log E_N(\gamma)$  is non-increasing for any  $N \in \mathbb{N}$ . It is well-known that those assumptions put together with the pointwise convergence imply the locally uniform convergence—see e.g. [139, Prop. 2.1] for a proof.

Notice that  $\frac{1}{\sigma_\beta} I$  is a random walk on  $\frac{1}{\sigma_\beta} \mathbb{Z}$  with variance 1, and upcoming computations still hold when replacing  $\gamma$  with  $\sigma_\beta \gamma$ . Hence we can assume without loss generality that  $\sigma_\beta = 1$ .

*Upper bound.* Let  $N \in \mathbb{N}$  and  $T > 0$ , and let us denote  $N_T := \lfloor TN^{2/3} \rfloor$ . Set also  $a_N := \lfloor N/N_T \rfloor$ . We decompose a trajectory contributing to  $E_N(\gamma)$  into  $a_N$  blocks of length  $N_T$  and a remaining block of length at most  $N_T$ . We apply Markov property at times  $jN_T$  for  $j \in \{1, \dots, a_N\}$  and we bound the contribution of the very last block by 1 to obtain

$$E_N(\gamma) \leq \left( \sup_{x \in \mathbb{N}} \mathbf{E}_{\beta,x} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mathbf{1}_{\{I \in B_{N_T}^+\}} \right] \right)^{a_N}, \quad (1.4.44)$$

where for all  $n \in \mathbb{N}$ ,  $B_n^+ := \{(X_k)_{k=1}^n \in \mathbb{Z}^n; X_k \geq 0 \forall 1 \leq k \leq n\}$ .

**Lemma 1.4.9.** For  $\alpha > 0$  and  $m \in \mathbb{N}_0$ ,

$$x \mapsto \mathbf{E}_{\beta,x} \left[ e^{-\alpha A_m(I)} \mid I \in B_m^+ \right],$$

is non-increasing on  $\mathbb{N}_0$ . This also holds when conditioning over  $B_m^{0,+}$  instead of  $B_m^+$ .

We postpone the proof of this lemma to Section 1.6.2. Choosing  $\alpha = \gamma/N$  and applying  $\frac{1}{N^{1/3}} \log$  to (1.4.44) (and because  $\mathbf{P}_{\beta,x}(B_m^+) \leq 1$ ), we obtain

$$\begin{aligned} \frac{1}{N^{1/3}} \log E_N(\gamma) &\leq \frac{a_N}{N^{1/3}} \log \left( \sup_{x \in \mathbb{N}} \mathbf{E}_{\beta,x} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mid I \in B_{N_T}^+ \right] \right) \\ &\leq \left( 1 + \frac{1}{N_T} \right) \frac{1}{T} \log \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mid I \in B_{N_T}^+ \right], \end{aligned} \quad (1.4.45)$$

for all  $N, T$ . Furthermore the process  $(\frac{1}{N^{1/3}} I_{\lfloor sN^{2/3} \rfloor})_{s \in [0, T]}$  conditioned to remain non-negative converges weakly as  $N \rightarrow \infty$  to the Brownian meander  $(\mathcal{M}_{s,T})_{s \in [0, T]}$  in the set of cadlag functions on  $[0, T]$  endowed with the uniform convergence topology (see [36, Theorem 3.2]), and the swiped area is a continuous function of the trajectory. Thereby,

$$\lim_{N \rightarrow \infty} \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mid I \in B_{N_T}^+ \right] = \mathbb{E} \left[ \exp \left( -\gamma \int_0^T \mathcal{M}_{s,T} ds \right) \right], \quad (1.4.46)$$

for all  $\gamma > 0$ . Recollecting (1.4.45), we have for any fixed  $T > 0$ ,

$$\limsup_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log E_N(\gamma) \leq \frac{1}{T} \log \left( \mathbb{E} \left[ \exp \left( -\gamma \int_0^T \mathcal{M}_{s,T} ds \right) \right] \right).$$

Choosing  $T > 0$  arbitrarily large, we conclude

$$\limsup_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log E_N(\gamma) \leq J(\gamma).$$

*Lower bound.* It remains to prove that

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log E_N(\gamma) \geq J(\gamma). \quad (1.4.47)$$

Let us settle some notations, i.e., for  $\kappa_1 < \kappa_2 \in \mathbb{R}$  and  $N \in \mathbb{N}$  we define  $\mathcal{O}_{\kappa_1, \kappa_2, N} := [\kappa_1 N^{1/3}, \kappa_2 N^{1/3}]$ . For  $\Delta > 0$ ,  $N, n \in \mathbb{N}$  and  $x, y \in \mathbb{N}_0$  we set

$$\begin{aligned} G_{N,n}^{x,y} &:= \mathbf{E}_{\beta,x} \left[ e^{-\frac{\gamma}{N} A_n(I)} \mathbf{1}_{\{I \in B_n^{y,+}\}} \right], \\ \text{and } \tilde{G}_{N,\Delta,n}^x &:= \sum_{y \in \mathcal{O}_{\Delta/2, \Delta, N}} G_{N,n}^{x,y}. \end{aligned} \quad (1.4.48)$$

Let also  $N \in \mathbb{N}$  and  $T > 0$ , and define  $N_T := \lfloor TN^{2/3} \rfloor$  and  $a_N := \lfloor (N - 2N_T)/(N_T + 2) \rfloor$  (where  $a_N \geq 0$  as soon as  $N \geq 8T^3$ ). We decompose a trajectory in  $E_N(\gamma)$  into  $a_N + 2$  blocks: the first block has length  $N_T$ , the  $a_N$  following blocks have length  $N_T + 2$ , and the final block has length  $\rho N_T$  for some  $\rho \in [1, 2 + \frac{1}{N_T}]$ . We restrict  $E_N(\gamma)$  to those trajectories located inside  $\mathcal{O}_{\Delta/2, \Delta, N}$  at times  $N_T$  and  $(j+1)N_T + 2j$  for every  $j \in \{1, \dots, a_N\}$ . Then, applying Markov property at those times we obtain

$$E_N(\gamma) \geq \tilde{G}_{N, \Delta, N_T}^0 \left[ \hat{G}_{N, \Delta, N_T+2} \right]^{a_N} \inf_{x \in \mathcal{O}_{\Delta/2, \Delta, N}} G_{N, \rho N_T}^{x, 0} \quad (1.4.49)$$

with

$$\hat{G}_{N, \Delta, n} := \inf_{x \in \mathcal{O}_{\Delta/2, \Delta, N}} \tilde{G}_{N, \Delta, n}^x. \quad (1.4.50)$$

We will prove (1.4.47) subject to claims 1.4.10, 1.4.11, 1.4.12 and 1.4.13 below. Before stating those claims we need some additional notations. For  $\Delta > 0$ ,  $n \in \mathbb{N}$  and  $N \in \mathbb{N}$  we set

$$U_{N, \Delta, n}^j := \mathbf{E}_{\beta, 0} \left[ e^{-\frac{\gamma}{N} A_n(I)} \mathbf{1}_{\{I \in \tilde{B}_{n, \Delta N^{1/3}}^+\}} \mathbf{1}_{\{I_n \in \tilde{\mathcal{O}}_{\Delta, N}^j\}} \right], \quad j \in \{1, 2, 3\}, \quad (1.4.51)$$

with

$$\tilde{\mathcal{O}}_{\Delta, N}^1 := \left( -\frac{\Delta N^{1/3}}{2}, -\frac{\Delta N^{1/3}}{4} \right) \quad \text{and} \quad \tilde{\mathcal{O}}_{\Delta, N}^2 := \left[ -\frac{\Delta N^{1/3}}{4}, 0 \right) \quad \text{and} \quad \tilde{\mathcal{O}}_{\Delta, N}^3 := \left[ 0, \frac{\Delta N^{1/3}}{4} \right].$$

and where we introduce for  $\kappa \geq 0$  and  $n \in \mathbb{N}$ ,

$$\tilde{B}_{n, \kappa}^+ := \left\{ (X_i)_{i=1}^n \in \mathbb{Z}^n; X_i \geq -\frac{\kappa}{2}, \forall 1 \leq i < n, X_n > -\frac{\kappa}{2} \right\}.$$

Claim 1.4.10 handles the first and last factors in (1.4.49). Claim 1.4.11 combined with Claim 1.4.12 will be useful to get rid of the infimum in the definition of  $\hat{G}_{N, \Delta, n}$  in (1.4.50) and replace it by  $U_{N, \Delta, n}^1 + U_{N, \Delta, n}^2$ . On this latter quantity one may apply an FKG inequality and get rid of the constraint  $\{I_n \in \mathcal{O}_{\Delta/2, \Delta, N}\}$  in  $\hat{G}_{N, \Delta, n}$ . Claim 1.4.13 will be used at the end of the proof to retrieve  $J(\gamma)$ .

**Claim 1.4.10.** For  $\Delta > 0$  and  $T > 0$ , one has

$$\lim_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log \tilde{G}_{N, \Delta, \rho N_T}^0 = \lim_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log \left( \inf_{x \in \mathcal{O}_{\Delta/2, \Delta, N}} G_{N, \rho N_T}^{x, 0} \right) = 0, \quad (1.4.52)$$

uniformly in  $\rho \in [1, 3]$ .

**Claim 1.4.11.** For  $\Delta > 0$  and  $N, n \in \mathbb{N}$ ,

$$\hat{G}_{N, \Delta, n} \geq \exp \left( -\frac{\gamma \Delta n}{N^{2/3}} \right) \min \left\{ U_{N, \Delta, n}^2, U_{N, \Delta, n}^3 \right\}. \quad (1.4.53)$$

**Claim 1.4.12.** There exists a  $C > 0$  such that for  $\Delta > 0$  and  $N, n \in \mathbb{N}$ ,

$$U_{N, \Delta, n}^j \leq C \exp \left( \frac{\gamma \Delta n}{N^{2/3}} \right) U_{N, \Delta, n+1}^{j+1}, \quad \text{for } j \in \{1, 2\}. \quad (1.4.54)$$

**Claim 1.4.13.** *One has*

$$\limsup_{\Delta \rightarrow 0^+} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{E} \left[ e^{-\gamma \int_0^T \mathcal{M}_{s,T}^\Delta ds} \right] \geq J(\gamma). \quad (1.4.55)$$

We resume the proof of the lower bound. We observe that by constraining the last increment of the random walk  $I$  to be null in  $U_{N,\Delta,n+1}^2$  we get the inequality

$$U_{N,\Delta,n}^2 \leq c_\beta U_{N,\Delta,n+1}^2. \quad (1.4.56)$$

Then, using Claim 1.4.12 for  $n = N_T$  and  $j = 1$  and using (1.4.56) for  $n = TN^{2/3}$ , we obtain that there exists a  $C_1 > 0$  such that for  $T, \Delta > 0$  and  $N \in \mathbb{N}$

$$U_{N,\Delta,N_T}^1 + U_{N,\Delta,N_T}^2 \leq C_1 e^{\gamma \Delta T} U_{N,\Delta,N_T+1}^2. \quad (1.4.57)$$

Then, it suffices to apply, on the one hand, Claim 1.4.12 for  $j = 2$  and  $n = N_T + 1$  to the r.h.s. in (1.4.57) and, on the other hand, (1.4.56) for  $n = N_T + 1$  to the r.h.s. in (1.4.57) to assert that there exists a  $C_2 > 0$  such that for  $T, \Delta > 0$  and  $N \in \mathbb{N}$ ,

$$V_{N_T} := U_{N,\Delta,N_T}^1 + U_{N,\Delta,N_T}^2 \leq C_2 e^{2\gamma \Delta T} \min\{U_{N,\Delta,N_T+2}^2, U_{N,\Delta,N_T+2}^3\}. \quad (1.4.58)$$

We note that  $V_{N_T}$  may be rewritten under the form

$$\begin{aligned} V_{N_T} &= \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mathbf{1}_{\{I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+\}} \mathbf{1}_{\{I_{N_T} \leq 0\}} \right] \\ &= \mathbf{P}_{\beta,0} \left( I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+ \right) \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mathbf{1}_{\{I_{N_T} \leq 0\}} \middle| I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+ \right]. \end{aligned} \quad (1.4.59)$$

By using the FKG inequality described in Proposition 1.B.1 with  $A := \tilde{B}_{N_T, \Delta N^{1/3}}^+$  we obtain

$$\begin{aligned} \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \mathbf{1}_{\{I_{N_T} \leq 0\}} \middle| I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+ \right] &\geq \\ \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{N_T}(I)} \middle| I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+ \right] \mathbf{P}_{\beta,0} \left[ I_{N_T} \leq 0 \middle| I \in \tilde{B}_{N_T, \Delta N^{1/3}}^+ \right]. \end{aligned} \quad (1.4.60)$$

By Donsker's invariance principle the rescaled process  $(\frac{1}{N^{1/3}} I_{\lfloor sN^{2/3} \rfloor})_{s \in [0, T]}$  converges in distribution towards  $(B_s)_{s \in [0, T]}$  a standard Brownian motion, in the set of cadlag functions on  $[0, T]$  endowed with the uniform convergence topology. Therefore, we set  $m_T := \min\{B_s, s \in [0, T]\}$  and we obtain

$$\liminf_{N \rightarrow \infty} V_{N_T} \geq \mathbb{P} \left[ B_T < 0, m_T > -\frac{\Delta}{2} \right] \mathbb{E} \left[ e^{-\gamma \int_0^T B_s ds} \middle| m_T > -\frac{\Delta}{2} \right], \quad (1.4.61)$$

Finally, using (1.4.52), (1.4.53), (1.4.58) and (1.4.61) we can deduce from (1.4.49) that

$$\begin{aligned} \liminf_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log E_N(\gamma) &\geq -\frac{\log C_3}{T} - 3\gamma\Delta + \frac{1}{T} \log \mathbb{P}\left[B_T < 0, m_T > -\frac{\Delta}{2}\right] \\ &\quad + \frac{1}{T} \log \mathbb{E}\left[e^{-\gamma \int_0^T B_s ds} \mid m_T > -\frac{\Delta}{2}\right]. \end{aligned} \quad (1.4.62)$$

Recalling that  $\Delta \leq 1$  and using [112, Prop. 8.1] we can state that there exists some  $c > 0$  such that for  $T$  large enough  $\mathbb{P}\left[B_T < 0, m_T > -\frac{\Delta}{2}\right] \geq c\Delta^2/T$  which, after taking  $\limsup_{T \rightarrow \infty}$  in the r.h.s. in (1.4.62) allows us to write that for every  $\Delta > 0$ ,

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1/3}} \log E_N(\gamma) \geq -3\gamma\Delta + \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{E}_{\Delta/2}\left[e^{-\gamma \int_0^T M_s ds}\right], \quad (1.4.63)$$

and we conclude by using Claim 1.4.13 and by taking  $\limsup_{\Delta \rightarrow 0^+}$  in the r.h.s. in (1.4.63). This proves (1.4.47) and completes the proof of (1.4.8).  $\square$

*Proof of Claim 1.4.10.* Let  $\Delta, T > 0$ . First we notice that we only need to prove a lower bound on  $G_{N,n}^{0,x}$  uniform in  $x \in \mathcal{O}_{\Delta/2,\Delta,N}$  as  $N \rightarrow \infty$ . Indeed, a time-reversal argument yields that  $G_{N,n}^{x,0} = G_{N,n}^{0,x}$  for all  $x \in \mathbb{N}_0, n \in \mathbb{N}$  (recall that  $\mathbf{P}_\beta$  is symmetric, see (1.2.8)), and

$$\inf_{x \in \mathcal{O}_{\Delta/2,\Delta,N}} G_{N,n}^{0,x} \leq \tilde{G}_{N,\Delta,n}^0 \leq 1. \quad (1.4.64)$$

Moreover for any  $x \in \mathcal{O}_{\Delta/2,\Delta,N}$  and  $\rho \in [1, 3]$  such that  $\rho N_T \in \mathbb{N}$ , we write

$$G_{N,\rho N_T}^{x,0} = \mathbf{E}_{\beta,x} \left[ e^{-\frac{\gamma}{N} A_{\rho N_T}(I)} \mid I \in B_{\rho N_T}^{0,+} \right] \mathbf{P}_{\beta,x} \left( B_{\rho N_T}^{0,+} \right). \quad (1.4.65)$$

Lemma 1.4.9 claims that the first factor is non-increasing in  $x$ . Moreover, the second factor is polynomial in  $N$  uniformly in  $x \in \mathcal{O}_{\Delta/2,\Delta,N}$  and  $\rho \in [1, 3] \cap \frac{1}{N_T} \mathbb{N}$ , see (1.6.1) below. Thereby, we deduce from (1.4.65) that

$$G_{N,\rho N_T}^{x,0} \geq \frac{C_{17}}{N^{2/3}} \mathbf{E}_{\beta,\Delta N^{1/3}} \left[ e^{-\frac{\gamma}{N} A_{\rho N_T}(I)} \mid I \in B_{\rho N_T}^{0,+} \right], \quad (1.4.66)$$

for  $N \in \mathbb{N}$ , uniformly in  $x \in \mathcal{O}_{\Delta/2,\Delta,N}$  and  $\rho \in [1, 3] \cap \frac{1}{N_T} \mathbb{N}$ —notice that we wrote  $\Delta N^{1/3}$  instead of  $\lceil \Delta N^{1/3} \rceil$ , and we will omit all ceil functions henceforth to lighten notations. Recalling that  $\frac{1}{2}TN^{2/3} \leq \rho N_T \leq 3TN^{2/3}$  and  $\rho \in [1, 3]$ , (1.4.66) implies

$$G_{N,\rho N_T}^{x,0} \geq \frac{C_{17}}{N^{2/3}} \mathbf{E}_{\beta,\Delta \sqrt{\frac{2}{T}} \sqrt{\rho N_T}} \left[ e^{-\gamma \left(\frac{3T}{\rho N_T}\right)^{3/2} A_{\rho N_T}(I)} \mid I \in B_{\rho N_T}^{0,+} \right]. \quad (1.4.67)$$

Furthermore, it is proven in [47, Thm. 2.4] that a properly rescaled random walk of length  $n$  starting from  $c\sqrt{n}$ ,  $c \in \mathbb{R}_+$ , conditioned to remain non-negative and end in 0, converges in distribution as  $n \rightarrow \infty$  to a Brownian bridge starting from  $c$ , conditioned to remain non-negative and to end in 0. Moreover  $\rho N_T \geq \frac{1}{2}TN^{2/3}$  for all  $\rho \in [1, 3]$ , so the expectation in the r.h.s of (1.4.67) converges as  $N \rightarrow \infty$  to some positive constant, uniformly in  $\rho \in [1, 3]$ . Recollecting (1.4.64), this concludes the proof of the claim.  $\square$

*Proof of Claim 1.4.11.* Recall (1.4.50) and note that

$$\tilde{G}_{N,\Delta,n}^x = \mathbf{E}_{\beta,x} \left[ e^{-\frac{\gamma}{N} A_n(I)} \mathbf{1}_{\{I \in \tilde{B}_{n,0}^+\}} \mathbf{1}_{\{I_n \in \mathcal{O}_{\Delta/2,\Delta,N}\}} \right]. \quad (1.4.68)$$

Let us first consider the case  $x \in \mathcal{O}_{\frac{\Delta}{2}, \frac{3\Delta}{4}, N}$ . We observe that if a trajectory  $I := (I_i)_{i=0}^n$  satisfies  $I_0 = 0$ ,  $I \in \tilde{B}_{n,\Delta N^{1/3}}^+$  and  $I_n \in \mathcal{O}_{0,\Delta/4,N}$  then  $x + I := (x + I_i)_{i=0}^n$  satisfies  $x + I \in \tilde{B}_{n,0}^+$  and  $x + I_n \in \mathcal{O}_{\frac{\Delta}{2},\Delta,N}$ . As a consequence

$$\begin{aligned} \tilde{G}_{N,\Delta,n}^x &\geq \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_n(x+I)} \mathbf{1}_{\{I \in \tilde{B}_{n,\Delta N^{1/3}}^+\}} \mathbf{1}_{\{I_n \in \mathcal{O}_{0,\Delta/4,N}\}} \right] \\ &\geq \exp\left(-\frac{\gamma\Delta n}{N^{2/3}}\right) U_{N,\Delta,n}^3. \end{aligned} \quad (1.4.69)$$

The case  $x \in \mathcal{O}_{\frac{3\Delta}{4},\Delta,N}$  is taken care of similarly. The only difference is that we consider  $I := (I_i)_{i=0}^n$  satisfies  $I_0 = 0$ ,  $I \in \tilde{B}_{n,\Delta N^{1/3}}^+$  and  $I_n \in \mathcal{O}_{-\frac{\Delta}{4},0,N}$  such that  $x + I \in \tilde{B}_{n,0}^+$  and  $x + I_n \in \mathcal{O}_{\frac{\Delta}{2},\Delta,N}$ . Then,

$$\begin{aligned} \tilde{G}_{N,\Delta,n}^x &\geq \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_n(x+I)} \mathbf{1}_{\{I \in \tilde{B}_{n,\Delta N^{1/3}}^+\}} \mathbf{1}_{\{I_n \in \mathcal{O}_{-\frac{\Delta}{4},0,N}\}} \right] \\ &\geq \exp\left(-\frac{\gamma\Delta n}{N^{2/3}}\right) U_{N,\Delta,n}^2, \end{aligned} \quad (1.4.70)$$

and this completes the proof of the claim.  $\square$

*Proof of Claim 1.4.12.* We will focus on proving the Claim for  $j = 2$ . The case  $j = 1$  can be taken care of similarly. We decompose  $U_{N,\Delta,n}^2$  depending on the time  $\tau$  at which the trajectory is above the  $x$ -axis for the last time before time  $n$ , that is  $\tau := \max\{i \leq n : I_i \geq 0\}$ . This gives

$$U_{N,\Delta,n}^2 = \sum_{k=0}^{n-1} U_{N,\Delta,n}^{2,k} := \sum_{k=0}^{n-1} \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_n(I)} \mathbf{1}_{\{I \in \tilde{B}_{n,\Delta N^{1/3}}^+\}} \mathbf{1}_{\{\tau=k\}} \mathbf{1}_{\{I_n \in \tilde{\mathcal{O}}_{\Delta,N}^2\}} \right]. \quad (1.4.71)$$

For  $k \in \{0, \dots, n-1\}$  we partition the trajectories contributing to  $U_{N,\Delta,n}^{2,k}$  depending on the values  $z$  and  $y$  taken by  $I_\tau$  and  $I_{\tau+1}$ , respectively. This gives

$$\begin{aligned} U_{N,\Delta,n}^{2,k} &:= \sum_{z \geq 0} \sum_{y=1}^{\Delta N^{1/3}/2} \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_k(I)} \mathbf{1}_{\{I \in \tilde{B}_{k,\Delta N^{1/3}}^+\}} \mathbf{1}_{\{I_k=z\}} \right] \mathbf{P}_\beta(I_1 = -z - y) \\ &\quad \mathbf{E}_{\beta,-y} \left[ e^{-\frac{\gamma}{N} A_{n-k-1}(I)} \mathbf{1}_{\{-\frac{\Delta N^{1/3}}{2} \leq I_i < 0, i \leq n-k-1\}} \mathbf{1}_{\{I_{n-k-1} \in \tilde{\mathcal{O}}_{\Delta,N}^2\}} \right]. \end{aligned} \quad (1.4.72)$$

We observe that  $\mathbf{P}_\beta(I_1 = -z - y) = c_\beta \mathbf{P}_\beta(I_1 = -z) \mathbf{P}_\beta(I_1 = y)$ . Since the increments of  $I$  have a symmetric law, we can rewrite the last expectation of the r.h.s. in (1.4.72) as

$$\begin{aligned} & \mathbf{E}_{\beta,y} \left[ e^{-\frac{\gamma}{N} A_{n-k-1}(-I)} \mathbf{1}_{\left\{-\frac{\Delta N^{1/3}}{2} \leq -I_i < 0, i \leq n-k-1\right\}} \mathbf{1}_{\left\{-I_{n-k-1} \in \tilde{\mathcal{O}}_{\Delta,N}^2\right\}} \right] \\ & \leq \exp\left(\frac{\gamma \Delta n}{N^{2/3}}\right) \mathbf{E}_{\beta,y} \left[ e^{-\frac{\gamma}{N} A_{n-k-1}(I)} \mathbf{1}_{\left\{0 < I_i \leq \frac{\Delta N^{1/3}}{2}, i \leq n-k-1\right\}} \mathbf{1}_{\left\{I_{n-k-1} \in \tilde{\mathcal{O}}_{\Delta,N}^3 \setminus \{0\}\right\}} \right]. \end{aligned} \quad (1.4.73)$$

where we have used that any trajectory  $I = (I_i)_{i=0}^{n-k-1}$  that contributes the expectation of the l.h.s. in (1.4.73) satisfies  $A_{n-k-1}(I) = -A_{n-k-1}(-I) \leq \frac{1}{2} \Delta N^{1/3} (n-k-1)$ . Thus, combining (1.4.72) with (1.4.73) we obtain

$$\begin{aligned} U_{N,\Delta,n}^{2,k} & \leq c_\beta \exp\left(\frac{\gamma \Delta n}{N^{2/3}}\right) \sum_{z \geq 0} \sum_{y=1}^{\Delta N^{1/3}/2} \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_k(I)} \mathbf{1}_{\left\{I \in \tilde{\mathcal{B}}_{k,\Delta N^{1/3}}^+\right\}} \mathbf{1}_{\{I_k=z\}} \right] \mathbf{P}_\beta(I_1 = -z) \\ & \quad \times \mathbf{P}_\beta(I_1 = y) \mathbf{E}_{\beta,y} \left[ e^{-\frac{\gamma}{N} A_{n-k-1}(I)} \mathbf{1}_{\left\{0 < I_i \leq \frac{\Delta N^{1/3}}{2}, i \leq n-k-1\right\}} \mathbf{1}_{\left\{I_{n-k-1} \in \tilde{\mathcal{O}}_{\Delta,N}^3 \setminus \{0\}\right\}} \right] \\ & \leq c_\beta \exp\left(\frac{\gamma \Delta n}{N^{2/3}}\right) \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{n+1}(I)} \mathbf{1}_{\left\{I \in \tilde{\mathcal{B}}_{n+1,\Delta N^{1/3}}^+\right\}} \mathbf{1}_{\{\tilde{\tau}=k+1\}} \mathbf{1}_{\left\{I_{n+1} \in \tilde{\mathcal{O}}_{\Delta,N}^3\right\}} \right], \end{aligned} \quad (1.4.74)$$

where  $\tilde{\tau} := \max\{0 \leq j \leq n+1 : I_j = 0\}$  is the last time before time  $n+1$  at which  $I$  touches the  $x$ -axis. Therefore,

$$\begin{aligned} \sum_{k=0}^{n-1} U_{N,\Delta,n}^{2,k} & \leq c_\beta \exp\left(\frac{\gamma \Delta n}{N^{2/3}}\right) \mathbf{E}_{\beta,0} \left[ e^{-\frac{\gamma}{N} A_{n+1}(I)} \mathbf{1}_{\left\{I \in \tilde{\mathcal{B}}_{n+1,\Delta N^{1/3}}^+\right\}} \mathbf{1}_{\left\{I_{n+1} \in \tilde{\mathcal{O}}_{\Delta,N}^3\right\}} \right] \\ & \leq c_\beta \exp\left(\frac{\gamma \Delta n}{N^{2/3}}\right) U_{N,\Delta,n+1}^3, \end{aligned}$$

and this completes the proof of the claim.  $\square$

*Proof of Claim 1.4.13.* For conciseness we set, for  $\Delta, T > 0$

$$R_\Delta(T) := \mathbb{E} \left[ e^{-\gamma \int_0^T \mathcal{M}_{s,T}^\Delta ds} \right]. \quad (1.4.75)$$

We assume by contradiction that there exists an  $\varepsilon > 0$  and a  $\Delta_0 \in (0, 1]$  such that for every  $\Delta \in (0, \Delta_0]$

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \log R_\Delta(T) \leq J(\gamma) - \varepsilon. \quad (1.4.76)$$



For  $\Delta, T > 0$ , we set  $\tau_\Delta := \inf\{s \geq 0: \mathcal{M}_{s,T} = \Delta\}$  ( $\tau_\Delta$  depends on  $T$  as well but we omit it for conciseness). We state a small ball inequality that will be proven at the end of the present proof, i.e., there exists  $c_1, c_2 > 0$  such that for  $T > 0$ ,  $\eta \in (0, 1]$  and  $\Delta > 0$ ,

$$\begin{aligned} \mathbb{P}(\tau_\Delta \geq \eta T) &= \mathbb{P}\left[\max_{s \in [0, \eta T]} \mathcal{M}_{s,T} \leq \Delta\right] \\ &= \mathbb{P}\left[\max_{s \in [0, \eta]} \mathcal{M}_s \leq \Delta/\sqrt{T}\right] \leq \frac{c_1}{\sqrt{\eta}} e^{-\frac{c_2}{\Delta^2} \eta T}, \end{aligned} \quad (1.4.77)$$

where we have used a standard scaling property of Brownian meander to write the second equality in (1.4.77).

At this stage, we choose  $\eta > 0$  and  $\Delta \in (0, \Delta_0]$  such that

$$(J(\gamma) - \frac{\varepsilon}{4})(1 - \eta) \leq J(\gamma) - \frac{\varepsilon}{8} \quad \text{and} \quad -\frac{c_2 \eta}{\Delta^2} \leq J(\gamma) - \frac{\varepsilon}{2}. \quad (1.4.78)$$

Applying Markov property at time  $\tau_\Delta$  we obtain

$$\mathbb{E}\left[e^{-\gamma \int_0^T \mathcal{M}_{s,T} ds}\right] \leq \mathbb{P}(\tau_\Delta \geq \eta T) + \mathbb{E}\left[\mathbf{1}_{\{\tau_\Delta < \eta T\}} R_\Delta(T - \tau_\Delta)\right] \quad (1.4.79)$$

Since  $\Delta \in (0, \Delta_0]$  we apply (1.4.76) and there exists a  $T_0 > 0$  such that  $R_\Delta(T) \leq e^{(J(\gamma) - \varepsilon/4)T}$  for  $T \geq T_0$ . It remains to apply (1.4.79) with  $T \geq T_0/(1 - \eta)$  in combination with (1.4.77) and (1.4.78) to obtain

$$\begin{aligned} \mathbb{E}\left[e^{-\gamma \int_0^T \mathcal{M}_{s,T} ds}\right] &\leq c_1 e^{(J(\gamma) - \frac{\varepsilon}{2})T} + \mathbb{E}\left[\mathbf{1}_{\{\tau_\Delta < \eta T\}} R_\Delta(T - \tau_\Delta)\right] \\ &\leq c_1 e^{(J(\gamma) - \frac{\varepsilon}{2})T} + \mathbb{E}\left[\mathbf{1}_{\{\tau_\Delta < \eta T\}} e^{(J(\gamma) - \varepsilon/4)(T - \tau_\Delta)}\right] \\ &\leq c_1 e^{(J(\gamma) - \frac{\varepsilon}{2})T} + e^{(J(\gamma) - \varepsilon/8)T}. \end{aligned} \quad (1.4.80)$$

Taking  $\frac{1}{T} \log$  on both sides in (1.4.80) and letting  $T \rightarrow \infty$  we obtain the contradiction  $J(\gamma) \leq J(\gamma) - \varepsilon/8$ . This completes the proof of the claim.

Let us quickly sketch the proof of the inequality in (1.4.77). We use that  $(\mathcal{M}_s)_{s \in [0,1]}$  is the limit in distribution of  $(B_s)_{s \in [0,1]}$  conditioned on  $\{m_1 > -u\}$  as  $u \rightarrow 0^+$  where we define  $m_t := \min_{s \in [0,t]}(B_s)$  for  $0 \leq t \leq 1$  (see [69, Section 2]). Therefore, for  $\Delta > 0$ ,  $u \in (0, \Delta)$  and  $\eta \in (0, 1)$ ,

$$\mathbb{P}\left[\max_{s \in [0, \eta]} \mathcal{M}_s \leq \Delta\right] = \lim_{u \rightarrow 0^+} \mathbb{P}\left[\max_{s \in [0, \eta]} B_s \leq \Delta \mid m_1 > -u\right]. \quad (1.4.81)$$

By applying Markov property at time  $\eta/2$ , we can bound from above the probability in the r.h.s. in (1.4.81) by

$$\frac{1}{\mathbb{P}(m_1 > -u)} \mathbb{P}(B_s \in (-u, \Delta], \forall s \in [0, \eta/2]) \sup_{x \in (-u, \Delta]} \mathbb{P}_x(B_s \in (-u, \Delta], \forall s \in [0, \eta/2]), \quad (1.4.82)$$

and since a Brownian motion  $(B_s)_{s \geq 0}$  of law  $\mathbb{P}_x$  has the same law as  $(x + B_s)_{s \geq 0}$  with  $(B_s)_{s \geq 0}$  of law  $\mathbb{P}_0$ , (1.4.82) is also smaller than

$$\frac{\mathbb{P}(m_{\eta/2} > -u)}{\mathbb{P}(m_1 > -u)} \mathbb{P} \left[ \max_{s \in [0, \eta/2]} |B_s| \leq 2\Delta \right] \quad (1.4.83)$$

At this stage, the inequality in the r.h.s. in (1.4.77) is obtained by combining (1.4.83) with the following two results:

$$\mathbb{P}(m_t > -u) = \sqrt{\frac{2}{\pi}} \int_0^{u/\sqrt{t}} e^{-\frac{x^2}{2}} dx, \quad \text{for } t, u > 0 \quad (1.4.84)$$

and there exist  $c_1, c_2 > 0$  such that for every  $\kappa > 0$

$$\mathbb{P}(\max_{s \in [0, 1]} |B_s| \leq \kappa) \leq c_1 e^{-c_2/\kappa^2}. \quad (1.4.85)$$

Note that (1.4.84) is obtained by observing that, since  $B$  is symmetric,  $\mathbb{P}(m_t > -u) = \mathbb{P}(M_t < u)$ , where  $M_t$  is the maximum of  $B$  on  $[0, t]$  and the law of  $M_t$  is well known (see e.g. [112, Proposition 8.1]). Proving (1.4.85) can be done by estimating the probability that  $|B|$  is smaller than  $\kappa$  at times  $\{j/k, 1 \leq j \leq k\}$  with  $k = \lfloor 1/\kappa^2 \rfloor$  and by applying Markov property at those times. □

### 1.4.5 Step 5 : horizontal extension inside the collapsed phase

We now prove that in the collapsed phase, the typical horizontal extension of the polymer is of order  $\sqrt{L}$  for  $L$  large. For any interval  $I \subseteq \mathbb{R}$ , define

$$\tilde{Z}_{L, \beta, \delta}^{\circ, +}(I) = \frac{1}{c_\beta e^{\beta L}} Z_{L, \beta, \delta}^{\circ, +}(I) := \sum_{\substack{1 \leq N \leq L/2, \\ N \in I \cap \mathbb{N}}} (\Gamma_\beta)^{2N} D_{N, q_N^L}^\circ, \quad (1.4.86)$$

where  $q_N^L$  and  $D_{N, q_N^L}^\circ$  are defined in Proposition 1.4.1. Recall that  $\mathcal{C}_{\text{bead}}$  denotes the collapsed phase of the single-bead model (see Proposition 1.3.1), and let  $(\mathcal{C}_{\text{bead}})^\circ$  denote its interior.

**Lemma 1.4.14.** *Let  $(\beta, \delta) \in (\mathcal{C}_{\text{bead}})^\circ$ , there exists  $(a_1, a_2) \in (0, \infty)^2$  such that*

$$\lim_{L \rightarrow \infty} \frac{Z_{L, \beta, \delta}^{\circ, +}([a_1, a_2]\sqrt{L})}{Z_{L, \beta, \delta}^{\circ, +}} = 1 \quad (1.4.87)$$

*Proof.* Recall (1.4.7) and (1.4.8). By relaxing all constraints on  $S, I$  but  $\{I \in B_N^{0, +}\}$  in (1.4.8), one has the obvious bound  $D_{N, q}^\circ \leq Z_{\text{wet}, N}^{\beta, \delta}$  (recall (1.A.1)). Moreover Proposition 1.A.1 (ii) implies that there exists some  $K > 0$  such that  $Z_{\text{wet}, N}^{\beta, \delta} \leq K e^{h_\beta(\delta)N}$  for  $N \geq 1$ . Thereby,

$$\tilde{Z}_{L, \beta, \delta}^{\circ, +}([a_2\sqrt{L}, L]) = \sum_{N=a_2\sqrt{L}}^{L/2} (\Gamma_\beta)^{2N} D_{N, q_N^L}^\circ \leq \sum_{N=a_2\sqrt{L}}^{L/2} (\Gamma_\beta)^{2N} K e^{h_\beta(\delta)N}.$$

In  $(\mathcal{C}_{\text{bead}})^\circ$ , one has  $2 \log \Gamma_\beta + h_\beta(\delta) < 0$  (recall (1.3.5)), so we conclude that there exist  $c_1, c_2 > 0$  such that

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+}([a_2\sqrt{L}, L]) \leq c_1 e^{-c_2 a_2 \sqrt{L}}, \quad (1.4.88)$$

for all  $L \geq 1$  and  $a_2 \in \frac{\mathbb{N}}{\sqrt{L}}$ .

Regarding trajectories with an horizontal extension smaller than  $a_1\sqrt{L}$ , we use  $\Gamma_\beta \leq 1$  and (1.4.7–1.4.8) to write

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+}([1, a_1\sqrt{L}]) \leq e^{\delta a_1 \sqrt{L}} \sum_{N \leq a_1 \sqrt{L}} \mathbf{P}_\beta(A_{N+1}(S) \geq \frac{L}{2} - N). \quad (1.4.89)$$

The inequality  $A_{N+1}(S) \geq \frac{L}{2} - N$  implies that  $\max\{S_i, i \leq N\} \geq \frac{L/2 - N}{N+1}$ . Therefore, for  $L$  large enough we can assert that for every  $a_1 \leq 1$  and  $N \leq a_1\sqrt{L}$

$$\mathbf{P}_\beta(A_{N+1}(S) \geq \frac{L}{2} - N) \leq \mathbf{P}_\beta\left(\sum_{i \leq a_1 \sqrt{L}} |S_i - S_{i-1}| \geq \frac{\sqrt{L}}{4a_1}\right). \quad (1.4.90)$$

Since  $|S_1 - S_0|$  (of law  $\mathbf{P}_\beta$ ) has finite small exponential moments, Chernov's bound guarantees us that there exists a function  $\rho : (0, \infty) \mapsto \mathbb{R}^+$  satisfying  $\lim_{a_1 \rightarrow 0^+} \rho(a_1) = \infty$  such that the r.h.s. in (1.4.90) is bounded above by  $e^{-\rho(a_1)\sqrt{L}}$ . Therefore, for  $L$  large enough and  $a_1 \leq 1$ ,

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+}([1, a_1\sqrt{L}]) \leq \sqrt{L} e^{-(\rho(a_1) - \delta a_1)\sqrt{L}}. \quad (1.4.91)$$

At this stage it remains to display a lower bound on  $Z_{L,\beta,\delta}^{\circ,+}$ . To that aim we only consider the term indexed by  $N = \sqrt{L}$  in the sum in (1.4.7). Applying Proposition (1.4.2) with  $q = \frac{1}{2} - \frac{1}{\sqrt{L}}$  and using that  $g$  is  $\mathcal{C}^1$  we deduce that there exists a  $C > 0$  such that

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+} \geq \frac{C}{L} (\Gamma_\beta)^{2\sqrt{L}} e^{-g(\frac{1}{2}, 0)\sqrt{L}} E_{\sqrt{L}, \frac{1}{2} - \frac{1}{\sqrt{L}}}^\circ. \quad (1.4.92)$$

It remains to bound  $E_{\sqrt{L}, \frac{1}{2} - \frac{1}{\sqrt{L}}}^\circ$  from below by choosing the trajectory  $(I_i)_{i=0}^{\sqrt{L}}$  that sticks to zero, i.e.,  $E_{\sqrt{L}, \frac{1}{2} - \frac{1}{\sqrt{L}}}^\circ \geq (e^\delta / c_\beta)^{\sqrt{L}}$ . We set  $\kappa := 2 \log \Gamma_\beta + \delta - g(1/2, 0) - \log c_\beta$  and we obtain that for  $L$  large enough,

$$\tilde{Z}_{L,\beta,\delta}^{\circ,+} \geq \frac{C}{L} e^{\kappa \sqrt{L}}. \quad (1.4.93)$$

Combining (1.4.88), (1.4.91) and (1.4.93), it suffices to choose  $a_1$  (resp.  $a_2$ ) small enough (resp. large enough) for (1.4.87) to hold.  $\square$

### 1.4.6 Step 6 : proof of Theorem 1.3.2

We finally have all required estimates to prove Theorem 1.3.2. Assume  $\beta > \beta_c$  and  $\delta < \delta_c^\circ(\beta)$ . Applying Proposition 1.4.1 and Lemma 1.4.14, we can restrict the sum in (1.4.7) to  $N \in [a_1, a_2] \sqrt{L} \cap \mathbb{N}$ . In particular this implies that  $q_N^L = \frac{L-2N}{2N^2} \in [\frac{1}{2a_2^2} - \frac{1}{a_1\sqrt{L}}, \frac{1}{2a_1^2}]$ , and  $\frac{1}{2a_2^2} - \frac{1}{a_1\sqrt{L}} > 0$  for  $L$  sufficiently large. Thereby Proposition 1.4.2 yields

$$Z_{L,\beta,\delta}^{\circ,+} \asymp \frac{e^{\beta L}}{L} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} (\Gamma_\beta)^{2N} e^{-Ng\left(\frac{L}{2N^2}, 0\right)} E_{N,q_N^L}^\circ, \quad (1.4.94)$$

uniformly in  $L \in \mathbb{N}$  sufficiently large, and where we also used that  $g$  is  $\mathcal{C}^1$ , so  $g(q_N^L, 0) = g\left(\frac{L}{2N^2}, 0\right) + O(1/N)$  uniformly in  $N \in [a_1, a_2] \sqrt{L} \cap \mathbb{N}$ .

**Case**  $\tilde{\delta}_c(\beta) < \delta < \delta_c^\circ(\beta)$

We apply Proposition 1.4.6 to (1.4.94), to write

$$Z_{L,\beta,\delta}^{\circ,+} \asymp \frac{e^{\beta L}}{L} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} e^{\sqrt{L}\varphi\left(\frac{N}{\sqrt{L}}\right)}, \quad (1.4.95)$$

where we define for all  $a > 0$ ,

$$\varphi(a) := a \left( 2 \log \Gamma_\beta + h_\beta(\delta) - g\left(\frac{1}{2a^2}, 0\right) \right). \quad (1.4.96)$$

Notice that  $\varphi$  is  $\mathcal{C}^2$  and negative on  $(0, \infty)$  (recall that  $g$  is non-negative and  $2 \log \Gamma_\beta + h_\beta(\delta) < 0$  in  $(\mathcal{C}_{\text{bead}})^\circ$ ). We claim that it is strictly concave, and that it reaches its maximum at some  $\tilde{a} \in (0, \infty)$ . Indeed we have

$$\begin{aligned} \varphi'(a) &= 2 \log \Gamma_\beta + h_\beta(\delta) - g\left(\frac{1}{2a^2}, 0\right) + \frac{1}{a^2} \partial_q g\left(\frac{1}{2a^2}, 0\right), \\ \varphi''(a) &= -\frac{1}{a^3} \partial_q g\left(\frac{1}{2a^2}, 0\right) - \frac{1}{a^5} \partial_q^2 g\left(\frac{1}{2a^2}, 0\right). \end{aligned}$$

Recall that  $\partial_q g\left(\frac{1}{2a^2}, 0\right) = \tilde{h}_0^{\left(\frac{1}{2a^2}, 0\right)} > 0$  (see (1.4.10) and Section 1.6.3), and  $g$  is convex hence  $\partial_q^2 g\left(\frac{1}{2a^2}, 0\right) \geq 0$ ; therefore  $\varphi''(a) < 0$  for all  $a \in (0, \infty)$ . Besides,  $\varphi'(a)$  can be written

$$\varphi'(a) = 2 \log \Gamma_\beta + h_\beta(\delta) + \frac{1}{2a^2} \tilde{h}_0^{\frac{1}{2a^2}, 0} + \mathcal{L}_\Lambda\left(\tilde{\mathbf{h}}^{\frac{1}{2a^2}, 0}\right). \quad (1.4.97)$$

Recall that  $2 \log \Gamma_\beta + h_\beta(\delta) < 0$  whenever  $(\beta, \delta) \in (\mathcal{C}_{\text{bead}})^\circ$ . Moreover  $\tilde{\mathbf{h}}$  is continuous and  $\tilde{\mathbf{h}}^{0,0} = (0, 0)$ , so  $\varphi'(a) < 0$  for  $a$  sufficiently large. As  $a \searrow 0$ , one has  $\tilde{h}_0^{\frac{1}{2a^2}, 0} \rightarrow \frac{\beta}{2} > 0$  and  $\mathcal{L}_\Lambda\left(\tilde{\mathbf{h}}^{\frac{1}{2a^2}, 0}\right) \geq 0$ ; hence  $\varphi'(a) > 0$  for  $a$  sufficiently small.

This proves that  $\sup_{(0,\infty)} \varphi = \varphi(\tilde{a})$  for some  $\tilde{a} = \tilde{a}(\beta, \delta) \in (0, \infty)$ . Provided that  $a_1$  (resp.  $a_2$ ) is sufficiently small (resp. large), we can assume  $\tilde{a} \in (a_1, a_2)$ . Because  $f$  is strictly concave and  $C^2$  on  $[a_1, a_2]$ , there are constants  $c, c' > 0$  such that

$$-c(a - \tilde{a})^2 \leq \varphi(a) - \varphi(\tilde{a}) \leq -c'(a - \tilde{a})^2, \quad (1.4.98)$$

for all  $a \in [a_1, a_2]$ . Thus we can rewrite (1.4.95) as

$$\begin{aligned} & C_{18} \frac{e^{\beta L}}{L} e^{\sqrt{L}\varphi(\tilde{a})} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} e^{-c\sqrt{L}\left(\frac{N}{\sqrt{L}} - \tilde{a}\right)^2} \\ & \leq Z_{L,\beta,\delta}^{\circ,+} \leq C_{19} \frac{e^{\beta L}}{L} e^{\sqrt{L}\varphi(\tilde{a})} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} e^{-c'\sqrt{L}\left(\frac{N}{\sqrt{L}} - \tilde{a}\right)^2}. \end{aligned} \quad (1.4.99)$$

Finally, we observe that for any  $c > 0$  and  $b_1 < 0 < b_2$ , one has

$$\sum_{\substack{b_1\sqrt{L} \leq N \leq b_2\sqrt{L}, \\ N \in \mathbb{Z}}} e^{-c\frac{N^2}{\sqrt{L}}} \asymp L^{1/4}. \quad (1.4.100)$$

Indeed, let  $R > 0$ , and define  $I_{R,L} = [-RL^{1/4}, RL^{1/4}]$ . We decompose the above sum into  $S_1^{L,\beta,R} + S_2^{L,\beta,R}$ , where

$$S_1^{L,\beta,R} := \sum_{N \in I_{R,L} \cap \mathbb{Z}} e^{-c\frac{N^2}{\sqrt{L}}}, \quad \text{and} \quad S_2^{L,\beta,R} := \sum_{\substack{N \in [b_1\sqrt{L}, b_2\sqrt{L}] \setminus I_{R,L}, \\ N \in \mathbb{Z}}} e^{-c\frac{N^2}{\sqrt{L}}}.$$

Let us first handle  $S_1^{L,\beta,R}$ . With a Riemann sum approximation, we have

$$L^{-1/4} S_1^{L,\beta,R} = L^{-1/4} \sum_{N=-RL^{1/4}}^{RL^{1/4}} e^{-c\left(\frac{N}{L^{1/4}}\right)^2} \xrightarrow{L \rightarrow \infty} \int_{-R}^R e^{-cx^2} dx, \quad (1.4.101)$$

therefore  $S_1^{L,\beta,R} \sim L^{1/4} \int_{-R}^R e^{-cx^2} dx$  as  $L \rightarrow \infty$ . Regarding  $S_2^{L,\beta,R}$ , we have

$$\sum_{\substack{N \in [b_1\sqrt{L}, b_2\sqrt{L}] \setminus I_{R,L}, \\ N \in \mathbb{Z}}} e^{-c\frac{N^2}{\sqrt{L}}} \leq 2 \sum_{N > RL^{1/4}} e^{-c\left(\frac{N}{L^{1/4}}\right)^2},$$

and a Riemann sum approximation yields that

$$\lim_{L \rightarrow \infty} L^{-1/4} \sum_{N > RL^{1/4}} e^{-c\left(\frac{N}{L^{1/4}}\right)^2} = \int_R^\infty e^{-cx^2} dx. \quad (1.4.102)$$

Hence there is  $C_R > 0$  such that  $S_2^{L,\beta,R} \leq C_R L^{1/4}$  for all  $L \geq 1$ . Together with (1.4.101), this concludes the proof of (1.4.100). Plugging (1.4.100) into (1.4.99), this proves Theorem 1.3.2 in  $(\mathcal{AC})^\circ$ .

**Case  $\delta = 0$**

We now prove the theorem when  $\delta = 0$  —then the case  $0 < \delta \leq \tilde{\delta}_c(\beta)$  will be a straightforward consequence of the other two. Let  $\varepsilon > 0$ . Recollecting (1.4.94) and Proposition 1.4.8, and noticing that  $q \mapsto J(\sigma_\beta \partial_q g(q, 0))$  is  $\mathcal{C}^1$  on  $(0, \infty)$ ; we can write the bounds

$$\begin{aligned} & C_{20} \frac{e^{\beta L}}{L} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} \exp\left(\sqrt{L}\varphi\left(\frac{N}{\sqrt{L}}\right) + L^{1/6}\left(\psi\left(\frac{N}{\sqrt{L}}\right) - \varepsilon\right)\right) \\ & \leq Z_{L,\beta,0}^{\circ,+} \leq C_{21} \frac{e^{\beta L}}{L} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} \exp\left(\sqrt{L}\varphi\left(\frac{N}{\sqrt{L}}\right) + L^{1/6}\left(\psi\left(\frac{N}{\sqrt{L}}\right) + \varepsilon\right)\right), \end{aligned} \quad (1.4.103)$$

for  $L$  sufficiently large, where  $\varphi$  is defined in (1.4.96), and we define for all  $a \in (0, \infty)$ ,

$$\psi(a) := a^{1/3} J\left(\sigma_\beta \partial_q g\left(\frac{1}{2a^2}, 0\right)\right) < 0. \quad (1.4.104)$$

Notice that we can drop the constant and polynomial factors by slightly adjusting  $\varepsilon$ . Let us fix some  $R > 0$  and define

$$I_{R,L}^1 := \tilde{a}\sqrt{L} + [-RL^{1/3}, RL^{1/3}] \quad \text{and} \quad I_{R,L}^2 := [a_1\sqrt{L}, a_2\sqrt{L}] \setminus I_{R,L}^1,$$

so (1.4.103) yields

$$e^{\beta L} Q_{L,-\varepsilon}^1 \leq Z_{L,\beta,0}^{\circ,+} \leq e^{\beta L} (Q_{L,\varepsilon}^1 + Q_{L,\varepsilon}^2), \quad (1.4.105)$$

where we define for any  $i \in \{1, 2\}$  and  $\eta \in \{-\varepsilon, \varepsilon\}$ ,

$$Q_{L,\eta}^i := \sum_{N \in I_{R,L}^i} \exp\left(\sqrt{L}\varphi\left(\frac{N}{\sqrt{L}}\right) + L^{1/6}\left(\psi\left(\frac{N}{\sqrt{L}}\right) + \eta\right)\right).$$

Regarding the lower bound, recall that  $\psi$  is  $\mathcal{C}^1$  on  $(0, \infty)$ , so there is some  $C_{22} > 0$  such that

$$\left|\psi\left(\frac{N}{\sqrt{L}}\right) - \psi(\tilde{a})\right| \leq C_{22} R L^{-1/6}, \quad (1.4.106)$$

uniformly in  $L$  sufficiently large and  $N \in I_{R,L}^1$ . Hence for  $L$  large,

$$\begin{aligned} Q_{L,-\varepsilon}^1 & \geq \exp\left(\sqrt{L}\varphi(\tilde{a}) + L^{1/6}(\psi(\tilde{a}) - \varepsilon) - C_{22}R\right) \\ & \quad \times \sum_{N=\tilde{a}\sqrt{L}-L^{1/3}}^{\tilde{a}\sqrt{L}+L^{1/3}} \exp\left(\sqrt{L}\left(\varphi\left(\frac{N}{\sqrt{L}}\right) - \varphi(\tilde{a})\right)\right), \end{aligned} \quad (1.4.107)$$

and recalling (1.4.98) to (1.4.102), the latter sum is of order  $C_{23}L^{1/4} \geq 1$  when  $L \rightarrow \infty$ , which yields the expected lower bound (by slightly adjusting  $\varepsilon$ ).

Similarly to the lower bound just above, we deduce from (1.4.106) and from (1.4.98) to (1.4.102) that

$$Q_{L,\varepsilon}^1 \leq \exp\left(\sqrt{L}\varphi(\tilde{a}) + L^{1/6}(\psi(\tilde{a}) + 2\varepsilon)\right), \quad (1.4.108)$$

for  $L$  sufficiently large. Regarding  $Q_{L,\varepsilon}^2$ , (1.4.98) yields that for all  $N \in I_{R,L}^2$ ,

$$\sqrt{L} \left( \varphi\left(\frac{N}{\sqrt{L}}\right) - \varphi(\tilde{a}) \right) \leq -c'R^2 L^{1/6},$$

Moreover  $\psi(N/\sqrt{L}) + \varepsilon \leq \frac{1}{2} \sup_{[a_1, a_2]} \psi < 0$  for all  $N \in [a_1\sqrt{L}, a_2\sqrt{L}]$ ,  $L \in \mathbb{N}$  provided that  $\varepsilon$  is sufficiently small. Thereby,

$$Q_{L,\varepsilon}^2 \leq (a_2 - a_1)\sqrt{L} \times \exp\left(\sqrt{L}\varphi(\tilde{a}) - c'R^2 L^{1/6}\right).$$

Finally, noticing that  $Q_{L,\varepsilon}^1 \geq Q_{L,-\varepsilon}^1$  and recalling (1.4.107), we have  $Q_{L,\varepsilon}^2/Q_{L,\varepsilon}^1 = o(1)$ , provided that  $R$  satisfies  $-c'R^2 < \inf_{[a_1, a_2]} \psi - \varepsilon$ . Recollecting (1.4.105) this completes the proof of the upper bound, and concludes the case  $\delta = 0$ .

**Case**  $0 < \delta \leq \tilde{\delta}_c(\beta)$

When  $\delta \in (0, \tilde{\delta}(\beta)]$ , we do not provide sharp estimates but we can still claim the following: first,  $\delta \mapsto Z_{L,\beta,\delta}^{\circ,+}$  is clearly non-decreasing, so  $Z_{L,\beta,\delta}^{\circ,+} \geq Z_{L,\beta,0}^{\circ,+}$ , and the lower bound from the case  $\delta = 0$  also holds for any  $\delta > 0$ . Secondly, notice that we have  $E_{N,q}^{\circ,\delta} \leq Z_{\text{wet},N}^{\beta,\delta}$  for all  $q > 0$ ,  $N \in \mathbb{N}$  (see (1.4.12) and (1.A.1)), so Proposition 1.A.1 (ii) implies that there exists  $C_{24} > 0$  (which depends on  $\beta, \delta$ ) such that  $E_{N,q}^{\circ,\delta} \leq C_{24}$  for  $N \in \mathbb{N}$ . Thus (1.4.94) yields

$$Z_{L,\beta,\delta}^{\circ,+} \leq \frac{e^{\beta L}}{L} C_{24} \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} e^{\sqrt{L}\varphi\left(\frac{N}{\sqrt{L}}\right)},$$

where  $\varphi$  is defined in (1.4.96), with  $h_\beta(\delta) = 0$  for all  $\delta \in (0, \tilde{\delta}(\beta)]$ . Then the behavior of the sum above is the same as in the case  $\delta \in (\tilde{\delta}(\beta), \delta_c^\circ(\beta))$ , which yields the same upper bound. This fully concludes the proof of Theorem 1.3.2, subject to Lemmas 1.4.5 and 1.4.9, and Proposition 1.4.3.

### 1.4.7 Proof of Corollary 1.3.3

In this section we prove Corollary 1.3.3, which follows directly from Theorem 1.3.2. Let  $\beta > \beta_c$ . We first focus on the case  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ . Recall the expression of  $\varphi'(a) = \varphi'_{(\beta,\delta)}(a)$  in (1.4.97) and the definition of  $\tilde{a}(\beta, \delta)$ , then notice that  $(\delta, a) \mapsto \varphi'_{(\beta,\delta)}(a)$  is  $\mathcal{C}^1$  on  $(\tilde{\delta}_c(\beta), \delta_c^\circ(\beta)) \times (0, +\infty)$  and  $\varphi''_{(\beta,\delta)}(\tilde{a}(\beta, \delta)) < 0$ . It follows from the implicit function theorem that  $\delta \mapsto \tilde{a}(\beta, \delta)$  is  $\mathcal{C}^1$  on  $(\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ . Hence  $\delta \mapsto \Phi(\beta, \delta)$  is  $\mathcal{C}^1$  on  $(\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$  too.

Let  $\delta \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$  and  $\varepsilon > 0$ , and let us denote for  $\ell \in \Omega_L^{\circ,+}$ ,

$$Q(\ell) := \sum_{k=1}^{N_\ell} \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}},$$

for conciseness. Let  $t > 0$  be such that  $\delta + t \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ , and write with Chernov's bound

$$\begin{aligned} \mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( Q(\ell) \geq (\partial_\delta \Phi(\beta, \delta) + \varepsilon) \sqrt{L} \right) &\leq \mathbf{E}_{L,\beta,\delta}^{\circ,+} \left[ e^{tQ(\ell)} \right] e^{-t(\partial_\delta \Phi(\beta, \delta) + \varepsilon) \sqrt{L}} \\ &\leq \frac{Z_{L,\beta,\delta+t}^{\circ,+}}{Z_{L,\beta,\delta}^{\circ,+}} e^{-t(\partial_\delta \Phi(\beta, \delta) + \varepsilon) \sqrt{L}}. \end{aligned}$$

Applying Theorem 1.3.2 (i), we therefore get that there exists some  $C_{25} > 0$  such that for  $L \geq 1$ ,

$$\mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( Q(\ell) \geq (\partial_\delta \Phi(\beta, \delta) + \varepsilon) \sqrt{L} \right) \leq C_{25} e^{(\Phi(\beta, \delta+t) - \Phi(\beta, \delta) - t \partial_\delta \Phi(\beta, \delta) - t\varepsilon) \sqrt{L}}.$$

Finally,  $\Phi(\beta, \delta + t) - \Phi(\beta, \delta) - t \partial_\delta \Phi(\beta, \delta) = o(t)$  as  $t \rightarrow 0$  (since  $\delta \mapsto \Phi(\beta, \delta)$  is  $\mathcal{C}^1$ ), so we can fix  $t > 0$  such that the r.h.s. above goes to 0 as  $L \rightarrow \infty$ . Similarly, we can write for  $s > 0$  such that  $\delta - s \in (\tilde{\delta}_c(\beta), \delta_c^\circ(\beta))$ ,

$$\begin{aligned} \mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( Q(\ell) \leq (\partial_\delta \Phi(\beta, \delta) - \varepsilon) \sqrt{L} \right) &\leq \frac{Z_{L,\beta,\delta-s}^{\circ,+}}{Z_{L,\beta,\delta}^{\circ,+}} e^{s(\partial_\delta \Phi(\beta, \delta) - \varepsilon) \sqrt{L}} \\ &\leq C_{26} e^{(\Phi(\beta, \delta-s) - \Phi(\beta, \delta) + s \partial_\delta \Phi(\beta, \delta) - s\varepsilon) \sqrt{L}}, \end{aligned}$$

for some  $C_{26} > 0$ , where we also applied Chernov's bound and Theorem 1.3.2 (i). We can then choose  $s > 0$  such that this term also goes to 0 as  $L \rightarrow \infty$ . This concludes the proof of Corollary 1.3.3-(i).

Regarding the case  $\delta \in [0, \tilde{\delta}_c(\beta))$ , let  $K > 0$ ,  $\varepsilon > 0$  and  $t > 0$  be such that  $\delta + t < \tilde{\delta}_c(\beta)$ . Using once again Chernov's bound and Theorem 1.3.2 (ii), there exist  $C_{27} > 0$  and  $L_0 \in \mathbb{N}$  such that for  $L \geq L_0$ ,

$$\begin{aligned} \mathbf{P}_{L,\beta,\delta}^{\circ,+} \left( Q(\ell) \geq KL^{1/6} \right) &\leq \mathbf{E}_{L,\beta,\delta}^{\circ,+} \left[ e^{tQ(\ell)} \right] e^{-tKL^{1/6}} \\ &\leq \frac{Z_{L,\beta,\delta+t}^{\circ,+}}{Z_{L,\beta,\delta}^{\circ,+}} e^{-tKL^{1/6}} \leq C_{27} L^{-3/4} e^{-(\Psi(\beta) + \varepsilon - tK)L^{1/6}}. \end{aligned}$$

This goes to 0 as  $L \rightarrow \infty$  provided that  $K$  has been fixed sufficiently large, and concludes the proof of Corollary 1.3.3-(ii).

## 1.5 Proof of Theorem 1.2.2 and Proposition 1.3.1

### 1.5.1 Proof of Theorem 1.2.2

The proof of Theorem 1.2.2 also relies on the random walk representation introduced in [130] and adapted in the proof of Theorem 1.3.2 (see Section 1.4), but it is much less involved. We divide the proof into 4 steps. First, we prove that the free energy is not



changed when we additionally constrain the polymer to end on the horizontal axis (that is  $\sum_{i=1}^N \ell_i = 0$ ) —in particular the *constrained* partition function  $Z_{L,\beta,\delta}^{+,c}$  is super-multiplicative, which implies the well-posedness of  $f$ . Next, we adapt to the present model the random-walk representation of IPDSAW. As in Section 1.4, we derive a probabilistic representation of the partition function by rewriting, for every  $N \leq L$ , the contribution to the partition function of those trajectories made of  $N$  stretches, in terms of two auxiliary (coupled) random walks  $S$  and  $I$ . Then we compute the generating function of the partition function  $Z_{L,\beta,\delta}^{+,c}$ . With our random walk representation, we can rewrite it as the partition function of a wetting model for two independent random walks, with the in-between area constraint of  $S$  and  $I$  becoming an in-between area penalization in the generating function. This allows us to characterize  $\tilde{f}$  in terms of the free energy of this “coupled, in-between area penalized” wetting model. Finally we place ourselves on the critical curve between  $\mathcal{C}$  and  $\mathcal{E}$  where the area penalization vanishes, and we apply well-known results on the wetting model to derive the equation of the curve.

### Step 1: constraining the partition function

Let us define the *constrained partition function*:

$$Z_{L,\beta,\delta}^{+,c} := \sum_{N=1}^L \sum_{\substack{\ell \in \mathcal{L}_{N,L}^+ \\ \ell_0 = \ell_{N+1} = 0}} e^{H(\ell)} \mathbf{1}_{\{\sum_{i=1}^N \ell_i = 0\}}. \quad (1.5.1)$$

Notice that this partition function is super-multiplicative: indeed for any  $L_1, L_2 \geq 0$ , we bound  $Z_{L_1+L_2,\beta,\delta}^{+,c}$  from below by constraining it to touch the axis between the  $(L_1 - 1)$ -th and  $L_1$ -th monomers; then we notice that the contribution to  $H(\ell)$  from self-touching between the two parts of the polymer (before and after the segment  $(L_1 - 1, L_1)$ ) are non-negative; and we separate the trajectory in two at the  $L_1$ -th monomer (recall that all our trajectories end with a horizontal step) to finally write

$$Z_{L_1+L_2,\beta,\delta}^{+,c} \geq Z_{L_1,\beta,\delta}^{+,c} Z_{L_2,\beta,\delta}^{+,c}. \quad (1.5.2)$$

Hence  $\frac{1}{L} \log Z_{L,\beta,\delta}^{+,c}$  converges as  $L \rightarrow \infty$ , by Fekete’s lemma.

We now claim that  $Z_{L,\beta,\delta}^+$  and  $Z_{L,\beta,\delta}^{+,c}$  are comparable. More precisely for  $L \in \mathbb{N}$ ,

$$\frac{1}{L^2} \left( Z_{L,\beta,\delta}^+ \right)^2 \leq Z_{2L,\beta,\delta}^{+,c} \leq Z_{2L,\beta,\delta}^+. \quad (1.5.3)$$

The upper bound is straightforward. For the lower bound, we constrain  $Z_{2L,\beta,\delta}^{+,c}$  to make a horizontal step between the  $(L - 1)$ -th and  $L$ -th monomers, and sum over its possible heights. Writing  $z_{L-1} = (x_{L-1}, y_{L-1})$  and  $z_L = (x_L, y_L)$  the coordinates of those monomers, we have

$$Z_{2L+2,\beta,\delta}^{+,c} \geq \sum_{y=0}^{L-1} Z_{2L,\beta,\delta}^+(y_{L-1} = y_L = y, x_L = 1 + x_{L-1}).$$

We then separate the trajectory at the  $L_1$ -th monomer, so that the first half gives the term  $Z_{L,\beta,\delta}^+(y_L = y)$  (recall that the last step is always horizontal). As for the second half, we shift the last horizontal step to just before  $z_L$  (which costs at most a factor  $e^\delta$ ), then we reverse the trajectory horizontally, to obtain that it is bounded from below by  $Z_{L,\beta,\delta}^+(y_L = y)$ . Therefore,

$$Z_{2L,\beta,\delta}^{+,c} \geq \sum_{y=0}^{L-1} \left( Z_{L,\beta,\delta}^+(y_L = y) \right)^2 \geq \left( \sup_{0 \leq y \leq L-1} Z_{L,\beta,\delta}^+(y_L = y) \right)^2,$$

and we conclude the proof of (1.5.3) by writing  $Z_{L,\beta,\delta}^+ \leq L \sup_{0 \leq y \leq L-1} Z_{L,\beta,\delta}^+(y_L = y)$ . In particular, this implies that

$$f(\beta, \delta) = \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\beta,\delta}^+ = \lim_{L \rightarrow \infty} \frac{1}{L} \log Z_{L,\beta,\delta}^{+,c}, \quad (1.5.4)$$

is well defined, and that we can replace  $Z_{L,\beta,\delta}^+$  with  $Z_{L,\beta,\delta}^{+,c}$  to prove Theorem 1.2.2.

### Step 2: a random-walk representation

We provide a probabilistic description of the constrained partition function  $Z_{L,\beta,\delta}^{+,c}$  as in Section 1.4 —which also applies to the free counterpart  $Z_{L,\beta,\delta}^+$ . Recalling (1.5.1), (1.2.2) and (1.4.1), we can rewrite the constrained partition function similarly to (1.4.2) to obtain

$$Z_{L,\beta,\delta}^{+,c} = e^{\beta L} \sum_{N=1}^L e^{-\beta N} \sum_{\substack{\ell \in \mathcal{L}_{N,L}^+ \\ \ell_0 = \ell_{N+1} = 0}} e^{-\frac{\beta}{2} \sum_{i=0}^N |\ell_i + \ell_{i+1}|} e^{\delta \sum_{k=1}^N \mathbf{1}_{\{\sum_{i=1}^k \ell_i = 0\}}} \mathbf{1}_{\{\sum_{i=1}^N \ell_i = 0\}}. \quad (1.5.5)$$

Recall the definition of  $\mathbf{P}_\beta$  in (1.2.8): similarly to (1.4.3), we consider two independent one-dimensional random walks  $S := (S_i)_{i \geq 0}$  and  $I := (I_i)_{i \geq 0}$  starting from 0 and such that  $(S_{i+1} - S_i)_{i \geq 0}$  and  $(I_{i+1} - I_i)_{i \geq 0}$  are i.i.d. sequences of random variables of law  $\mathbf{P}_\beta$ . We notice that for every  $\ell \in \mathcal{L}_{N,L}^+$  (with  $\ell_0 = \ell_{N+1} = 0$ ) the first factor in the second sum in (1.5.5) satisfies

$$e^{-\frac{\beta}{2} \sum_{i=0}^N |\ell_i + \ell_{i+1}|} = c_\beta^{N+1} \mathbf{P}_\beta \left( S_k = \sum_{i=0}^{2k-1} \ell_i, \forall k \leq N_S \right) \mathbf{P}_\beta \left( I_k = \sum_{i=0}^{2k} \ell_i, \forall k \leq N_I \right), \quad (1.5.6)$$

with  $N_S := \lfloor N/2 \rfloor + 1$  and  $N_I := \lceil N/2 \rceil$ . Hence we define  $S$  and  $I$  as in Section 1.4.1. Similarly to (1.4.4), define

$$G_N(S, I) := \sum_{k=1}^{N_I} |I_k - S_k| + \sum_{k=1}^{N_S} |S_k - I_{k-1}|, \quad (1.5.7)$$

and our definition of  $S, I$  implies  $\sum_{i=1}^N |\ell_i| = G_N(S, I)$ . Finally the constraint  $\sum_{i=1}^N \ell_i = 0$  is equivalent to  $S_{N_S} = I_{N_I} = 0$  (recall  $\ell_{N+1} = 0$ ).

Defining  $\tilde{Z}_{L,\beta,\delta}^{+,c} := \frac{1}{c_\beta} e^{-\beta L} Z_{L,\beta,\delta}^{+,c}$ , and plugging (1.5.6) into (1.5.5), we obtain

$$\tilde{Z}_{L,\beta,\delta}^{+,c} = \sum_{N=1}^L (\Gamma_\beta)^N \mathbf{E}_\beta \left[ e^{\delta \sum_{k=1}^{N_S} \mathbf{1}_{\{s_k=0\}}} e^{\delta \sum_{k=1}^{N_I} \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{G_N(S,I)=L-N\}} \mathbf{1}_{\{S \in B_{N_S}^{0,+}\}} \mathbf{1}_{\{I \in B_{N_I}^{0,+}\}} \right], \quad (1.5.8)$$

where we recall  $\Gamma_\beta = c_\beta e^{-\beta}$ ,  $G_N(S, I)$  is defined in (1.5.7), and  $B_n^{0,+}$  is the set of non-negative trajectories of length  $n$  ending on the horizontal axis (see (1.2.10)).

### Step 3: the generating function of $\tilde{Z}_{L,\beta,\delta}^{+,c}$

Recalling (1.5.4), the excess free energy satisfies

$$\tilde{f}(\beta, \delta) = \sup \left\{ \gamma \geq 0; \sum_{L \geq 1} \tilde{Z}_{L,\beta,\delta}^{+,c} e^{-\gamma L} = +\infty \right\}, \quad (1.5.9)$$

and  $\tilde{f}(\beta, \delta) = 0$  if this set is empty. Let us compute the generating function of the tilted partition function. Recalling (1.5.8) and inverting the sums over  $L$  and  $N$ , we obtain

$$\sum_{L \geq 1} \tilde{Z}_{L,\beta,\delta}^{+,c} e^{-\gamma L} = \sum_{N \geq 1} (\Gamma_\beta e^{-\gamma})^N Q_N^{\beta,\delta,\gamma}, \quad (1.5.10)$$

where we define

$$Q_N^{\beta,\delta,\gamma} := \mathbf{E}_{\beta,0} \left[ e^{-\gamma G_N(S,I)} e^{\delta \sum_{k=1}^{N_S} \mathbf{1}_{\{s_k=0\}}} e^{\delta \sum_{k=1}^{N_I} \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{S \in B_{N_S}^{0,+}, I \in B_{N_I}^{0,+}\}} \right], \quad (1.5.11)$$

which is the partition function of a coupled wetting model, with an additional “in-between area” penalization. Note that  $(Q_N^{\beta,\delta,\gamma})_{N \geq 1}$  is super-multiplicative, therefore we can define the free energy of this model:

$$h_\beta(\delta, \gamma) := \lim_{N \rightarrow \infty} \frac{1}{N} \log Q_N^{\beta,\delta,\gamma}, \quad (1.5.12)$$

Notice that  $h_\beta(\delta, \gamma) \leq \delta$  (in particular  $h_\beta$  is finite),  $h_\beta$  is non-increasing in  $\gamma$  and non-decreasing in  $\delta$ , and it is continuous.

Recollecting (1.5.9), we conclude that  $\tilde{f}(\beta, \delta)$  is the only positive solution (if it exists) in  $\gamma$  of:

$$\log \Gamma_\beta - \gamma + h_\beta(\delta, \gamma) = 0, \quad (1.5.13)$$

and  $\tilde{f}(\beta, \delta) = 0$  otherwise (this equation has at most one solution because  $\gamma \mapsto -\gamma + h_\beta(\delta, \gamma)$  is decreasing).

**Step 4: characterizing the critical curve**

Let  $(\beta, \delta) \in \bar{\mathcal{C}} \cap \bar{\mathcal{E}}$  be a point of the critical curve: then  $\tilde{f}(\beta, \delta) = 0$  (because  $\tilde{f}$  is continuous) and  $\log \Gamma_\beta + h_\beta(\delta, 0) = 0$  (because (1.5.13) is continuous in  $\gamma \searrow 0$ ). In particular there is no area constraint in  $Q_N^{\beta, \delta, 0}$ , and because  $S$  and  $I$  are independent we can uncouple them and write

$$Q_N^{\beta, \delta, 0} = Z_{\text{wet}, N_S}^{\beta, \delta} \times Z_{\text{wet}, N_I}^{\beta, \delta}, \quad (1.5.14)$$

where  $Z_{\text{wet}, N}^{\beta, \delta}$ ,  $N \geq 1$  is the partition function of a wetting model (see (1.A.1)). By applying  $\frac{1}{N} \log$  to (1.5.14), we notice that  $h_\beta(\delta, 0)$  matches exactly the free energy of the wetting model:  $h_\beta(\delta, 0) = h_\beta(\delta)$  (see (1.2.9)).

The asymptotic behavior of  $Z_{\text{wet}, N}^{\beta, \delta}$  is already well-known —see Proposition 1.A.1. Thus we can finally conclude the proof of Theorem 1.2.2. Recall that  $(\beta, \delta) \in \bar{\mathcal{C}} \cap \bar{\mathcal{E}}$  implies  $\log \Gamma_\beta + h_\beta(\delta) = 0$ , where  $\Gamma_\beta$  is decreasing in  $\beta$ , and  $\beta_c$  is the only solution of  $\Gamma_\beta = 1$ . Therefore,

- if  $\beta < \beta_c$ , then  $\log \Gamma_\beta + h_\beta(\delta) \geq \log \Gamma_\beta > 0$  so  $(\beta, \delta) \notin \bar{\mathcal{C}} \cap \bar{\mathcal{E}}$ ,
- if  $\beta = \beta_c$ , then  $(\beta, \delta) \in \bar{\mathcal{C}} \cap \bar{\mathcal{E}}$  if and only if  $h_\beta(\delta) = 0$ , that is  $\delta \leq \tilde{\delta}_c(\beta_c)$ ,
- if  $\beta > \beta_c$ , then there is a unique solution in  $\delta \geq \tilde{\delta}_c(\beta)$  to  $h_\beta(\delta) = -\log \Gamma_\beta$ , which we note  $\delta_c(\beta)$ .

Moreover, recall that the l.h.s. of (1.5.13) is decreasing and continuous in  $\gamma$ , implying that  $\tilde{f}(\beta, \delta) = 0$  if and only if  $\log \Gamma_\beta + h_\beta(\delta) \leq 0$ , that is  $\beta \geq \beta_c$  and  $\delta \leq \delta_c(\beta)$ ; which fully characterizes  $\mathcal{C}$ . Finally, the analytic expression of  $\delta_c(\beta)$  follows directly from (1.2.11) and (1.A.4) by solving a quadratic equation; details are left to the reader.

**1.5.2 Proof of Proposition 1.3.1**

This is very similar to Theorem 1.2.2. The single-bead partition  $Z_{L, \beta, \delta}^{\circ, +}$  function is already constrained to return to 0, hence it is super-multiplicative (and  $f^\circ$  is well-posed) and we do not need to replicate Step 1. The random walk representation is already laid out in Proposition 1.4.7 —notice that the main difference with (1.5.8) is the constraint  $\{S \succ I\}$  (recall (1.4.5)), in particular the random walk  $S$  cannot touch the wall. Thereby the generating function of  $\tilde{Z}_{L, \beta, \delta}^{\circ, +}$  can be written (similarly to (1.5.10)) as

$$\sum_{L \geq 1} \tilde{Z}_{L, \beta, \delta}^{\circ, +} e^{-\gamma L} = \sum_{N \geq 1} (\Gamma_\beta e^{-\gamma})^{2N} R_N^{\beta, \delta, \gamma}, \quad (1.5.15)$$

where we define

$$R_N^{\beta, \delta, \gamma} := \mathbf{E}_{\beta, 0} \left[ e^{-2\gamma(A_{N+1}(S) - A_N(I))} e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{S \succ I\}} \mathbf{1}_{\{S \in B_{N+1}^{\circ, +}, I \in B_N^{\circ, +}\}} \right]. \quad (1.5.16)$$

$(R_N^{\beta, \delta, \gamma})_{N \geq 1}$  is super-multiplicative, thereby we define its free energy

$$\kappa_\beta(\delta, \gamma) := \lim_{N \rightarrow \infty} \frac{1}{N} \log R_N^{\beta, \delta, \gamma},$$

and we note that  $\tilde{f}^\circ(\beta, \delta)$  is the only positive solution (if it exists) in  $\gamma$  of

$$2 \log \Gamma_\beta + 2\gamma + \kappa_\beta(\delta, \gamma) = 0 ,$$

and  $\tilde{f}^\circ(\beta, \delta) = 0$  otherwise.

Let  $(\beta, \delta) \in \partial \mathcal{C}_{\text{bead}}$  be on the critical curve (which implies  $\tilde{f}^\circ(\beta, \delta) = 0$  and  $2 \log \Gamma_\beta + \kappa_\beta(\delta, 0) = 0$ ). We notice  $R_N^{\beta, \delta, \gamma} \leq Z_{\text{wet}, N}^{\beta, \delta}$ , so  $\kappa_\beta(\delta, \gamma) \leq h_\beta(\delta)$ , and we will now compute a lower bound on  $R_N^{\beta, \delta} := R_N^{\beta, \delta, 0}$  to deduce  $\kappa_\beta(\delta, 0) = h_\beta(\delta)$ .

Let  $\alpha > 0$  and recall (1.5.16) with  $\gamma = 0$ . Constraining  $I_k$  to remain below  $\lfloor \alpha \sqrt{N} \rfloor$  for all  $1 \leq k < N$ , and constraining  $S$  to  $S_1 = S_N = \lfloor \alpha \sqrt{N} \rfloor + 1$ , and  $S_k \geq \lfloor \alpha \sqrt{N} \rfloor + 1$  for all  $2 \leq k < N$ , we obtain the lower bound

$$\begin{aligned} R_N^{\beta, \delta} &\geq \mathbf{E}_{\beta, 0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I_k \leq \lfloor \alpha \sqrt{N} \rfloor, \forall 1 \leq k < N\}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \right] \\ &\quad \times \mathbf{P}_{\beta, 0} \left( S_1 = S_N = \lfloor \alpha \sqrt{N} \rfloor + 1, S_{N+1} = 0, S_k \geq \lfloor \alpha \sqrt{N} \rfloor + 1, \forall 2 \leq k < N \right) \\ &\geq \mathbf{E}_{\beta, 0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I_k \leq \lfloor \alpha \sqrt{N} \rfloor, \forall 1 \leq k < N\}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \right] \left( \frac{e^{-\frac{\beta}{2} \lfloor \alpha \sqrt{N} \rfloor + 1}}{c_\beta} \right)^2 \mathbf{P}_{\beta, 0} \left( B_{N-1}^{0,+} \right), \end{aligned}$$

where we can estimate  $\mathbf{P}_{\beta, 0}(B_{N-1}^{0,+}) \asymp N^{-3/2}$  with (1.6.1) below. Moreover we notice that

$$(x_k)_{k=1}^N \mapsto e^{\delta \sum_{k=1}^N \mathbf{1}_{\{x_k=0\}}}, \quad \text{and} \quad (x_k)_{k=1}^N \mapsto \mathbf{1}_{\{x_k \leq \lfloor \alpha \sqrt{N} \rfloor, \forall 1 \leq k < N\}},$$

are both bounded, non-increasing functions on  $B_N^{0,+}$ , hence we can apply an FKG inequality (see Proposition 1.B.1) to claim

$$\begin{aligned} &\mathbf{E}_{\beta, 0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{I_k=0\}}} \mathbf{1}_{\{I_k \leq \lfloor \alpha \sqrt{N} \rfloor, \forall 1 \leq k < N\}} \mathbf{1}_{\{I \in B_N^{0,+}\}} \right] \\ &\geq Z_{\text{wet}, N}^{\beta, \delta} \times \mathbf{P}_{\beta, 0} \left( I_k \leq \lfloor \alpha \sqrt{N} \rfloor, \forall 1 \leq k < N \mid I \in B_N^{0,+} \right), \end{aligned} \tag{1.5.17}$$

where we recall that  $Z_{\text{wet}, N}^{\beta, \delta}$ ,  $N \geq 1$  is the partition function of a wetting model (defined in (1.A.1)), and the second factor in the r.h.s. is bounded from below by some positive constant provided that  $\alpha$  is large (see (1.6.2) below).

All this implies  $\kappa_\beta(\delta, 0) \geq h_\beta(\delta)$ , so  $\kappa_\beta(\delta, 0) = h_\beta(\delta)$ . Therefore we conclude the proof of Proposition 1.3.1 similarly to Theorem 1.2.2.

## 1.6 Proofs of technical results

In this section we prove Lemmas 1.4.5 and 1.4.9, then the much more involved Proposition 1.4.3.

### 1.6.1 Proof of Lemma 1.4.5

Before proving Lemma 1.4.5, we provide a useful estimate on non-negative random walk bridges proven in [47]. Recall the definition of  $B_n^{y,+}$  in (1.2.10), and that  $\mathbf{P}_{\beta,x}$  is the law of a random walk with increments distributed as  $\mathbf{P}_\beta$  (see (1.2.8)) starting from  $x \geq 0$ . One has

$$\mathbf{P}_{\beta,0}(B_n^{x,+}) = \mathbf{P}_{\beta,x}(B_n^{0,+}) \asymp \frac{\max(x, 1)}{n^{3/2}}, \quad (1.6.1)$$

uniformly in  $n \geq 1$  and  $0 \leq x \leq C_1\sqrt{n}$  for any  $C_1 > 0$ . The first identity is obtained by reversing the walk in time (notice that  $\mathbf{P}_\beta$  is symmetric), and the asymptotic behavior is derived in [47, (4.3)].

It is also proven in [47, Corol. 2.5] that a properly rescaled centered random walk with finite variance conditioned to remain non-negative and to finish in 0 converges (in distribution) to a Brownian excursion. The maximum being a continuous function on  $\mathcal{C}([0, 1], \mathbb{R})$ , we thereby deduce that for any  $\eta > 0$ , there exist  $C_A > 0$  and  $N_0 \in \mathbb{N}$  such that

$$\mathbf{P}_{\beta,0}\left(\max_{1 \leq k \leq N}(X_k) > C_A\sqrt{N} \mid X \in B_N^{0,+}\right) \leq \eta, \quad (1.6.2)$$

for all  $N \geq N_0$ .

*Proof of Lemma 1.4.5.* Notice that  $\max_{1 \leq k \leq N}(X_k) \leq C_A\sqrt{N}$  implies that  $A_N(X) \leq C_A N^{3/2}$ . Thus we have the lower bound,

$$\begin{aligned} & \mathbf{P}_{\beta,0}\left(\forall 1 \leq k < N, X_k \leq \tilde{f}_I(k); A_N(X) \leq C_A N^{3/2} \mid X \in B_N^{0,+}\right) \\ & \geq 1 - \mathbf{P}_{\beta,0}\left(\exists 1 \leq k < N, X_k > \tilde{f}_I(k); \max_{1 \leq k \leq N}(X_k) \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) \\ & \quad - \mathbf{P}_{\beta,0}\left(\max_{1 \leq k \leq N}(X_k) > C_A\sqrt{N} \mid X \in B_N^{0,+}\right). \end{aligned} \quad (1.6.3)$$

Recalling (1.6.2), the last term in (1.6.3) is not larger than some  $\eta > 0$  arbitrarily small. Regarding the other term, we write

$$\begin{aligned} & \mathbf{P}_{\beta,0}\left(\exists 1 \leq k < N, X_k > \tilde{f}_I(k); \max_{1 \leq k \leq N}(X_k) \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) \\ & \leq 2 \sum_{k=1}^{\lceil N/2 \rceil} \mathbf{P}_{\beta,0}\left(\tilde{f}_I(k) < X_k \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right), \end{aligned} \quad (1.6.4)$$

where we reversed in time all terms with index larger than  $\lceil N/2 \rceil$ . For any  $k \leq \lceil N/2 \rceil$ , we partition the  $k$ -th term over possible values of  $X_k$ , then we apply Markov property at time  $k$  to write:

$$\mathbf{P}_{\beta,0}\left(\tilde{f}_I(k) < X_k \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) = \sum_{z=\tilde{f}_I(k)+1}^{C_A\sqrt{N}} \mathbf{P}_{\beta,0}(B_k^{z,+}) \frac{\mathbf{P}_{\beta,z}(B_{N-k}^{0,+})}{\mathbf{P}_{\beta,0}(B_N^{0,+})}. \quad (1.6.5)$$

Recalling (1.6.1), we can bound  $\mathbf{P}_{\beta,z}(B_{N-k}^{0,+})/\mathbf{P}_{\beta,0}(B_N^{0,+})$  from above by  $C_2 \times z$  with  $C_2 > 0$  a constant uniform in  $z \leq C_A\sqrt{N}$  (recall that  $N-k \geq \frac{N}{2}$ ). Hence, (1.6.5) becomes

$$\mathbf{P}_{\beta,0}\left(\tilde{f}_I(k) < X_k \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) \leq C_2 \mathbf{E}_{\beta,0}\left[X_k \mathbf{1}_{\{\tilde{f}_I(k) < X_k \leq C_A\sqrt{N}\}} \mathbf{1}_{\{X_i \geq 0, \forall i \leq k\}}\right].$$

Recall the definition of  $\tilde{f}_I$  from (1.4.13). Noticing that for any  $0 < \lambda < \beta/2$ ,  $\mathbf{E}_{\beta,0}[e^{\lambda X_k}] = e^{k\mathcal{L}(\lambda)}$  and  $\mathbf{E}_{\beta,0}[X_k e^{\lambda X_k}] = \partial_\lambda \mathbf{E}_{\beta,0}[e^{\lambda X_k}] = k\mathcal{L}'(\lambda)e^{k\mathcal{L}(\lambda)}$ , and choosing  $\lambda$  sufficiently small so that  $\mathcal{L}(\lambda) < \lambda\gamma$  (recall  $\gamma > 0$  and  $\mathcal{L}(\lambda) \sim \lambda^2\sigma_\beta^2/2$  as  $\lambda \searrow 0$ ), we obtain with Markov's inequality,

$$\begin{aligned} \mathbf{P}_{\beta,0}\left(\tilde{f}_I(k) < X_k \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) &\leq C_2 \mathbf{E}_{\beta,0}\left[X_k e^{\lambda(X_k - \gamma k - K_I)}\right] \\ &\leq C_3 k e^{-\lambda K_I} e^{-k(\lambda\gamma - \mathcal{L}(\lambda))}. \end{aligned}$$

The r.h.s. being summable in  $k$ , there exists  $C_4 > 0$  such that,

$$\sum_{k=1}^{\lceil N/2 \rceil} \mathbf{P}_{\beta,0}\left(\tilde{f}_I(k) < X_k \leq C_A\sqrt{N} \mid X \in B_N^{0,+}\right) \leq C_4 e^{-\lambda K_I}, \quad (1.6.6)$$

uniformly in  $N \geq N_0$ . Assuming  $K_I$  is sufficiently large, this term is smaller than any fixed  $\eta > 0$ . Recollecting (1.6.3), (1.6.2), (1.6.4) and (1.6.6) and fixing  $\eta = \varepsilon/3$ , this concludes the proof.  $\square$

## 1.6.2 Proof of Lemma 1.4.9

Let us define for any  $x, m \in \mathbb{N}_0$  some  $\alpha > 0$ ,

$$D_{x,m} := \mathbf{E}_{\beta,x}\left[e^{-\alpha A_m(I)} \mid I \in B_m^+\right]. \quad (1.6.7)$$

Let us prove that  $x \mapsto D_{x,m}$  is non-increasing for any  $\alpha > 0$ ,  $m \in \mathbb{N}_0$  by induction on  $m$ . Notice that the proof also holds when conditioning by  $B_m^{0,+}$  instead of  $B_m^+$  without further changes (this is required in the proof of Claim 1.4.10 in Section 1.4.4).

*Proof of Lemma 1.4.9.* When  $m = 0$ ,  $D_{x,0} = e^{-\alpha x}$  for all  $x \geq 0$ , which is non-increasing. Let  $m \in \mathbb{N}$  and assume  $x \mapsto D_{x,m-1}$  is non-increasing. For any  $x \geq 0$  and  $\alpha > 0$ , one has by Markov's property,

$$\begin{aligned} \mathbf{E}_{\beta,x}\left[e^{-\alpha A_m(I)} \mid B_m^+\right] &= \sum_{y \geq 0} \mathbf{E}_{\beta,x}\left[e^{-\alpha x} e^{-\alpha \sum_{k=1}^m I_k} \mathbf{1}_{\{I_1=y\}} \mid B_m^+\right] \\ &= e^{-\alpha x} \sum_{y \geq 0} \mathbf{E}_{\beta,y}\left[e^{-\alpha A_{m-1}(I)} \mid B_{m-1}^+\right] R_x(y), \end{aligned} \quad (1.6.8)$$

where

$$R_x(y) := \frac{\mathbf{P}_{\beta,x}(I_1 = y) \mathbf{P}_{\beta,y}(B_{m-1}^+)}{\mathbf{P}_{\beta,x}(B_m^+)} = \mathbf{P}_{\beta,x}(I_1 = y \mid B_m^+). \quad (1.6.9)$$

For any  $x, y \geq 0$ , let  $\bar{R}_x(y) := \sum_{t \geq y} R_x(t)$ , so (1.6.8) becomes

$$\begin{aligned} e^{\alpha x} \mathbf{E}_{\beta, x} \left[ e^{-\alpha A_m(I)} \middle| B_m^+ \right] &= \sum_{y \geq 0} \mathbf{E}_{\beta, y} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \left( \bar{R}_x(y) - \bar{R}_x(y+1) \right) \\ &= \bar{R}_x(0) \mathbf{E}_{\beta, 0} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \\ &\quad + \sum_{y \geq 1} \bar{R}_x(y) \left( \mathbf{E}_{\beta, y} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] - \mathbf{E}_{\beta, y-1} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \right). \end{aligned} \quad (1.6.10)$$

Recall that we assumed that  $y \mapsto D_{y, m-1}$  is non-increasing, so

$$\mathbf{E}_{\beta, y} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] - \mathbf{E}_{\beta, y-1} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \leq 0,$$

for all  $y \geq 1$ .

Moreover we claim that  $\bar{R}_{x+1}(y) \geq \bar{R}_x(y)$  for all  $x, y, m \geq 0$ . To prove this, it suffices to notice that for all  $x, m \in \mathbb{N}_0$ , and for any trajectory  $(s_k)_{k=1}^m \in \mathbb{Z}^m$ ,

$$\mathbf{P}_{\beta, x+1} \left( (I_k)_{k=1}^m = (s_k)_{k=1}^m \right) = e^{\pm \frac{\beta}{2}} \mathbf{P}_{\beta, x} \left( (I_k)_{k=1}^m = (s_k)_{k=1}^m \right),$$

where the sign  $\pm$  is  $+$  if  $s_1 > x$  and  $-$  otherwise; then to distinguish the cases  $y \geq x+1$  (where the proof is instantaneous), and  $y \leq x+1$  (where it is obtained by induction on  $0 \leq y \leq x+1$  for a fixed  $x \in \mathbb{N}_0$ ).

Noticing also that  $\bar{R}_x(0) = \bar{R}_{x+1}(0) = 1$  for all  $x \geq 0$ , we obtain the lower bound

$$\begin{aligned} e^{\alpha x} \mathbf{E}_{\beta, x} \left[ e^{-\alpha A_m(I)} \middle| B_m^+ \right] &\geq \bar{R}_{x+1}(0) \mathbf{E}_{\beta, 0} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \\ &\quad + \sum_{y \geq 1} \bar{R}_{x+1}(y) \left( \mathbf{E}_{\beta, y} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] - \mathbf{E}_{\beta, y-1} \left[ e^{-\alpha A_{m-1}(I)} \middle| B_{m-1}^+ \right] \right), \end{aligned}$$

and the identity (1.6.10) finally yields

$$e^{\alpha x} \mathbf{E}_{\beta, x} \left[ e^{-\alpha A_m(I)} \middle| B_m^+ \right] \geq e^{\alpha(x+1)} \mathbf{E}_{\beta, x+1} \left[ e^{-\alpha A_m(I)} \middle| B_m^+ \right],$$

which concludes the induction.  $\square$

### 1.6.3 Proof of Proposition 1.4.3

This proposition is an improvement of [52, Prop. 2.5], where a polynomial lower bound is displayed for the probability that an  $n$ -step walk remains positive, comes back to 0 at time  $n$  and enclose an area  $qn^2$ . To improve this result, we recall some tools introduced in [52] and we keep in mind that all the upcoming claims are proven in [52]. Recall that  $\mathcal{L}(h)$  is the logarithmic moment generating function of the increments of the random walk  $X$ , see (1.3.8). To lighten upcoming formulae, let us define for any  $n \in \mathbb{N}$ ,

$$\Lambda_n := \left( A_n(X)/n, X_n \right), \quad (1.6.11)$$



(notice that the area is normalized by  $n$  in this definition), as well as the parallelograms

$$\mathcal{D}_{\beta,n} := \left\{ (h_0, h_1) \in \mathbb{R}^2; |h_1| < \beta/2, |(1 - 1/n)h_0 + h_1| < \beta/2 \right\}. \quad (1.6.12)$$

For any  $\mathbf{h} \in \mathcal{D}_{\beta,n}$ , we define the tilted law:

$$\frac{d\mathbf{P}_{n,\mathbf{h}}}{d\mathbf{P}_{\beta,0}}(X) = e^{\mathbf{h} \cdot \Lambda_n - \mathcal{L}_{\Lambda_n}(\mathbf{h})}, \quad \mathcal{L}_{\Lambda_n}(\mathbf{h}) := \log \mathbf{E}_{\beta,0}[e^{\mathbf{h} \cdot \Lambda_n}], \quad (1.6.13)$$

where  $\mathbf{h} \cdot \Lambda_n := h_0 A_n/n + h_1 X_n$  denotes the scalar product. Note that

$$\mathbf{h} \cdot \Lambda_n = \sum_{k=1}^n \left( (1 - k/n)h_0 + h_1 \right) (X_k - X_{k-1}), \quad (1.6.14)$$

so this change of measure is equivalent to tilt all increments of  $X$  independently, with an intensity depending on  $k$ . For any  $q, p \in \mathbb{R}$  and  $n \in \mathbb{N}$ , let  $\mathbf{h}_n^{q,p} = (h_{n,0}^{q,p}, h_{n,1}^{q,p})$  be the unique solution in  $\mathbf{h} \in \mathcal{D}_{\beta,n}$  of the equation:

$$\mathbf{E}_{n,\mathbf{h}} \left[ \frac{1}{n} \Lambda_n \right] = \nabla \left[ \frac{1}{n} \mathcal{L}_{\Lambda_n} \right] (\mathbf{h}) = (q, p), \quad (1.6.15)$$

where it is proven in [52, Lem. 5.4] that  $\nabla[\frac{1}{n}\mathcal{L}_{\Lambda_n}]$  is a  $\mathcal{C}^1$  diffeomorphism from  $\mathcal{D}_{\beta,n}$  to  $\mathbb{R}^2$ . Notice that

$$\mathbf{P}_{\beta,0}(\Lambda_n = (nq, np)) = \mathbf{P}_{n,\mathbf{h}_n^{q,p}}(\Lambda_n = (nq, np)) e^{-n\mathbf{h}_n^{q,p} \cdot (q,p) + \mathcal{L}_{\Lambda_n}(\mathbf{h}_n^{q,p})}. \quad (1.6.16)$$

Moreover [52, Prop. 6.1] gives a uniform local central limit theorem for  $\mathbf{P}_{n,\mathbf{h}_n^{q,p}}$ : for any  $q_1 < q_2, p_1 < p_2, t_1 < t_2$  and  $s_1 < s_2$ , there are constants  $C'_5, C_5 > 0$  and  $n_0 \in \mathbb{N}$  such that for any  $q \in [q_1, q_2], p \in [p_1, p_2], t \in [t_1, t_2], s \in [s_1, s_2]$  and  $n \geq n_0$ ,

$$\frac{C_5}{n^2} \leq \mathbf{P}_{n,\mathbf{h}_n^{q,p}}(\Lambda_n = (qn + t\sqrt{n}, pn + s\sqrt{n})) \leq \frac{C'_5}{n^2}. \quad (1.6.17)$$

Notice that the authors of [52] only state this for  $q \in [q_1, q_2]$  and  $p = 0$ , but their proof can be extended to  $p \in [p_1, p_2]$  without trouble since the local limit theorem and all uniformity arguments for  $q$  hold similarly for  $p \neq 0$ . Details are left to the reader.

The asymptotic of  $\mathbf{h}_n^{q,p}$  for large  $n$  can be described sharply. For any  $\mathbf{h} \in \mathcal{D}_\beta$ , Recall that we defined  $\mathcal{L}_\Lambda(\mathbf{h}) := \int_0^1 \mathcal{L}(h_0 x + h_1) dx$  for all  $\mathbf{h} \in \mathcal{D}_\beta$  (see (1.3.9)), and let  $\tilde{\mathbf{h}}^{q,p} = (\tilde{h}_0^{q,p}, \tilde{h}_1^{q,p})$  be the unique solution in  $\mathbf{h} \in \mathcal{D}_\beta$  of the equation  $\nabla \mathcal{L}_\Lambda(\mathbf{h}) = (q, p)$ , where

$$\begin{aligned} \nabla \mathcal{L}_\Lambda(\mathbf{h}) &= (\partial_{h_0} \mathcal{L}_\Lambda, \partial_{h_1} \mathcal{L}_\Lambda)(\mathbf{h}) \\ &= \left( \int_0^1 x \mathcal{L}'(xh_0 + h_1) dx, \int_0^1 \mathcal{L}'(xh_0 + h_1) dx \right), \end{aligned} \quad (1.6.18)$$

is a  $\mathcal{C}^1$  diffeomorphism from  $\mathcal{D}_\beta$  to  $\mathbb{R}^2$  (see [52, Lem. 5.3]). If  $h_0 \neq 0$ , an integration by parts yields

$$\nabla \mathcal{L}_\Lambda(\mathbf{h}) = \frac{1}{h_0} \left( \mathcal{L}(h_0 + h_1) - \mathcal{L}_\Lambda(\mathbf{h}), \mathcal{L}(h_0 + h_1) - \mathcal{L}(h_1) \right), \quad (1.6.19)$$

and if  $h_0 = 0$ ,  $\nabla \mathcal{L}_\Lambda(\mathbf{h}) = (\mathcal{L}'(h_1)/2, \mathcal{L}'(h_1))$ .

For any  $q_1 < q_2$  and  $p_1 < p_2 \in \mathbb{R}$ , [52, Prop. 2.3] gives constants  $C_6, C'_6 > 0$  and  $n_0$ , such that for any  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  and  $n \geq n_0$ , one has <sup>†</sup>

$$\left| \left[ \frac{1}{n} \mathcal{L}_{\Lambda_n}(\mathbf{h}_n^{q,p}) - \mathbf{h}_n^{q,p} \cdot (q, p) \right] - \left[ \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q,p}) - \tilde{\mathbf{h}}^{q,p} \cdot (q, p) \right] \right| \leq \frac{C_6}{n}, \quad (1.6.20)$$

and

$$\|\mathbf{h}_n^{(q,p)} - \tilde{\mathbf{h}}^{q,p}\| \leq \frac{C'_6}{n}. \quad (1.6.21)$$

Notice that [52, Prop. 2.3] is only formulated for  $p = 0$  and  $q \in [q_1, q_2]$  with  $q_1 > 0$ , but all uniformity arguments also hold for  $p \in [p_1, p_2]$  and any  $[q_1, q_2] \subset \mathbb{R}$ . These estimates, combined with (1.6.16) and (1.6.17), give the precise asymptotic of  $\mathbf{P}_\beta(\Lambda_n = (nq, np))$  (up to constant factors) uniformly in  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$ .

Before we start the proof, let us point out that  $\tilde{h}_0^{q,p} > 0$  for all  $q, p \in \mathbb{R}$  with  $p \leq 2q - \varepsilon$ . Indeed, one notices that  $\tilde{h}_0^{q,p} = 0$  and (1.6.18) imply  $p = 2q$ . Because  $\tilde{\mathbf{h}}$  is continuous, it has constant sign on each sets  $\{p < 2q\}$  and  $\{p > 2q\}$ , and it is already stated in [52, Rem. 5.5] that  $\tilde{h}_0^{q,0} > 0$  for all  $q > 0$ .

Finally we also define the i.i.d. uniform tilt of the increments of  $X$  for any  $|\delta| < \beta/2$ :

$$\frac{d\tilde{\mathbf{P}}_\delta}{d\mathbf{P}_{\beta,0}}(X_1) = e^{\delta X_1 - \mathcal{L}(\delta)}, \quad (1.6.22)$$

where we defined  $\mathcal{L}$  in (1.3.8).

*Proof of Proposition 1.4.3.* Here is the strategy of the proof: recollecting (1.6.16), (1.6.17) and (1.6.20), we already have

$$\mathbf{P}_{\beta,0}(\Lambda_N = (qN, pN)) \asymp \frac{1}{N^2} e^{-N \left( \tilde{\mathbf{h}}^{q,p} \cdot (q,p) - \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q,p}) \right)}, \quad (1.6.23)$$

(recall (5.0.1) for the definition of  $\asymp$ ). Hence the proof will be complete when we show that

$$\mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < N \mid \Lambda_N = (qN, pN))$$

is bounded from below by some positive constant uniform in  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  and  $N \in \mathbb{N}$ .

For that purpose we decompose a trajectory  $X$  with  $X_0 = 0$ ,  $X_N = pN$  and area  $A_N(X) = qN^2$  into three parts: both ends of length  $u_N := N^{1/3}$ , and the middle part of length  $N' := N - 2u_N$ . To estimate the probability to go under the fixed curve  $\tilde{f}(k)$ , we study the middle part under the (non-uniform) tilt  $\mathbf{P}_{N',\mathbf{h}}$ , and we handle both ends with uniform tilts  $\tilde{\mathbf{P}}_{\delta_1}, \tilde{\mathbf{P}}_{\delta_2}$ . Then we take advantage of the fact that uniformly tilted walks remain

<sup>†</sup>. We actually prove an improvement of this result in Chapter 2, Proposition 2.5.1 in the case  $p = 0$ . In particular it allows us to compute the constant  $K_\beta$  in Theorem 2.2.1; however, it is not sufficient to improve Theorem 1.3.2.

positive (and even above certain affine curves) with positive probability, and we handle the middle part with aforementioned estimates.

For  $i \in \{1, 2\}$ , define

$$Q_{i,u_N} := \left( \left[ a_i u_N / 2, b_i u_N / 2 + 3K + K_S \right] \cap \left( \frac{1}{u_N} \mathbb{Z} \right) \right) \times \left( \left[ a_i (u_N - 1) + K + K_S, b_i (u_N - 1) + 3K + K_S \right] \cap \mathbb{Z} \right), \quad (1.6.24)$$

where  $a_i < b_i, i \in \{1, 2\}$  and  $K$  are convenient positive constants which will be fixed below. We constrain the first bit of the walk (until index  $u_N$ ) to grow by  $X_{u_N} = x_1$  and have an area  $A_{u_N} = A_1$ , with  $(A_1/u_N, x_1) \in Q_{1,u_N}$ . Similarly, we reverse the third bit in time (from  $N$  to  $N - u_N$ ), and we constrain it to grow by  $X_{N-u_N} - pN = x_2$  and have an area  $(A_N - A_{N-u_N}) - pNu_N = A_2$ , with  $(A_2/u_N, x_2) \in Q_{2,u_N}$ .

Therefore we have the decomposition

$$\begin{aligned} & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < N; \Lambda_N = (qN, pN)) \\ & \geq \sum_{\substack{(A_1/u_N, x_1) \in Q_{1,u_N}, \\ (A_2/u_N, x_2) \in Q_{2,u_N}}} \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} = (A_1/u_N, x_1)) \\ & \quad \times \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k + u_N) - x_1, \forall 0 < k < N'; \Lambda_{N'} = (q'N', p'N')) \\ & \quad \times \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} = (A_2/u_N, x_2)), \end{aligned} \quad (1.6.25)$$

where  $N' := N - 2u_N$ , and we define  $(q', p')$  such that  $q'N'^2 = qN^2 - A_1 - A_2 - x_1N' - pNu_N$  and  $p'N' = pN + x_2 - x_1$ . Notice that with our choice of  $A_1, A_2 = O(u_N^2)$ ,  $x_1, x_2 = O(u_N)$ , with  $u_N = N^{1/3}$ , and because  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$ , then for  $N$  sufficiently large  $q'$  and  $p'$  are contained in some compact sets  $[q'_1, q'_2]$  and  $[p'_1, p'_2]$ . Moreover we have the following estimates:

$$\begin{aligned} q - q' &= (p' - 4q') \frac{u_N}{N} + \frac{x_1}{N} + o(1/N), \\ p - p' &= -2p' \frac{u_N}{N} + \frac{x_1 - x_2}{N}. \end{aligned} \quad (1.6.26)$$

As mentioned above, we tilt uniformly the first and last factors in the r.h.s. of (1.6.25) with respective parameters  $\delta_1, \delta_2$  (which will be explicitly fixed later),

$$\begin{aligned} & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} = (A_1/u_N, x_1)) \\ & = e^{-\delta_1 x_1 + u_N \mathcal{L}(\delta_1)} \tilde{\mathbf{P}}_{\delta_1}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} = (A_1/u_N, x_1)) ; \\ & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} = (A_2/u_N, x_2)) \\ & = e^{-\delta_2 x_2 + u_N \mathcal{L}(\delta_2)} \tilde{\mathbf{P}}_{\delta_2}(X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} = (A_2/u_N, x_2)). \end{aligned} \quad (1.6.27)$$

The second factor in the r.h.s. of (1.6.25) is bounded from below by the following lemma.

**Lemma 1.6.1.** *Fix some  $\varepsilon > 0$  and  $q'_1 < q'_2$ ,  $p'_1 < p'_2$  in  $\mathbb{R}$ . Then there are constants  $c, C > 0$  and  $N_0 \in \mathbb{N}$  such that for any  $N' \geq N_0$ ,  $q' \in [q'_1, q'_2]$  and  $p' \in [p'_1, p'_2]$  satisfying  $p' \leq 2q' - \varepsilon$ , one has*

$$\begin{aligned} & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k + u_N) - x_1, \forall 0 < k < N'; \Lambda_{N'} = (q'N', p'N')) \\ & \geq e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q', p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} \left( \mathbf{P}_{N', \mathbf{h}_{N'}^{q',p'}}(\Lambda_{N'} = (q'N', p'N')) - C e^{-c u_N} \right). \end{aligned} \quad (1.6.28)$$

This lemma will be proven afterwards. Writing (1.6.25) with (1.6.27) and (1.6.28), we have

$$\begin{aligned} & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < N; \Lambda_N = (qN, pN)) \\ & \geq \sum_{\substack{(A_1/u_N, x_1) \in Q_{1,u_N}, \\ (A_2/u_N, x_2) \in Q_{2,u_N}}} e^{-\delta_1 x_1 + u_N \mathcal{L}(\delta_1)} e^{-\delta_2 x_2 + u_N \mathcal{L}(\delta_2)} e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q', p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} \\ & \quad \times \left( \mathbf{P}_{N', \mathbf{h}_{N'}^{q',p'}}(\Lambda_{N'} = (q'N', p'N')) - C e^{-c u_N} \right) \\ & \quad \times \tilde{\mathbf{P}}_{\delta_1}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} = (A_1/u_N, x_1)) \\ & \quad \times \tilde{\mathbf{P}}_{\delta_2}(X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} = (A_2/u_N, x_2)). \end{aligned} \quad (1.6.29)$$

Then we divide by  $\mathbf{P}_{\beta,0}(\Lambda_N = (qN, pN))$ , which we express with the tilt  $\mathbf{P}_{N, \mathbf{h}_N^{q',p'}}$  (rather than  $\mathbf{P}_{N, \mathbf{h}_N^{q,p}}$ ), so we have

$$\begin{aligned} & \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < N \mid \Lambda_N = (qN, pN)) \\ & \geq \sum_{\substack{(A_1/u_N, x_1) \in Q_{1,u_N}, \\ (A_2/u_N, x_2) \in Q_{2,u_N}}} e^{g(x_1, a_1, x_2, a_2)} \left( \frac{\mathbf{P}_{N', \mathbf{h}_{N'}^{q',p'}}(\Lambda_{N'} = (q'N', p'N')) - C e^{-c u_N}}{\mathbf{P}_{N, \mathbf{h}_N^{q',p'}}(\Lambda_N = (qN, pN))} \right) \\ & \quad \times \tilde{\mathbf{P}}_{\delta_1}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} = (A_1/u_N, x_1)) \\ & \quad \times \tilde{\mathbf{P}}_{\delta_2}(X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} = (A_2/u_N, x_2)), \end{aligned} \quad (1.6.30)$$

where we define

$$\begin{aligned} g(x_1, a_1, x_2, a_2) & := -\delta_1 x_1 + u_N \mathcal{L}(\delta_1) - \delta_2 x_2 + u_N \mathcal{L}(\delta_2) - N' \mathbf{h}_{N'}^{q',p'} \cdot (q', p') \\ & \quad + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'}) + N \mathbf{h}_N^{q',p'} \cdot (q, p) - \mathcal{L}_{\Lambda_N}(\mathbf{h}_N^{q',p'}). \end{aligned} \quad (1.6.31)$$

Let us handle the factor in parenthesis in (1.6.30) first. We can apply (1.6.17) to both probabilities (notice that  $q - q'$ ,  $p - p'$  are of order  $u_N/N \ll N^{-1/2}$ , recall (1.6.26)) to bound it from below by  $(C_7 \frac{N^2}{N^2} - C_8 N^2 e^{-c u_N}) \geq C_9 > 0$  for  $N$  sufficiently large, uniformly in  $q, p$ .

Now let us focus on  $g(x_1, a_1, x_2, a_2)$ , where we fix  $\delta_1 = \tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}$  and  $\delta_2 = -\tilde{h}_1^{q',p'}$ . Notice that they match respectively the value of the exponential tilt applied on the first and last increments in the second piece of trajectory and in the limit  $N \rightarrow \infty$  (recall (1.6.14),

and the sign of  $\delta_2$  is inverted because we reversed in time the third bit of the trajectory). We introduce in (1.6.31) a term  $N\mathbf{h}_N^{q',p'} \cdot (q', p')$  and we apply (1.6.20) for  $(q', p', N')$  and  $(q', p', N)$  to bound from below

$$\begin{aligned} g(x_1, a_1, x_2, a_2) &\geq -\delta_1 x_1 + u_N \mathcal{L}(\delta_1) - \delta_2 x_2 + u_N \mathcal{L}(\delta_2) + N\mathbf{h}_N^{q',p'} \cdot (q - q', p - p') \\ &\quad - N' \left( \tilde{\mathbf{h}}^{q',p'} \cdot (q', p') - \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q',p'}) \right) + N \left( \tilde{\mathbf{h}}^{q',p'} \cdot (q', p') - \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q',p'}) \right) - C_{10}, \end{aligned} \quad (1.6.32)$$

for some uniform constant  $C_{10} < \infty$ . Recall  $N' = N - 2u_N$ , and apply (1.6.21) to the term  $N\mathbf{h}_N^{q',p'} \cdot (q - q', p - p')$  to obtain

$$\begin{aligned} g(x_1, a_1, x_2, a_2) &\geq -\delta_1 x_1 + u_N \mathcal{L}(\delta_1) - \delta_2 x_2 + u_N \mathcal{L}(\delta_2) + N\tilde{\mathbf{h}}^{q',p'} \cdot (q - q', p - p') \\ &\quad + 2u_N \left( \tilde{\mathbf{h}}^{q',p'} \cdot (q', p') - \mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q',p'}) \right) - C_{11}. \end{aligned} \quad (1.6.33)$$

Recalling that  $\delta_1 = \tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}$  and  $\delta_2 = -\tilde{h}_1^{q',p'}$ , and also (1.6.26), we have

$$\begin{aligned} g(x_1, a_1, x_2, a_2) &\geq -(\tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'})x_1 + h_1^{q',p'} x_2 \\ &\quad + u_N \left( 2\tilde{\mathbf{h}}^{q',p'} \cdot (q', p') - 2\mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q',p'}) + \mathcal{L}(\tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}) + \mathcal{L}(-\tilde{h}_1^{q',p'}) \right) \\ &\quad + \tilde{h}_0^{q',p'} \left( (p' - 4q')u_N + x_1 \right) + \tilde{h}_1^{q',p'} \left( -2p'u_N + x_1 - x_2 \right) - C_{12}. \end{aligned} \quad (1.6.34)$$

So,

$$\begin{aligned} g(x_1, a_1, x_2, a_2) &\geq u_N \left( -2\tilde{h}_0^{q',p'} q' + \tilde{h}_0^{q',p'} p' - 2\mathcal{L}_\Lambda(\tilde{\mathbf{h}}^{q',p'}) \right) \\ &\quad + \mathcal{L}(\tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}) + \mathcal{L}(-\tilde{h}_1^{q',p'}) - C_{12}, \end{aligned} \quad (1.6.35)$$

and we conclude by recalling the definition of  $\tilde{\mathbf{h}}^{q',p'}$  and (1.6.19): the first term is zero, and  $g(x_1, a_1, x_2, a_2)$  is bounded from below uniformly by some constant.

Therefore, we can bound (1.6.30) from below by

$$\begin{aligned} \mathbf{P}_{\beta,0}(X_k \geq \tilde{f}(k), \forall 0 < k < N \mid \Lambda_N = (qN, pN)) \\ &\geq C_{13} \times \tilde{\mathbf{P}}_{\delta_1} \left( X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} \in \mathcal{Q}_{1,u_N} \right) \\ &\quad \times \tilde{\mathbf{P}}_{\delta_2} \left( X_k \geq \tilde{f}(N - k) - pN, \forall 0 < k < u_N; \Lambda_{u_N} \in \mathcal{Q}_{2,u_N} \right), \end{aligned} \quad (1.6.36)$$

with  $\delta_1 = \tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}$ ,  $\delta_2 = -\tilde{h}_1^{q',p'}$ , and  $C_{13}$  is some uniform positive constant. The proof will be over once we show that those factors are uniformly bounded away from 0. We will focus on the first factor, because we notice in the second one  $\tilde{f}(N - k) - pN = Nf(\frac{k}{N}) - kp + K_S$ , which matches  $\tilde{f}(k)$  with an inverted slope  $-p$  (recall we reversed in time that bit); hence it is handled by the exact same arguments.

Recall that  $\nabla \mathcal{L}_\Lambda$  is a  $\mathcal{C}^1$  diffeomorphism, in particular  $\tilde{\mathbf{h}}$  is Lipschitz on compact sets, and  $(q - q')$ ,  $(p - p')$  are of order  $u_N/N = N^{-2/3}$  (recall (1.6.26)), so there exists some

constant  $C_{14} > 0$  such that for all  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  and  $N \in \mathbb{N}$ ,

$$\left| \delta_1 - \tilde{h}_0^{p,q} - \tilde{h}_1^{p,q} \right| \leq \frac{C_{14} u_N}{N}. \quad (1.6.37)$$

In particular for  $N$  large enough,  $\delta_1$  is contained in some compact set  $[\underline{\delta}, \bar{\delta}] \subset (-\beta/2, \beta/2)$  uniformly in  $q, p$ . As claimed earlier we have  $\tilde{h}_0^{q,p} > 0$  under our assumptions—it is even bounded away from 0—so we also have  $\tilde{h}_0^{q',p'} > 0$  for  $N$  large (recall (1.6.26) again and  $\tilde{h}$  is locally Lipschitz). Combined with (1.6.19) and with the strict convexity of  $\mathcal{L}$ , this implies

$$\mathcal{L}'(\tilde{h}_1^{q',p'}) < p' < \mathcal{L}'(\tilde{h}_0^{q',p'} + \tilde{h}_1^{q',p'}). \quad (1.6.38)$$

In particular there is some constant  $C_{15} > 0$  such that  $\mathcal{L}'(\delta_1) - p' \geq C_{15}$  and  $p' + \mathcal{L}'(\delta_2) \geq C_{15}$  uniformly in  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  and  $N \in \mathbb{N}$  (recall that those functions are continuous, and  $\mathcal{L}'(\tilde{h}_1^{q',p'}) = -\mathcal{L}'(-\tilde{h}_1^{q',p'})$ ).

Fix some  $a, C > 0$ . For any  $\delta$  such that  $|\delta| < \beta/2$ , and any  $0 < t < \beta/2 + \delta$ ,  $k \in \mathbb{N}$ , Markov's inequality implies that

$$\tilde{\mathbf{P}}_\delta(X_k \leq ak - C) \leq \tilde{\mathbf{E}}_\delta[e^{t(ak-C) - tX_k}] = e^{-Ct} e^{k(\tilde{\mathcal{L}}_\delta(-t) + at)}, \quad (1.6.39)$$

where  $\tilde{\mathcal{L}}_\delta(-t) := \log \tilde{\mathbf{E}}_\delta[e^{-tX_1}] = \mathcal{L}(\delta - t) - \mathcal{L}(\delta)$  for any  $t \in (-\beta/2 + \delta, \beta/2 + \delta)$ . Choosing  $a = \frac{3}{4}\mathcal{L}'(\delta_1) + \frac{1}{4}p'$  and writing a Taylor expansion of  $\mathcal{L}$ , there is some (uniform)  $c > 0$  such that for  $t$  small,

$$\tilde{\mathcal{L}}_{\delta_1}(-t) + at \leq \frac{t}{4}(p' - \mathcal{L}'(\delta_1)) + ct^2. \quad (1.6.40)$$

Recall that  $p' - \mathcal{L}'(\delta_1) < 0$ , and it is even bounded away from 0 (uniformly). So for  $t$  sufficiently small, we have  $\tilde{\mathcal{L}}_{\delta_1}(-t) + at \leq \frac{t}{8}(p' - \mathcal{L}'(\delta_1))$ . Hence there exists a  $t_0 > 0$  such that,

$$\begin{aligned} \tilde{\mathbf{P}}_{\delta_1}(\exists k \geq 1, X_k \leq ak - C) &\leq \sum_{k \geq 1} \tilde{\mathbf{P}}_{\delta_1}(X_k \leq ak - C) \\ &\leq e^{-Ct_0} \frac{e^{t_0(p' - \mathcal{L}'(\delta_1))/8}}{1 - e^{t_0(p' - \mathcal{L}'(\delta_1))/8}} \leq C_{16} e^{-Ct_0}, \end{aligned} \quad (1.6.41)$$

where we bounded the fraction by some constant uniform in  $(q', p')$ .

Similarly, we have

$$\tilde{\mathbf{P}}_\delta(X_k \geq bk + C) \leq \tilde{\mathbf{E}}_\delta[e^{wX_k - w(bk+C)}] = e^{-Cw} e^{k(\tilde{\mathcal{L}}_\delta(w) - bw)}, \quad (1.6.42)$$

for some  $b, C > 0$  fixed, and any  $k \in \mathbb{N}$ ,  $\delta$  such that  $|\delta| < \beta/2$ , and  $0 < w < \beta/2 - \delta$ . Because  $\tilde{\mathcal{L}}_\delta(w) \sim \mathcal{L}'(\delta)w$  as  $w \rightarrow 0$ , we fix some  $b > \sup\{\mathcal{L}'(\delta), \delta \in [\underline{\delta}, \bar{\delta}]\}$ , and for some  $w_0 > 0$  sufficiently small we have

$$\tilde{\mathbf{P}}_{\delta_1}(\exists k \geq 1, X_k \geq bk + C) \leq e^{-Cw_0} \frac{e^{\tilde{\mathcal{L}}_{\delta_1}(w_0) - bw_0}}{1 - e^{\tilde{\mathcal{L}}_{\delta_1}(w_0) - bw_0}} \leq C_{17} e^{-Cw_0}, \quad (1.6.43)$$

where we bounded the fraction by some uniform constant (notice that we could pick  $w_0 > 0$  such that  $\tilde{\mathcal{L}}_\delta(w_0) - bw_0 < 0$  for all  $\delta \in [\underline{\delta}, \bar{\delta}]$ ).

Finally, we estimate the first factor in (1.6.36) by constraining the trajectory  $X$  to have a first step  $X_1 = 2K + K_S$ , then to remain between the two curves  $a(k-1) + K + K_S$  and  $b(k-1) + 3K + K_S$ ,  $k \geq 2$ . Moreover for  $\gamma > 0$  sufficiently small, we have

$$a - (p + \gamma) = \frac{3}{4}(\mathcal{L}'(\delta_1) - p') - \gamma + (p' - p) > 0, \quad (1.6.44)$$

(recall that  $p' - p = O(N^{-2/3})$  and  $\mathcal{L}'(\delta_1) - p'$  is uniformly bounded away from 0), and this  $\gamma$  can be chosen independently of  $K_S$ . This means that our constraint implies that  $X_k$  remains above  $\tilde{f}(k) = \gamma k + kp + K_S$  for all  $1 \leq k \leq u_N$ . It also implies  $\Lambda_{u_N} \in Q_{1,u_N}$  by setting  $a_1 = a = \frac{3}{4}\mathcal{L}'(\delta_1) + \frac{1}{4}p'$  and  $b_1 = b$  in the definition of  $Q_{1,u_N}$ . Recollecting (1.6.41) and (1.6.43), we finally obtain

$$\begin{aligned} & \tilde{\mathbf{P}}_{\delta_1}(X_k \geq \tilde{f}(k), \forall 0 < k < u_N; \Lambda_{u_N} \in Q_{1,u_N}) \\ & \geq \tilde{\mathbf{P}}_{\delta_1}(X_1 = 2K + K_S; a(k-1) + K + K_S \leq X_k \leq b(k-1) + 3K + K_S, \forall k \geq 2) \\ & = \tilde{\mathbf{P}}_{\delta_1}(X_1 = 2K + K_S) \times \tilde{\mathbf{P}}_{\delta_1}(ak - K \leq X_k \leq bk + K, \forall k \geq 1) \\ & \geq \tilde{\mathbf{P}}_{\delta_1}(X_1 = 2K + K_S) \times (1 - C_{18}e^{-K \min(t_0, w_0)}) \geq C_{19} > 0. \end{aligned} \quad (1.6.45)$$

where we choose  $K > 0$  such that the second factor in (1.6.45) is greater than  $\frac{1}{2}$ , then bound  $\tilde{\mathbf{P}}_{\delta_1}(X_1 = 2K + K_S)$  uniformly over  $\delta_1 \in [\underline{\delta}, \bar{\delta}]$ . We can reproduce the same proof to handle the other factor in (1.6.36), by setting  $a_2 = \frac{3}{4}\mathcal{L}'(\delta_2) - \frac{1}{4}p'$  and a convenient  $b_2 > a_2$  in the definition of  $Q_{2,u_N}$ . Therefore (1.6.36) is bounded from below by some uniform positive constant, and this concludes the proof of the proposition (subject to Lemma 1.6.1).  $\square$

Let us now prove Lemma 1.6.1. This result relies on estimates that are similar to those displayed at the end of the proof of Proposition 1.4.3, where we used Markov inequalities for  $\tilde{\mathbf{P}}_\delta$ . First we prove an upper bound on the moment generating function of the difference between the random walk with law  $\mathbf{P}_{N', h_N^{q,p}}$  and the straight line with slope  $p$ .

**Lemma 1.6.2.** *There exists  $\lambda_0 > 0$  such that for any  $\lambda \in (0, \lambda_0)$ , there are  $c > 0$  and  $N_0 \in \mathbb{N}$  such that*

$$\mathbf{E}_{N, h_N^{q,p}} \left[ e^{-\lambda(X_k - pk)} \right] \leq e^{-ck}, \quad (1.6.46)$$

uniformly in  $0 \leq k \leq N$ ,  $N \geq N_0$ ,  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  satisfying  $p \leq 2q - \varepsilon$ .

*Proof.* Under the law  $\mathbf{P}_{N, h_N^{q,p}}$ , the increments of the random walk  $X$  are still independent (but no more identically distributed), so we define  $h_N^i := (1 - i/N)h_{N,0}^{q,p} + h_{N,1}^{q,p}$ , the tilt

parameter for the increment  $X_i - X_{i-1}$ ,  $i \geq 1$ . Then we have

$$\begin{aligned}
 \log \mathbf{E}_{N, \mathbf{h}_N^{q,p}} \left[ e^{-\lambda(X_k - pk)} \right] &= \lambda pk + \sum_{i=1}^k \left( \mathcal{L}(-\lambda + h_N^i) - \mathcal{L}(h_N^i) \right) \\
 &= -\lambda \sum_{i=1}^k \left( \frac{\mathcal{L}(h_N^i) - \mathcal{L}(-\lambda + h_N^i)}{\lambda} - p \right) \\
 &\leq -\lambda \sum_{i=1}^k \left( \mathcal{L}'(-\lambda + h_N^i) - p \right),
 \end{aligned} \tag{1.6.47}$$

where we used the convexity of  $\mathcal{L}$ , and the r.h.s. is well-defined for all  $|\lambda| < \lambda_0$ , for some  $\lambda_0 > 0$  uniform in  $(q, p)$  (recall that  $\mathbf{h}_N^{q,p}$  is contained in a compact subset of  $\mathcal{D}_{\beta,n}$ ). Notice that  $h_N^i$  is affine, decreasing in  $i$  (recall that  $p \leq 2q - \varepsilon$  implies  $\tilde{h}_0^{q,p} > 0$ ), and recall that  $\mathcal{L}'$  is increasing. Therefore it suffices to prove the claim for  $k = N$ , that is to prove

$$\frac{1}{N} \sum_{i=1}^N \mathcal{L}'(-\lambda + h_N^i) - p \geq c > 0, \tag{1.6.48}$$

and we conclude by noticing that if  $(u_n)_{n \geq 1}$  is a decreasing sequence, then  $(\frac{1}{n} \sum_{i=1}^n u_i)_{n \geq 1}$  is decreasing too, hence the claim holds for  $k \leq N$ .

Notice that the first term in the l.h.s. of (1.6.48) is a Riemann sum, and we have

$$\frac{1}{N} \sum_{i=1}^N \mathcal{L}'(-\lambda + h_N^i) \xrightarrow{N \rightarrow \infty} \int_0^1 \mathcal{L}'(x \tilde{h}_0^{q,p} + \tilde{h}_1^{q,p} - \lambda) dx. \tag{1.6.49}$$

We also claim that the convergence holds uniformly in  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$  —this follows from a Riemann-sum approximation of the r.h.s., from (1.6.21) and from the fact that  $\mathcal{L}'$  is locally Lipschitz. Moreover we have  $p = \int_0^1 \mathcal{L}'(x \tilde{h}_0^{q,p} + \tilde{h}_1^{q,p}) dx$  (recall (1.6.18)), so this concludes the proof of (1.6.48) for  $N$  sufficiently large and some  $c > 0$  uniform in  $q \in [q_1, q_2]$ ,  $p \in [p_1, p_2]$ .  $\square$

*Proof of Lemma 1.6.1.* Let us bound from above

$$\mathbf{P}_{\beta,0} \left( X_k < \tilde{f}(k + u_N) - x_1; \Lambda_{N'} = (q'N', p'N') \right),$$

for any  $0 < k < N'$ , where we recall  $x_1 \in [a_1(u_N - 1) + K + K_S, b_1(u_N - 1) + 3K + K_S]$ ,  $u_N = N^{1/3}$ ,  $N' = N - 2u_N$  and  $a_1$  satisfies (1.6.44) with  $a = a_1$ . We have

$$\begin{aligned}
 &\mathbf{P}_{\beta,0} \left( X_k < \tilde{f}(k + u_N) - x_1; \Lambda_{N'} = (q'N', p'N') \right) \\
 &\leq \mathbf{P}_{\beta,0} \left( X_k < -\zeta u_N + pk + \gamma(k \wedge (N' - k)) - K + a_1; \Lambda_{N'} = (q'N', p'N') \right),
 \end{aligned} \tag{1.6.50}$$



where we define  $\zeta := a_1 - p - \gamma$  (which is positive and bounded away from 0 uniformly in  $p$ ). Recalling (1.6.13) and applying Markov's inequality, we have

$$\begin{aligned}
 & \mathbf{P}_{\beta,0} \left( X_k < \tilde{f}(k + u_N) - x_1; \Lambda_{N'} = (q'N', p'N') \right) \\
 & \leq e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q',p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} \\
 & \quad \times \mathbf{P}_{N', \mathbf{h}_{N'}^{q',p'}} \left( X_k < -\zeta u_N + pk + \gamma(k \wedge (N' - k)) - K + a_1; \Lambda_{N'} = (q'N', p'N') \right) \\
 & \leq e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q',p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} \\
 & \quad \times e^{-\lambda \zeta u_N + \lambda \gamma (k \wedge (N' - k)) + \lambda (a_1 - K)} \mathbf{E}_{N', \mathbf{h}_{N'}^{q',p'}} \left[ e^{-\lambda (X_k - p'k)} \right] e^{\lambda (p - p')k}, \tag{1.6.51}
 \end{aligned}$$

for some  $\lambda \in (0, \lambda_0)$ . Applying Lemma 1.6.2 and the inequality  $k \wedge (N' - k) \leq k$ , we obtain

$$\begin{aligned}
 & \mathbf{P}_{\beta,0} \left( X_k < \tilde{f}(k + u_N) - x_1; \Lambda_{N'} = (q'N', p'N') \right) \\
 & \leq e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q',p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} e^{-\lambda \zeta u_N + \lambda \gamma k + \lambda (a_1 - K)} e^{-ck} e^{\lambda (p - p')k}. \tag{1.6.52}
 \end{aligned}$$

Choosing  $\gamma < c/\lambda$  (independently of  $K_S$ ), and recalling that  $p - p'$  goes uniformly to 0 (see (1.6.26)), there is  $N_0 \in \mathbb{N}$  such that  $\lambda \gamma - c + \lambda(p - p') \leq -C_{20} < 0$  uniformly in  $(q, p)$  and  $N \geq N_0$ . Therefore we have

$$\begin{aligned}
 & \mathbf{P}_{\beta,0} \left( X_k \geq \tilde{f}(k + u_N) - x_1, \forall 0 < k < N'; \Lambda_{N'} = (q'N', p'N') \right) \\
 & \geq \mathbf{P}_{\beta,0} \left( \Lambda_{N'} = (q'N', p'N') \right) - \sum_{k=1}^{N'} \mathbf{P}_{\beta,0} \left( X_k < \tilde{f}(k + u_N) - x_1; \Lambda_{N'} = (q'N', p'N') \right) \\
 & \geq e^{-N' \mathbf{h}_{N'}^{q',p'} \cdot (q',p') + \mathcal{L}_{\Lambda_{N'}}(\mathbf{h}_{N'}^{q',p'})} \\
 & \quad \times \left( \mathbf{P}_{N', \mathbf{h}_{N'}^{q',p'}} \left( \Lambda_{N'} = (q'N', p'N') \right) - e^{-\lambda \zeta u_N} e^{\lambda (a_1 - K)} \sum_{k=1}^{N'} e^{-C_{20}k} \right), \tag{1.6.53}
 \end{aligned}$$

and  $\sum_{k \geq 1} e^{-C_{20}k} < \infty$ , which concludes the proof.  $\square$

# Appendices to Chapter 1

## 1.A The wetting model

In this appendix we provide well-known results and estimates on the wetting model which are proven in [84]. Recall that we defined  $\mathbf{P}_\beta$  in (1.2.8), and for all  $N \geq 1$ , let

$$Z_{\text{wet},N}^{\beta,\delta} := \mathbf{E}_{\beta,0} \left[ e^{\delta \sum_{k=1}^N \mathbf{1}_{\{X_k=0\}}} \mathbf{1}_{\{X \in B_N^{0,+}\}} \right], \quad (1.A.1)$$

be the (constrained) partition function of a wetting model associated with a random walk of law  $\mathbf{P}_\beta$ . Let  $h_\beta(\delta)$  be its free energy —recall (1.2.9), and notice that it is well-defined because  $(Z_{\text{wet},N}^{\beta,\delta})_{N \geq 1}$  is super-multiplicative. Note that  $h_\beta(\delta)$  is non-decreasing in  $\delta$ , that  $h_\beta(\delta) \leq \delta$ , and recalling (1.6.1), we have  $h_\beta(0) = 0$ ; hence  $h_\beta(\delta) \geq 0$  for all  $\delta \geq 0$ , and we define  $\tilde{\delta}_c(\beta) := \inf\{\delta \geq 0, h_\beta(\delta) > 0\} \in [0, \infty]$ .

Let  $\tau := \inf\{t \geq 1; X_t \leq 0\}$  and let  $K_\beta(t) := \mathbf{P}_{\beta,0}(\tau = t, X_t = 0)$ ,  $t \geq 1$ . Notice that conditionally to  $\tau$ ,  $-X_\tau$  follows a geometric distribution on  $\mathbb{N}_0$  (in particular  $\tau$  and  $X_\tau$  are independent). Therefore one has  $\sum_{t \geq 1} K_\beta(t) = \mathbf{P}_{\beta,0}(X_\tau = 0) = 1 - e^{-\beta/2}$ , and  $K_\beta(t) = \mathbf{P}_{\beta,0}(\tau = t)(1 - e^{-\beta/2})$ . It is a straightforward application of [47, (4.5)] that there exists  $c > 0$  depending on  $\beta$  only such that

$$K_\beta(t) \sim \frac{c}{t^{3/2}} \quad \text{as } t \rightarrow \infty. \quad (1.A.2)$$

**Proposition 1.A.1** ([84, Thm. 2.2 & Eq. 2.2]).

(i) For  $\beta > 0$ , the free energy of this wetting model  $h_\beta(\delta)$  is the only solution in  $\zeta \geq 0$  of the equation

$$\sum_{t \geq 1} K_\beta(t) e^{-\zeta t} = e^{-\delta}, \quad (1.A.3)$$

if it exists (that is if  $\delta > \tilde{\delta}_c(\beta) = -\log(1 - e^{-\beta/2})$ ), and 0 otherwise. An explicit expression of  $h_\beta(\delta)$  is given by

$$h_\beta(\delta) = \log \left( \frac{(e^\delta - 1)(1 - e^{-\beta/2})^2}{1 - e^{-\delta} - e^{-\beta}} \right), \quad \text{for } \delta > -\log(1 - e^{-\beta/2}), \quad (1.A.4)$$

and 0 otherwise.

(ii) The asymptotics of the partition function are given by:

(a) if  $\delta > \tilde{\delta}_c(\beta)$ , then as  $N \rightarrow \infty$ ,

$$Z_{\text{wet},N}^{\beta,\delta} \sim C_{\text{wet}} e^{h_\beta(\delta)N} \quad \text{where} \quad C_{\text{wet}} := \left( e^\delta \sum_{t \geq 1} t K_\beta(t) e^{-h_\beta(\delta)t} \right)^{-1},$$

(b) if  $\delta < \tilde{\delta}_c(\beta)$ , then as  $N \rightarrow \infty$ ,

$$Z_{\text{wet},N}^{\beta,\delta} \sim \frac{c e^{-\delta}}{(e^{-\delta} - 1 + e^{-\beta/2})^2} N^{-3/2},$$

(c) if  $\delta = \tilde{\delta}_c(\beta)$ , then as  $N \rightarrow \infty$ ,

$$Z_{\text{wet},N}^{\beta,\tilde{\delta}_c(\beta)} \sim \frac{1 - e^{-\beta/2}}{2\pi c} N^{-1/2}.$$

Recall (1.A.2) for the definition of  $c$ . The first claim follows from [84, (2.2)] and above observations regarding  $K_\beta$ . The explicit formula for  $h_\beta(\delta)$  is obtained by computing

$$\sum_{t \geq 1} K_\beta(t) e^{-\zeta t} = (1 - e^{-\beta/2}) \mathbf{E}_{\beta,0}[e^{-\zeta \tau}] = 1 - e^a e^{-\beta/2},$$

for some  $a > 0$ , where we used that  $(e^{-aX_n - \zeta n})_{n \geq 0}$  is a martingale when  $a > 0$  is the unique solution of  $\mathbf{E}_{\beta,0}[e^{-aX_1 - \zeta}] = 1$ , along with a stopping time argument; details are left to the reader. The asymptotics of  $Z_{\text{wet},N}^{\beta,\delta}$  as  $N \rightarrow \infty$  are given in [84, Thm. 2.2], where we recall (1.A.2).

## 1.B An FKG inequality

We provide here a (conditional) FKG inequality for random walks of length  $N \in \mathbb{N}$  with increment distributed as  $\mathbf{P}_\beta$ , similarly to more general, already well-known FKG inequalities as developed in [136]. Define a partial order on  $\mathbb{Z}^N$ : for all  $u, v \in \mathbb{Z}^N$ ,  $u \preceq v$  if and only if  $u_k \leq v_k$  for all  $1 \leq k \leq N$ . A function  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  is *non-decreasing* (resp. *non-increasing*) if for all  $u, v \in \mathbb{Z}^N$ ,  $u \preceq v$  implies  $f(u) \leq f(v)$  (resp.  $f(u) \geq f(v)$ ). Define also for every  $u, v \in \mathbb{Z}^N$ ,

$$\begin{aligned} u \wedge v &:= (\min(u_1, v_1), \dots, \min(u_N, v_N)), \\ u \vee v &:= (\max(u_1, v_1), \dots, \max(u_N, v_N)). \end{aligned} \tag{1.B.1}$$

**Proposition 1.B.1.** *Let  $A \subseteq \mathbb{Z}^N$  be such that*

- $\mathbf{P}_{\beta,0}(A) > 0$ ,
- for any  $u, v \in A$ , one has  $u \wedge v \in A$  and  $u \vee v \in A$ .

Let  $f, g : A \rightarrow \mathbb{R}$  be two non-increasing (or two non-decreasing), non-negative and bounded functions. Then,

$$\mathbf{E}_{\beta,0}[f(X)g(X) \mid X \in A] \geq \mathbf{E}_{\beta,0}[f(X) \mid X \in A] \mathbf{E}_{\beta,0}[g(X) \mid X \in A]. \quad (1.B.2)$$

Note that this claim holds in particular for  $A = B_N^{0,+}$  or  $A = \mathbb{Z}^N$ .

*Proof.* We prove the proposition for  $f, g$  both non-increasing. Let  $u, v \in A$ . For every subset  $B \subseteq A$ , let us define  $\mu_1(B) := \mathbf{P}_{\beta,0}(B \mid A)$ , and

$$\mu_2(B) := \frac{\mathbf{E}_{\beta,0}[g(X)1_B(X) \mid X \in A]}{\mathbf{E}_{\beta,0}[g(X) \mid X \in A]},$$

where we assume without loss of generality that  $\mathbf{E}_{\beta,0}[g(X) \mid X \in A] > 0$  —otherwise  $g \equiv 0$  on  $A$  and the proposition is trivial. Notice that  $\mu_1$  and  $\mu_2$  are probability measures on  $A$ . To apply an FKG inequality from [136], we must prove that for any  $u, v \in A$ ,

$$\mu_1(u \vee v) \mu_2(u \wedge v) \geq \mu_1(u) \mu_2(v). \quad (1.B.3)$$

To that end we multiply the left hand side in (1.B.3) by  $\mathbf{E}_{\beta,0}[g(X) \mid A] \times \mathbf{P}_{\beta,0}(A)^2$ , to obtain

$$\begin{aligned} & \mathbf{E}_{\beta,0}[g(X) \mid A] \mathbf{P}_{\beta,0}(A)^2 \mu_1(u \vee v) \mu_2(u \wedge v) \\ &= g(u \wedge v) \mathbf{P}_{\beta,0}(X = u \vee v) \mathbf{P}_{\beta,0}(X = u \wedge v) \\ &= g(u \wedge v) c_\beta^{-2N} \prod_{k=1}^N e^{-\frac{\beta}{2}|(u_k \vee v_k) - (u_{k-1} \vee v_{k-1})|} \prod_{k=1}^N e^{-\frac{\beta}{2}|(u_k \wedge v_k) - (u_{k-1} \wedge v_{k-1})|}, \end{aligned} \quad (1.B.4)$$

where  $u_0 = v_0 := 0$ . Moreover we claim that for any  $a, b, c, d \in \mathbb{R}$ ,

$$|a \vee b - c \vee d| + |a \wedge b - c \wedge d| \leq |a - c| + |b - d|. \quad (1.B.5)$$

To prove this, observe that  $|a - c| = a + c - 2(a \wedge c)$  for any  $a, c \in \mathbb{R}$ , hence (1.B.5) is equivalent to

$$a \wedge c + b \wedge d \leq (a \vee b) \wedge (c \vee d) + (a \wedge b) \wedge (c \wedge d).$$

Notice that  $(a \wedge b) \wedge (c \wedge d) = \min\{a, b, c, d\}$  is either equal to  $a \wedge c$  or  $b \wedge d$ . Assume  $a \wedge b \wedge c \wedge d = a \wedge c$  without loss of generality. Then one obviously has  $b \wedge d \leq (a \vee b) \wedge (c \vee d)$ , which concludes the proof of (1.B.5).

Using (1.B.5) in (1.B.4) and recalling that  $g$  is non-increasing, we conclude

$$\begin{aligned} & \mathbf{E}_{\beta,0}[g(X) \mid A] \mathbf{P}_{\beta,0}(A)^2 \mu_1(u \vee v) \mu_2(u \wedge v) \\ & \geq g(v) c_\beta^{-2N} \prod_{k=1}^N e^{-\frac{\beta}{2}|u_k - u_{k-1}|} \prod_{k=1}^N e^{-\frac{\beta}{2}|v_k - v_{k-1}|} \\ & \geq \mathbf{E}_{\beta,0}[g(X) \mid A] \mathbf{P}_{\beta,0}(A)^2 \mu_1(u) \mu_2(v), \end{aligned}$$

which eventually proves (1.B.3).

Since  $f$  is also a non-increasing function on  $A$ , we can use a generalized FKG inequality claimed in [136, Theorem 3] to obtain

$$\int_A f d\mu_1 \leq \int_A f d\mu_2, \quad (1.B.6)$$

which concludes the proof. □



# Sharp asymptotics of the partition function for the collapsed Interacting Partially Directed Self-Avoiding Walk

## 2.1 Introduction

Identifying the behavior of the partition function of a lattice polymer model is in general a challenging question that sparked interest in both the physical and mathematical literature. In a recent survey [94], some polymer models are reviewed that have in common that their partition functions is of the form

$$Q_L \underset{L \rightarrow \infty}{\sim} B \mu^L \mu_1^{L^\sigma} L^g \quad (2.1.1)$$

where  $B, \mu, \mu_1, \sigma, g$  are real constants depending on the coupling parameters of the model. Among these model, the Interacting Partially Directed Self-Avoiding Walk (referred to under the acronym IPDSAW) which accounts for an homopolymer dipped in a poor (i.e., repulsive) solvent is conjectured to satisfy (2.1.1) inside its collapsed phase. To be more specific, based on numerics displayed in [94] and in [131] (with simulations up to size  $L = 6000$ ), the values of  $\sigma$  and  $g$  are conjectured to be  $1/2$  and  $-3/4$ . In an earlier paper [53, Theorem 2.1], analytic expressions where displayed for  $\mu$  and  $\mu_1$  while  $\sigma$  was proven to be  $1/2$ . In the present paper, we give a full proof of (2.1.1) for the IPDSAW in its collapsed phase, in particular we prove that  $g = -3/4$  and give an analytic expression of  $B$ .

---

This chapter is joint work with Nicolas Pétrélis, and is currently in reviewing process [120].

### 2.1.1 Model

The IPDSAW was initially introduced in [155]. The spatial configurations of the polymer are modeled by the trajectories of a *self-avoiding* random walk on  $\mathbb{Z}^2$  that only takes unitary steps *upwards, downwards and to the right* (see Fig. 2.1). To take into account the monomer-solvent interactions, one considers that, when dipped in a poor solvent, the monomers try to exclude the solvent and therefore attract one another. For this reason, any non-consecutive vertices of the walk though adjacent on the lattice are called *self-touchings* (see Fig. 2.1) and the interactions between monomers are taken into account by assigning an energetic reward  $\beta \geq 0$  to the polymer for each self-touching.

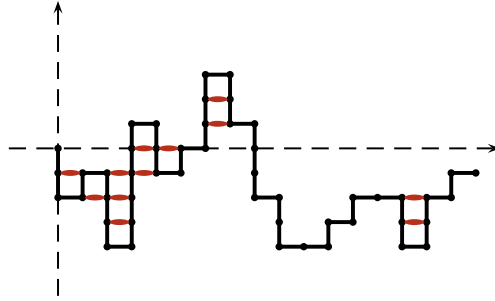


Figure 2.1: Representation of an IPDSAW trajectory with length  $L = 48$  and horizontal extension  $N = 17$ . Its self-touchings are shown in red.

It is convenient to represent the configurations of the model as families of oriented vertical stretches separated by horizontal steps. To be more specific, for a polymer made of  $L \in \mathbb{N}$  monomers, the set of allowed path is  $\Omega_L := \bigcup_{N=1}^L \mathcal{L}_{N,L}$ , where  $\mathcal{L}_{N,L}$  consists of all families made of  $N$  vertical stretches that have a total length  $L - N$ , that is

$$\mathcal{L}_{N,L} = \left\{ \ell := (\ell_i)_{i=1}^N \in \mathbb{Z}^N : \sum_{n=1}^N |\ell_n| + N = L \right\}. \quad (2.1.2)$$

With this representation, the modulus of a given stretch corresponds to the number of monomers constituting this stretch (and the sign gives the direction upwards or downwards). For convenience, we require every configuration to end with an horizontal step, and we note that any two consecutive vertical stretches are separated by a step placed horizontally. The latter explains why  $\sum_{n=1}^N |\ell_n|$  must equal  $L - N$  in order for  $\ell = (\ell_i)_{i=1}^N$  to be associated with a polymer made of  $L$  monomers (see Fig. 2.1).

As mentioned above, the attraction between monomers is taken into account in the Hamiltonian associated with each path  $\ell \in \Omega_L$ , by rewarding energetically those pairs of consecutive stretches with opposite directions, i.e.,

$$H_{L,\beta}(\ell_1, \dots, \ell_N) = \beta \sum_{n=1}^{N-1} (\ell_n \tilde{\wedge} \ell_{n+1}), \quad (2.1.3)$$

where

$$x \tilde{\wedge} y = \begin{cases} |x| \wedge |y| & \text{if } xy < 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2.1.4)$$

One can already note that large Hamiltonians will be assigned to trajectories made of few but long vertical stretches with alternating signs. Such paths will be referred to as collapsed configurations. With the Hamiltonian in hand we can define the polymer measure as

$$P_{L,\beta}(\ell) = \frac{e^{H_{L,\beta}(\ell)}}{Z_{L,\beta}}, \quad \ell \in \Omega_L, \quad (2.1.5)$$

where  $Z_{L,\beta}$  is the partition function of the model, i.e.,

$$Z_{L,\beta} = \sum_{N=1}^L \sum_{\ell \in \mathcal{L}_{N,L}} e^{H_{L,\beta}(\ell)}. \quad (2.1.6)$$

## 2.2 Main results

The IPDSAW undergoes a collapse transition at some  $\beta_c > 0$  (see [131] or [130, Theorem 1.3]) dividing  $[0, \infty)$  into an extended phase  $\mathcal{E} := [0, \beta_c)$  and a collapsed phase  $\mathcal{C} := [\beta_c, \infty)$ . Note that, when the monomer-monomer attraction is switched off ( $\beta = 0$ ), a typical configuration (sampled from  $P_{L,\beta}$ ) has an horizontal extension  $O(L)$  since, roughly speaking, every step has a positive probability (bounded from below) to be horizontal. In the extended phase, the interaction intensity  $\beta$  is not yet strong enough to bring this typical horizontal extension from  $O(L)$  to  $o(L)$ . Inside the collapsed phase, in turn, the interaction intensity is large enough to change dramatically the geometric features of a typical trajectory, which roughly looks like a compact ball with an horizontal extension  $o(L)$ . The asymptotics of  $(Z_{L,\beta})_{L \geq 1}$  are displayed and proven in [53, Theorem 2.1 (1) and (2)] for the extended phase and at criticality ( $\beta = \beta_c$ ). Inside the collapsed phase, although the exponential terms of the partition function growth rate were identified via upper and lower bounds (see [53, Theorem 2.1 (iii)]), a full proof of such asymptotics was missing. We close this gap with Theorem 2.2.1 below, by identifying the polynomial prefactor. This improvement relies on a sharp local limit theorem displayed in Proposition 2.4.5 for some random walk in a large deviation regime (more precisely it is constrained to be positive, cover a large area and end in 0).

Let us settle some notations that are required to state Theorem 2.2.1. For  $\beta > 0$  we let  $\mathbf{P}_\beta$  be the following discrete Laplace probability law on  $\mathbb{Z}$ :

$$\mathbf{P}_\beta(\cdot = k) = \frac{e^{-\frac{\beta}{2}|k|}}{c_\beta}, \quad c_\beta := \sum_{k \in \mathbb{Z}} e^{-\frac{\beta}{2}|k|} = \frac{1 + e^{-\beta/2}}{1 - e^{-\beta/2}}, \quad (2.2.1)$$

and we denote by  $\mathcal{L}(h)$  the logarithmic moment generating function of  $Z$  a random variable of law  $\mathbf{P}_\beta$ , i.e.,

$$\mathcal{L}(h) := \log \mathbf{E}_\beta[e^{hZ}], \quad h \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right). \quad (2.2.2)$$



Moreover,  $\beta_c$  is the unique positive solution to the equation  $\frac{c_\beta}{e^\beta} = 1$  (see [130, Theorem 1.3]).

We define below the escape probability of a certain class of drifted random walks and we display the exponential decay rate of the probability that some random walk trajectories enclose an atypically large area.

**Tilting and escape probability.**

For  $|h| < \beta/2$  we let  $\tilde{\mathbf{P}}_h$  be the probability law on  $\mathbb{Z}$  defined by perturbing  $\mathbf{P}_\beta$  as

$$\frac{d\tilde{\mathbf{P}}_h}{d\mathbf{P}_\beta}(k) = e^{hk - \mathcal{L}(h)} \quad k \in \mathbb{Z}. \quad (2.2.3)$$

For  $h \in (0, \beta/2)$ , we consider a random walk  $X := (X_i)_{i \in \mathbb{N}}$  such that  $X_0 = 0$  and whose increments  $(X_i - X_{i-1})_{i \in \mathbb{N}}$  are i.i.d. with law  $\tilde{\mathbf{P}}_h$  (i.e., with a positive drift). We denote by  $\kappa(h)$  the probability that  $X$  never returns to the lower half plan, that is

$$\kappa(h) := \tilde{\mathbf{P}}_h(X_i > 0, \forall i \in \mathbb{N}) = \frac{e^{2h} - 1}{e^{h+\beta/2} - 1}, \quad (2.2.4)$$

where the second identity will be proven in Lemma 2.5.13.

**Rate function for the area of a positive excursion.**

We let  $\mathcal{G} : (-\beta, \beta) \rightarrow \mathbb{R}$  be defined as

$$\mathcal{G}(h) := \int_0^1 \mathcal{L}(h(\frac{1}{2} - x)) dx, \quad \text{for } h \in (-\beta, \beta). \quad (2.2.5)$$

We will prove in Lemma 2.5.3 that  $\mathcal{G}'$  is a  $\mathcal{C}^1$  diffeomorphism from  $(-\beta, \beta)$  to  $\mathbb{R}$  and we let  $q \in \mathbb{R} \mapsto \tilde{h}^q$  be its inverse function. Considering  $X := (X_i)_{i \in \mathbb{N}}$  a random walk starting from the origin, whose increments  $(X_i - X_{i-1})_{i \in \mathbb{N}}$  are i.i.d. with law  $\mathbf{P}_\beta$ , we denote by  $A_N(X)$  (or  $A_N$  when there is no risk of confusion) the algebraic area enclosed by  $X$  up to time  $N$ , i.e.,

$$A_N(X) := X_1 + \dots + X_N. \quad (2.2.6)$$

We will prove in Proposition 2.4.5 below that, for  $q \in [0, \infty) \cap \frac{\mathbb{N}}{N^2}$ , the exponential decay rate of the event  $\{A_N = qN^2, X_N = 0, X_i > 0, 1 \leq i \leq N - 1\}$  is given by  $\tilde{\psi} : (0, \infty) \mapsto \mathbb{R}$  taking value

$$\tilde{\psi}(q) := q\tilde{h}^q - \mathcal{G}(\tilde{h}^q), \quad \text{for } q \in (0, \infty). \quad (2.2.7)$$

At this stage, we define a positive constant that depends on  $\beta$  and  $q > 0$  as

$$C_{\beta,q} := \frac{1}{2\pi} \vartheta(\tilde{h}^q)^{\frac{1}{2}} \kappa\left(\frac{\tilde{h}^q}{2}\right)^2, \quad (2.2.8)$$

where  $\vartheta : h \in (-\beta/2, \beta/2) \mapsto \mathbb{R}$  takes value

$$\vartheta(h) = \int_0^1 x^2 \mathcal{L}''[h(x - \frac{1}{2})] dx \int_0^1 \mathcal{L}''[h(x - \frac{1}{2})] dx - \left[ \int_0^1 x \mathcal{L}''[h(x - \frac{1}{2})] dx \right]^2. \quad (2.2.9)$$

**Theorem 2.2.1.** *For any  $\beta > \beta_c$ , one has*

$$Z_{L,\beta} \underset{L \rightarrow \infty}{\sim} \frac{K_\beta}{L^{3/4}} e^{\beta L + \tilde{\mathcal{G}}(a_\beta) \sqrt{L}}, \quad (2.2.10)$$

with

$$\tilde{\mathcal{G}}(x) := x \log \frac{c_\beta}{e^\beta} - x \tilde{\psi}(x^{-2}), \quad x > 0, \quad (2.2.11)$$

and

$$a_\beta := \arg \max \{ \tilde{\mathcal{G}}(x) : x \in ]0, \infty[ \}, \quad (2.2.12)$$

and

$$K_\beta := \frac{2\sqrt{2\pi} C_{\beta, a_\beta^{-2}} e^{\tilde{\psi}'(a_\beta^{-2})}}{\left[ (1 + e^{-\beta}) e^{\operatorname{arccosh}(e^{-\beta/2} \cosh(\beta))} - e^{\beta/2} (1 - e^{-\beta}) \right]^2 a_\beta^2 |\tilde{\mathcal{G}}''(a_\beta)|^{1/2}}. \quad (2.2.13)$$

Theorem 2.2.1 is proven in Section 2.4. The proof will require to decompose a trajectory into a succession of *beads* that are sub-trajectories made of non-zero vertical stretches of alternating signs. Inside the collapsed phase, an issue raised by physicists was to understand whether a typical trajectory contains a unique macroscopic bead or not. Thus, for every  $\ell \in \Omega_L$  we let  $N_\ell$  be its horizontal extension (i.e.,  $\ell \in \mathcal{L}_{N_\ell, L}$ ) and also  $|I_{\max}(\ell)|$  be the length of its largest bead, i.e.,

$$|I_{\max}(\ell)| := \max \left\{ \sum_{i=u}^v 1 + |\ell_i| : 1 \leq u \leq v \leq N_\ell, \ell_i \ell_{i+1} < 0 \ \forall u \leq i \leq v-1 \right\}. \quad (2.2.14)$$

With [52, Thm. C] it is known that a typical trajectory indeed contains a unique macroscopic bead, and that at most  $(\log L)^4$  monomers lay outside this large bead. We improve this result with the following theorem, by showing that only finitely many monomers are to be found outside the unique macroscopic bead.

**Theorem 2.2.2.** *For any  $\beta > \beta_c$ ,*

$$\lim_{k \rightarrow \infty} \liminf_{L \rightarrow \infty} \mathbf{P}_{L,\beta}(|I_{\max}(\ell)| \geq L - k) = 1. \quad (2.2.15)$$

## Outline of the paper

With Section 2.3 below, we introduce some mathematical tools of particular importance for the rest of the paper. Thus, in Section 2.3.1 we define rigorously the set containing the single-bead trajectories. Such beads allow us to decompose any paths in  $\Omega_L$  into sub-trajectories that are not interacting with each other. Under the polymer measure, the cumulated lengths of those beads form a renewal process which will be of key importance throughout the paper. Section 2.3.2 is dedicated to the random walk representation of the model, that was introduced initially in [130] and provides a probabilistic expression for every partition function introduced in the paper. In Section 2.4 we prove Theorem 2.2.1 subject to Proposition 2.4.1 which gives sharp asymptotics for those partition functions associated with single bead trajectories. Proposition 2.4.1 is proven afterwards subject to Proposition 2.4.5, which is the main feature of this paper since it strongly improves previous estimates (such as [52, Prop. 2.4–2.5]) by providing an equivalent to the probability that a random walk  $X := (X_i)_{i=0}^N$  of law  $\mathbf{P}_\beta$  describes a positive excursion, ends up at 0 and encloses an area  $qN^2$  for  $q > 0$ . The proof of Proposition 2.4.5 is divided into 4 steps, displayed in Section 2.5 after we introduce a tilted law for the random walk  $X$  in such a way that the event  $\{A_N(X) = qN^2, X_N = 0\}$  becomes typical. Note that Steps 3 and 4 from Section 2.5 require local limit estimates which are displayed in Appendix 2.A, and Proposition 2.5.4 is proven in Appendix 2.B. Finally, Theorem 2.2.2 is proven in Section 2.6 by using mostly the asymptotics provided by Theorem 2.2.1.

## 2.3 Preparations

### 2.3.1 One bead trajectories and renewal structure

Let  $L \in \mathbb{N}$  and denote by  $\Omega_L^\circ$  the subset of  $\Omega_L$  gathering those single-bead trajectories, i.e., trajectories made of non-zero vertical stretches that alternate orientations. Thus, we set  $\Omega_L^\circ = \cup_{N=1}^{L/2} \mathcal{L}_{N,L}^\circ$  with

$$\mathcal{L}_{N,L}^\circ := \left\{ (\ell_i)_{i=1}^N \in \mathbb{Z}^N : \sum_{i=1}^N |\ell_i| = L - N, \quad \ell_i \ell_{i+1} < 0 \quad \forall 1 \leq i < N \right\}. \quad (2.3.1)$$

We also denote by  $Z_{L,\beta}^\circ$  the partition function restricted to those trajectories in  $\Omega_L^\circ$ , i.e.,

$$Z_{L,\beta}^\circ := \sum_{N=1}^{L/2} \sum_{\ell \in \mathcal{L}_{N,L}^\circ} e^{\beta H(\ell)}. \quad (2.3.2)$$

In order to decompose a general trajectory into single-bead subpaths, we need to deal with the zero-length vertical stretches. This has to be done in such a way that the decomposition gives rise to a true renewal structure, namely that the realization of a single-bead trajectory has no influence on the value of the partition function associated with the

next bead. To that aim, we will integrate the zero-length stretches at the beginning of beads and for that we define  $\widehat{\Omega}^\circ$  the set of *extended beads* as

$$\widehat{\Omega}^\circ := \bigcup_{L \geq 2} \widehat{\Omega}_L^\circ \quad (2.3.3)$$

where  $\widehat{\Omega}_L^\circ$  is the subset of  $\Omega_L$  gathering those trajectories which may or may not start with a sequence of zero-length stretches and form subsequently a unique bead, i.e.,

$$\widehat{\Omega}_L^\circ := \bigcup_{k=0}^{L-2} \widehat{\Omega}_L^{\circ, k} \quad (2.3.4)$$

with

$$\widehat{\Omega}_L^{\circ, 0} := \left\{ \ell \in \Omega_L^\circ : \ell_1 > 0 \right\} \quad (2.3.5)$$

$$\widehat{\Omega}_L^{\circ, k} := \left\{ \ell \in \Omega_L : N_\ell > k, \ell_1 = \dots = \ell_k = 0, (\ell_{i+k})_{i=1}^{N_\ell-k} \in \Omega_{L-k}^\circ \right\}, \quad k \in \{1, \dots, L-2\}.$$

We recall (2.3.2) and (2.3.4) and the partition function restricted to those trajectories in  $\widehat{\Omega}_L^\circ$  becomes:

$$\widehat{Z}_{L,\beta}^\circ := \sum_{k=0}^{L-2} \left[ \frac{1}{2} \mathbf{1}_{\{k=0\}} Z_{L,\beta}^\circ + \mathbf{1}_{\{k \in \mathbb{N}\}} Z_{L-k,\beta}^\circ \right]. \quad (2.3.6)$$

Note that, in the definition of  $\widehat{\Omega}_L^{\circ, 0}$  in (2.3.5) the sign of  $\ell_1$  is prescribed because if an extended bead does not start with a zero-length stretch, then the sign of its first stretch must be the same as that of the last stretch of the preceding bead; hence the factor  $\frac{1}{2}$  in (2.3.6) (see Fig. 2.1). Of course this latter restriction does not apply to the very first extended bead of a trajectory and this is why we define

$$\bar{Z}_{L,\beta}^\circ := \sum_{k=0}^{L-2} Z_{L-k,\beta}^\circ. \quad (2.3.7)$$

For convenience, we define  $\Omega_L^c$  the subset of  $\Omega_L$  containing those trajectories ending with a non-zero stretch, i.e.,

$$\Omega_L^c := \{ \ell \in \Omega_L : \ell_{N_\ell} \neq 0 \}. \quad (2.3.8)$$

At this stage we can decompose a given trajectory  $\ell \in \Omega_L^c$  into extended beads by cutting the trajectory at times  $(\tau_j)_{j=0}^{n(\ell)}$  defined as  $\tau_0 = 0$  and for  $j \in \mathbb{N}$  such that  $\tau_{j-1} < N_\ell$

$$\tau_j := \max \left\{ s > \tau_{j-1} : (\ell_i)_{i=1+\tau_{j-1}}^s \in \widehat{\Omega}^\circ \quad \text{or} \quad (-\ell_i)_{i=1+\tau_{j-1}}^s \in \widehat{\Omega}^\circ \right\}. \quad (2.3.9)$$

Then  $n(\ell)$  is the number of beads composing  $\ell$  and satisfies  $\tau_{n(\ell)} = N_\ell$ , thus

$$\ell = \odot_{j=1}^{n(\ell)} \mathcal{B}_j \quad \text{with} \quad \mathcal{B}_j := (\ell_{\tau_{j-1}+1}, \dots, \ell_{\tau_j}), \quad (2.3.10)$$

where  $\odot$  denotes the concatenation. We also set  $\mathfrak{X}_0 = 0$  and for  $j \in \{1, \dots, n_\ell\}$ , we denote by  $\mathfrak{X}_j - \mathfrak{X}_{j-1}$  the number of steps (or monomers) that the  $j$ -th bead is made of (also referred to as *total length* of the bead), that is,

$$\mathfrak{X}_j - \mathfrak{X}_{j-1} = \tau_j - \tau_{j-1} + |\ell_{\tau_{j-1}+1}| + \dots + |\ell_{\tau_j}|, \quad j \in \{1, \dots, n_\ell\}. \quad (2.3.11)$$

The set  $\mathfrak{X} := \{0, \mathfrak{X}_1, \dots, \mathfrak{X}_{n_\ell}\}$  contains the cumulated lengths of the beads forming  $\ell$ , in particular  $\mathfrak{X}_{n_\ell} = L$ .

**Remark 2.3.1.** Note that a trajectory  $\ell \in \Omega_L \setminus \Omega_L^c$  may also be decomposed into extended beads as in (2.3.9–2.3.11). The only difference is that the very last bead is followed by a sequence of zero-length vertical stretches, i.e.,  $\mathfrak{X}_{n_\ell} = L - k$  for some  $k \in \{1, \dots, L\}$  and the last  $k$  vertical stretches in  $\ell$  have zero-length.

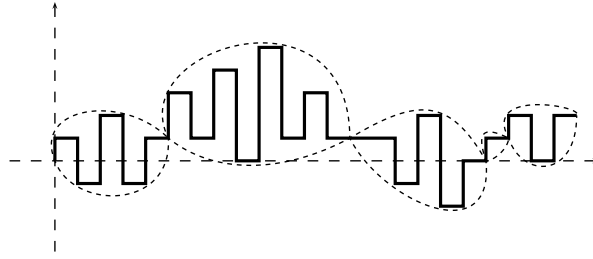


Figure 2.1: Decomposition of an IPDSAW trajectory into beads. The first bead of the decomposition may start with a stretch with any sign, whereas 2nd, 4th and 5th beads start with non-zero stretches with constrained signs, and the 3rd bead starts with two zero-length stretches. Notice that the last vertical stretch of this trajectory is non-zero, so it lies in  $Z_{L,\beta}^c$ .

By using this bead-decomposition we can rewrite  $Z_{L,\beta}^c$  the partition function restricted to  $\Omega_L^c$  as

$$\begin{aligned} Z_{L,\beta}^c &= \sum_{r=1}^{L/2} \sum_{t_1+\dots+t_r=L} Z_{L,\beta}^c(n(\ell) = r, \mathfrak{X}_i - \mathfrak{X}_{i-1} = t_i, \forall 1 \leq i \leq r) \\ &= \sum_{r=1}^{L/2} \sum_{t_1+\dots+t_r=L} \bar{Z}_{t_1,\beta}^\circ \prod_{j=2}^r \hat{Z}_{t_j,\beta}^\circ, \end{aligned} \quad (2.3.12)$$

where for  $D \subset \Omega_L^c$ , we denote by  $Z_{L,\beta}^c(D)$  the partition function restricted to  $D$  (see Fig. 2.1).

## 2.3.2 Probabilistic representation

The aim of this section is to give a probabilistic expression of the partition function  $Z_{L,\beta}^\circ$  and to use it subsequently to provide closed expression of the generating functions associated with  $(\hat{Z}_{L,\beta}^\circ)_{L \geq 2}$  and with  $(\bar{Z}_{L,\beta}^\circ)_{L \geq 2}$ .

We recall the definition of  $\mathbf{P}_\beta$  in (2.2.1) and for  $x \in \mathbb{Z}$  we denote by  $\mathbf{P}_{\beta,x}$  the law of a random walk  $X := (X_i)_{i \geq 0}$  starting from  $x$  (i.e.,  $X_0 = x$ ) and such that  $(X_{i+1} - X_i)_{i \geq 0}$  is an i.i.d. sequence of random variables with law  $\mathbf{P}_\beta$ . In the case  $x = 0$ , we will omit the  $x$ -dependence of  $\mathbf{P}_{\beta,x}$  when there is no risk of confusion. We also recall from (2.2.6) that  $A_N$  defines the algebraic area enclosed in-between a random walk trajectory and the  $x$ -axis after  $N$  steps.

Recall (2.3.1) and (2.3.2), and let us now briefly remind the transformation that allows us to give a probabilistic representation of  $Z_{L,\beta}^\circ$  (we refer to [54] for a review on the recent progress made on IPDSAW by using probabilistic tools). First, note that for  $x, y \in \mathbb{Z}$  one can write  $x \tilde{\wedge} y = \frac{1}{2}(|x| + |y| - |x + y|)$ . By using the latter equality to compute the Hamiltonian (recall (2.1.3)), we may rewrite (2.3.2) as

$$\begin{aligned} Z_{L,\beta}^\circ &= \sum_{N=1}^{L/2} \sum_{\substack{\ell \in \mathcal{L}_{N,L}^\circ \\ \ell_0 = \ell_{N+1} = 0}} \exp\left(\beta \sum_{n=1}^N |\ell_n| - \frac{\beta}{2} \sum_{n=0}^N |\ell_n + \ell_{n+1}|\right) \\ &= c_\beta e^{\beta L} \sum_{N=1}^{L/2} \left(\frac{c_\beta}{e^\beta}\right)^N \sum_{\substack{\ell \in \mathcal{L}_{N,L}^\circ \\ \ell_0 = \ell_{N+1} = 0}} \prod_{n=0}^N \frac{\exp\left(-\frac{\beta}{2} |\ell_n + \ell_{n+1}|\right)}{c_\beta}. \end{aligned} \quad (2.3.13)$$

Henceforth, for convenience, we will assume that any  $\ell \in \mathcal{L}_{N,L}^\circ$  satisfies  $\ell_0 = \ell_{N+1} = 0$ . At this stage, we denote by  $B_N^+$  the set of those  $N$ -step integer-valued random walk trajectories, starting and ending at 0 and remaining positive in-between, i.e.,

$$B_N^+ := \left\{ (x_i)_{i=0}^N \in \mathbb{Z}^{N+1} : X_0 = X_N = 0, X_i > 0 \ \forall 0 < i < N \right\}. \quad (2.3.14)$$

It remains to notice that the application

$$\begin{aligned} T_N : \{ \ell \in \mathcal{L}_{N,L}^\circ : \ell_1 > 0 \} &\mapsto \left\{ (x_i)_{i=0}^{N+1} \in B_{N+1}^+ : A_N(x) = L - N \right\} \\ (\ell_i)_{i=0}^{N+1} &\rightarrow ((-1)^{i-1} \ell_i)_{i=0}^{N+1} \end{aligned} \quad (2.3.15)$$

is a one-to-one correspondence, and that for  $\ell \in \mathcal{L}_{N,L}^\circ$  the increments of  $T_N(\ell)$  are in modulus equal to  $(|\ell_{i-1} + \ell_i|)_{i=1}^{N+1}$ . Therefore, set  $\Gamma_\beta := c_\beta/e^\beta$  and (2.3.13) becomes

$$Z_{L,\beta}^\circ e^{-\beta L} = 2 c_\beta \sum_{N=1}^{L/2} \Gamma_\beta^N \mathbf{P}_\beta \left( X \in B_{N+1}^+, A_N(X) = L - N \right) \quad (2.3.16)$$

where the factor 2 in the r.h.s. in (2.3.16) is required to take into account those  $\ell \in \mathcal{L}_{N,L}^\circ$  satisfying  $\ell_1 < 0$ .

With the next Lemma, we provide exact expressions for the two generating functions  $\sum_{L \geq 2} \widehat{Z}_{L,\beta}^\circ z^L$  and  $\sum_{L \geq 2} \bar{Z}_{L,\beta}^\circ z^L$  at  $z = e^{-\beta}$ . This is needed to determine  $K_\beta$  in Theorem 2.2.1.

**Lemma 2.3.2.** For  $\beta > 0$ ,

$$\begin{aligned}\delta_1(\beta) &:= \sum_{L \geq 2} \bar{Z}_{L,\beta}^\circ e^{-\beta L} = \frac{2e^\beta}{1 - e^{-\beta}} r_\beta \\ \delta_2(\beta) &:= \sum_{L \geq 2} \hat{Z}_{L,\beta}^\circ e^{-\beta L} = e^\beta c_{2\beta} r_\beta\end{aligned}\tag{2.3.17}$$

with  $r_\beta := \mathbf{E}_\beta[\mathbf{1}_{\{X_1 > 0\}} \mathbf{1}_{\{X_\rho = 0\}} (\Gamma_\beta)^\rho]$  where  $\rho := \inf\{i \geq 1 : X_i \leq 0\}$ . Moreover,

$$r_\beta = \begin{cases} +\infty & \text{if } \beta < \beta_c, \\ 1 - e^{-\beta} - e^{-\frac{\beta}{2} + \operatorname{arccosh}(e^{-\beta/2} \cosh(\beta))} & \text{if } \beta \geq \beta_c. \end{cases}\tag{2.3.18}$$

*Proof.* We recall (2.3.16) and we start by computing

$$\begin{aligned}\sum_{L \geq 2} Z_{L,\beta}^\circ e^{-\beta L} &= 2c_\beta \sum_{N \geq 1} \Gamma_\beta^N \sum_{L \geq 2N} \mathbf{P}_\beta(X \in B_{N+1}^+, A_N(X) = L - N) \\ &= 2c_\beta \sum_{N \geq 1} \Gamma_\beta^N \mathbf{P}_\beta(X \in B_{N+1}^+) \\ &= 2e^\beta r_\beta.\end{aligned}\tag{2.3.19}$$

Then, we use (2.3.6) to write

$$\begin{aligned}\sum_{L \geq 2} \hat{Z}_{L,\beta}^\circ e^{-\beta L} &:= \sum_{L \geq 2} \sum_{k=0}^{L-2} e^{-\beta L} \left[ \frac{1}{2} \mathbf{1}_{\{k=0\}} Z_{L,\beta}^\circ + \mathbf{1}_{\{k \in \mathbb{N}\}} Z_{L-k,\beta}^\circ \right] \\ &= \frac{1}{2} \sum_{L \geq 2} e^{-\beta L} Z_{L,\beta}^\circ + \sum_{k=1}^{\infty} e^{-\beta k} \sum_{t=2}^{\infty} e^{-\beta t} Z_{t,\beta}^\circ \\ &= \left( \frac{1}{2} + \frac{e^{-\beta}}{1 - e^{-\beta}} \right) \sum_{L \geq 2} e^{-\beta L} Z_{L,\beta}^\circ,\end{aligned}\tag{2.3.20}$$

and it remains to combine (2.3.19) and (2.3.20) to obtain the second equality in (2.3.17). Using (2.3.7), the very same computation allows us to obtain the first equality in (2.3.17).

Let us now compute  $\mathbf{E}_\beta[(\Gamma_\beta)^\rho]$ , which will yield a formula for  $r_\beta$ . Indeed, the fact that the increments of  $X$  follow a discrete Laplace law entails that  $(\rho, X_1, \dots, X_{\rho-1})$  and  $X_\rho$  are independent, and moreover that  $-X_\rho$  follows a geometric law on  $\mathbb{N} \cup \{0\}$  with parameter  $1 - e^{-\beta/2}$ . Thereby,

$$\begin{aligned}\mathbf{E}_\beta[(\Gamma_\beta)^\rho] &= \mathbf{E}_\beta[\mathbf{1}_{\{X_1 > 0\}} (\Gamma_\beta)^\rho] + \Gamma_\beta \mathbf{P}_\beta(X_1 \leq 0) \\ &= \frac{r_\beta}{\mathbf{P}_\beta(-X_\rho = 0)} + \frac{c_\beta}{e^\beta} \left( \frac{1}{c_\beta} + \frac{1}{2} \left( 1 - \frac{1}{c_\beta} \right) \right) \\ &= \frac{r_\beta + e^{-\beta}}{1 - e^{-\beta/2}},\end{aligned}\tag{2.3.21}$$

which will conclude the proof.

It is a straightforward application of [47, (4.5)] that there exists  $c > 0$  depending on  $\beta$  only such that

$$\mathbf{P}_\beta(\rho = t) \sim \frac{c}{t^{3/2}} \quad \text{as } t \rightarrow \infty. \quad (2.3.22)$$

Recall that  $\beta \mapsto \Gamma_\beta$  is decreasing on  $(0, \infty)$ ; in particular when  $\beta < \beta_c$ , one has  $\Gamma_\beta > 1$ , so (2.3.21) and (2.3.22) imply  $r_\beta = +\infty$ .

When  $\beta \geq \beta_c$ , notice that  $(e^{-\zeta X_n + \log(\Gamma_\beta)n})_{n \geq 0}$  is a martingale under  $\mathbf{P}_{\beta,0}$  if and only if  $\mathcal{L}(-\zeta) = -\log(\Gamma_\beta)$ . Recalling  $\Gamma_\beta \leq 1$  and that  $\mathcal{L}$  is decreasing, not bounded on  $(-\beta/2, 0]$ , there is a unique  $\zeta_\beta \in [0, \beta/2)$  satisfying this equality, given by

$$\zeta_\beta = \operatorname{arccosh}\left((1 - \Gamma_\beta) \cosh(\beta/2) + \Gamma_\beta\right) = \operatorname{arccosh}\left(e^{-\beta/2} \cosh(\beta)\right). \quad (2.3.23)$$

Thanks to a stopping time argument (we do not write the details here), we finally obtain

$$\mathbf{E}_\beta[(\Gamma_\beta)^\rho] = \mathbf{E}_\beta\left[e^{-\zeta_\beta X_\rho}\right]^{-1} = \frac{1 - e^{\zeta_\beta - \beta/2}}{1 - e^{-\beta/2}},$$

which concludes the proof by recollecting (2.3.21) and (2.3.23).  $\square$

To end this section we state and prove the following corollary that will be needed in the proof of Theorem 2.2.1 (see Section 2.4) below.

**Corollary 2.3.3.** *For any  $\beta > \beta_c$ , we have  $\delta_2(\beta) < 1$ .*

*Proof.* We observe that  $\beta \rightarrow \delta_2(\beta)$  is decreasing simply because for every  $L \geq 2$  and every  $\ell \in \Omega_L$  the quantity  $\beta \rightarrow H_{L,\beta}(\ell) - \beta L$  is non-increasing and even decreasing if  $H_{L,\beta}(\ell) < \beta L$  (recall (2.1.3)). Therefore, the corollary will be proven once we show  $\delta_2(\beta_c) = 1$ . Recall that  $\Gamma_{\beta_c} = 1$ , which ensures that  $e^{\beta_c/2}$  is a solution to  $X^3 - X^2 - X - 1 = 0$  and that  $\zeta_{\beta_c} = 0$  in (2.3.23). This implies both

$$r_{\beta_c} = 1 - e^{-\beta_c} - e^{-\beta_c/2} = e^{-3\beta_c/2},$$

and

$$\delta_2(\beta_c) = e^{-\beta_c/2} \left( \frac{1 + e^{-\beta_c}}{1 - e^{-\beta_c}} \right) = 1,$$

which concludes the proof.  $\square$

## 2.4 Proof of Theorem 2.2.1

In this section we provide a sharp estimate for the partition function of single-bead trajectories in Proposition 2.4.1, with which we prove Theorem 2.2.1. The proof of Proposition 2.4.1 is displayed afterwards subject to Proposition 2.4.5 and Lemma 2.4.7.



**Proposition 2.4.1.** For  $\beta > \beta_c$ , there exists  $K_\beta^\circ > 0$  such that

$$Z_{L,\beta}^\circ \underset{L \rightarrow \infty}{\sim} \frac{K_\beta^\circ}{L^{3/4}} e^{\beta L + \tilde{\mathcal{G}}(a_\beta)\sqrt{L}}. \quad (2.4.1)$$

**Corollary 2.4.2.** For  $\beta > \beta_c$ , there exist  $\widehat{K}_\beta > 0$  and  $\bar{K}_\beta > 0$  such that

$$\widehat{Z}_{L,\beta}^\circ \underset{L \rightarrow \infty}{\sim} \frac{\widehat{K}_\beta}{L^{3/4}} e^{\beta L + \tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \quad \text{and} \quad \bar{Z}_{L,\beta}^\circ \underset{L \rightarrow \infty}{\sim} \frac{\bar{K}_\beta}{L^{3/4}} e^{\beta L + \tilde{\mathcal{G}}(a_\beta)\sqrt{L}}. \quad (2.4.2)$$

We observe that (2.4.1) is a substantial improvement of [52, Prop. 4.2], where the polynomial factors were  $\frac{1}{L^\kappa}$  with  $\kappa > 1$  in the lower bound, and  $\frac{1}{\sqrt{L}}$  in the upper bound. Moreover Corollary 2.4.2 is a straightforward consequence of Proposition 2.4.1. Indeed, let us define for convenience

$$h(n) := \frac{1}{n^{3/4}} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{n}}, \quad n \in \mathbb{N}. \quad (2.4.3)$$

Recollecting (2.3.6), we write

$$\frac{e^{-\beta L}}{h(L)} \widehat{Z}_{L,\beta}^\circ = \frac{1}{2} \frac{e^{-\beta L}}{h(L)} Z_{L,\beta}^\circ + \sum_{k=1}^{L-2} e^{-\beta k} \frac{e^{-\beta(L-k)}}{h(L)} Z_{L-k,\beta}^\circ,$$

and similarly for (2.3.7). Noticing that  $h(L) \sim h(L-k)$  as  $L \rightarrow \infty$  for any  $k \in \mathbb{N}$ , (2.4.1) and dominated convergence imply that (2.4.2) holds true with

$$\widehat{K}_\beta = K_\beta^\circ \frac{1 + e^{-\beta}}{2(1 - e^{-\beta})} \quad \text{and} \quad \bar{K}_\beta = K_\beta^\circ \frac{1}{1 - e^{-\beta}}. \quad (2.4.4)$$

*Proof of Theorem 2.2.1 subject to Proposition 2.4.1.* Recalling the definitions of  $\delta_1(\beta)$  and  $\delta_2(\beta)$  from (2.3.17), we define two probability laws  $q_1$  and  $q_2$  on  $\mathbb{N}$  by

$$q_1(t) := \delta_1(\beta)^{-1} \bar{Z}_{t,\beta}^\circ e^{-\beta t} \quad t \geq 2, \quad (2.4.5)$$

$$q_2(t) := \delta_2(\beta)^{-1} \widehat{Z}_{t,\beta}^\circ e^{-\beta t} \quad t \geq 2. \quad (2.4.6)$$

For  $r \in \mathbb{N}$ , we denote by  $q_2^{r*}$  the convolution product of  $r$  times  $q_2$  and by  $q_1 * q_2^{r*}$  the convolution product between  $q_1$  and  $q_2^{r*}$ . This allows us to rewrite (2.3.12) under the form

$$\tilde{Z}_{L,\beta}^c := e^{-\beta L} Z_{L,\beta}^c = \frac{\delta_1(\beta)}{\delta_2(\beta)} \sum_{r \geq 1} \delta_2(\beta)^r \left[ q_1 * q_2^{(r-1)*} \right](L). \quad (2.4.7)$$

Recalling (2.4.3), Corollary 2.4.2 implies that

$$\begin{aligned} q_1(n) &\underset{n \rightarrow \infty}{\sim} u_\beta h(n), \quad \text{with} \quad u_\beta := \frac{\bar{K}_\beta}{\delta_1(\beta)} \\ q_2(n) &\underset{n \rightarrow \infty}{\sim} v_\beta h(n), \quad \text{with} \quad v_\beta := \frac{\widehat{K}_\beta}{\delta_2(\beta)} \end{aligned} \quad (2.4.8)$$

At this stage, Theorem 2.2.1 will be a straightforward consequence of Claims 2.4.3 and 2.4.4 below. Those claims are proven in [74, Corollary 4.13 and Theorem 4.14] in the case where  $q_1 \equiv q_2$ . However, (2.4.8) guarantees that the proof in [74] can easily be adapted to our case since  $q_1$  and  $q_2$  enjoy the same asymptotic behavior (up to a constant).

**Claim 2.4.3.** For  $\beta > 0$  and  $r \in \mathbb{N} \cup \{0\}$  it holds that  $q_1 * q_2^{r*}(n) \sim_{n \rightarrow \infty} (u_\beta + r v_\beta) h(n)$ .

**Claim 2.4.4.** For  $\beta > 0$  and  $\varepsilon > 0$  there exists  $n_0(\varepsilon) \in \mathbb{N}$  and  $C(\varepsilon) > 0$  such that

$$q_1 * q_2^{r*}(n) \leq C(\varepsilon) (1 + \varepsilon)^r h(n), \quad n \geq n_0(\varepsilon), r \in \mathbb{N} \cup \{0\} \quad (2.4.9)$$

It remains to use the dominated convergence Theorem to conclude from Claims 2.4.3 and 2.4.4 that for  $\delta < 1$

$$\lim_{n \rightarrow \infty} \frac{1}{h(n)} \sum_{r \geq 1} \delta^r \left[ q_1 * q_2^{(r-1)*} \right](n) = u_\beta \sum_{r \geq 1} \delta^r + \delta v_\beta \sum_{r \geq 1} r \delta^r = \frac{u_\beta \delta}{1 - \delta} + \frac{v_\beta \delta^2}{(1 - \delta)^2} \quad (2.4.10)$$

Combining (2.4.7) with (2.4.10) at  $\delta = \delta_2(\beta)$  (recall that  $\delta_2(\beta) < 1$  by Corollary 2.3.3) we obtain that

$$\lim_{L \rightarrow \infty} \frac{1}{h(L)} \tilde{Z}_{L,\beta}^c = \frac{\bar{K}_\beta}{1 - \delta_2(\beta)} + \frac{\widehat{K}_\beta \delta_1(\beta)}{(1 - \delta_2(\beta))^2}. \quad (2.4.11)$$

To complete the proof of Theorem 2.2.1, it remains to take into account trajectories that end with some zero-length stretches. To that aim, we recall (2.3.8) and we partition  $\Omega_L$  into subsets whose trajectories are ending with a prescribed number of zero-length stretches, i.e.,

$$\Omega_L = \bigcup_{k=0}^L \left\{ \ell \in \Omega_L : N_\ell \geq k, (\ell_i)_{i=1}^{N_\ell - k} \in \Omega_{L-k}^c, \ell_{N_\ell - k + 1} = \ell_{N_\ell - k + 2} = \dots = \ell_{N_\ell} = 0 \right\}. \quad (2.4.12)$$

By using this decomposition, we obtain that

$$\tilde{Z}_{L,\beta} := Z_{L,\beta} e^{-\beta L} = \sum_{k=0}^L e^{-\beta k} \tilde{Z}_{L-k,\beta}^c \quad (2.4.13)$$

and then, using (2.4.11) and dominated convergence we conclude that

$$K_\beta := \lim_{L \rightarrow \infty} \frac{1}{h(L)} \tilde{Z}_{L,\beta} = \frac{1}{1 - e^{-\beta}} \left[ \frac{\bar{K}_\beta}{1 - \delta_2(\beta)} + \frac{\widehat{K}_\beta \delta_1(\beta)}{(1 - \delta_2(\beta))^2} \right], \quad (2.4.14)$$

which completes the proof of Theorem 2.2.1, by combining (2.4.14), (2.4.4) and (2.3.17), and by writing

$$K_\beta = K_\beta^\circ e^{-\beta} \left[ (1 + e^{-\beta}) e^{\operatorname{arccosh}(e^{-\beta/2} \cosh(\beta))} - e^{\beta/2} (1 - e^{-\beta}) \right]^{-2}, \quad (2.4.15)$$

where  $K_\beta^\circ$  is computed below in (2.4.36).  $\square$

*Proof of Proposition 2.4.1.* Let us now prove Proposition 2.4.1 subject to Proposition 2.4.5 and Lemma 2.4.7 below. The proof of Proposition 2.4.5 is postponed to section 2.5, whereas Lemma 2.4.7 was already stated and proven in [52, Lemma 4.4] so we will not repeat the proof in the present paper.

For  $\beta > 0$  and  $q > 0$  recall the definitions of  $\tilde{\psi}(q)$  and  $C_{\beta,q}$  from (2.2.7) and (2.2.8).

**Proposition 2.4.5.** *Let  $[q_1, q_2] \subset (0, \infty)$  and  $N \in \mathbb{N}$ . We have that for  $N \in \mathbb{N}$  such that  $qN^2 \in \mathbb{N}$ ,*

$$\mathbf{P}_\beta(\mathcal{V}_{N, qN^2}) = \frac{C_{\beta,q}}{N^2} e^{-N\tilde{\psi}(q)}(1 + o(1)), \quad (2.4.16)$$

uniformly in  $q \in [q_1, q_2]$ , with

$$\mathcal{V}_{n,k} := \{X : X_n = 0, A_n = k, X_i > 0, 0 < i < n\}, \quad (n, k) \in (\mathbb{N}_0)^2. \quad (2.4.17)$$

**Remark 2.4.6.** For the proof of Proposition 2.4.5 (see Section 2.5), we took inspiration from [132] where a slightly different problem is considered. To be more specific, the authors consider a random walk  $(Y_i)_{i \geq 0}$  with a negative drift and a light tail such that the moment generating function  $\varphi(t) := E(e^{tY_1})$  satisfies that there exists a  $\lambda > 0$  such that  $\varphi(\lambda) = 1$ ,  $\varphi'(\lambda) < \infty$  and  $\varphi''(\lambda) < \infty$ . For such a walk, they provide the asymptotics of the joint law of  $\tau := \inf\{i \geq 1 : Y_i \leq 0\}$  and of  $A_{\tau-1}$  as the latter becomes large. When applied in our framework, [132, Theorem 1 and 2] prove that for  $p > 0$  and  $q > p/2$  there exist  $C_1, C_2 > 0$  such that

$$\mathbf{P}_\beta(X_N \leq pN, A_N = qN^2, X_i > pi, 0 < i < N) \underset{N \rightarrow \infty}{\sim} \frac{C_1}{N^2} e^{-C_2 N}.$$

The case  $p = 0$ , which is the object of Proposition 2.4.5 is not covered by [132] though, which is why we provide a complete proof in Section 2.5.

**Lemma 2.4.7.** *Let  $\beta > \beta_c$ , there exists  $(a_1, a_2) \in (0, \infty)^2$  such that*

$$\lim_{L \rightarrow \infty} \frac{Z_{L,\beta}^\circ([a_1, a_2]\sqrt{L})}{Z_{L,\beta,\delta}^{\circ,+}} = 1. \quad (2.4.18)$$

We start the proof of Proposition 2.4.1 by recalling (2.3.16) and the equality  $\frac{c_\beta}{\Gamma_\beta} = e^\beta$ , which allow us to write for  $L \in \mathbb{N}$ ,

$$Z_{L,\beta}^\circ = 2c_\beta e^{\beta L} \tilde{Z}_{L,\beta}^\circ := 2e^{\beta(1+L)} \sum_{N=2}^{1+L/2} \Gamma(\beta)^N \mathbf{P}_\beta(\mathcal{V}_{N, L-N+1}). \quad (2.4.19)$$

Lemma 2.4.7 guarantees us that it suffices to consider

$$\begin{aligned} Q_{L,\beta} &:= \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} \Gamma(\beta)^N \mathbf{P}_\beta(\mathcal{V}_{N, L-N+1}) \\ &= \sum_{x \in [a_1, a_2] \cap \frac{\mathbb{N}}{\sqrt{L}}} \Gamma(\beta)^{x\sqrt{L}} \mathbf{P}_\beta[\mathcal{V}_{x\sqrt{L}, q_L(x)(x\sqrt{L})^2}], \end{aligned} \quad (2.4.20)$$

with  $q_L(x) := \frac{L-x\sqrt{L+1}}{x^2L}$ . We note that  $x \in [a_1, a_2]$  yields that for  $L$  large enough,  $q_L(x) \in [\frac{1}{2a_2^2}, \frac{2}{a_1^2}]$  so that we can apply Proposition 2.4.5 for every  $x \in [a_1, a_2] \cap \frac{\mathbb{N}}{\sqrt{L}}$  in the r.h.s. of (2.4.20) and obtain

$$Q_{L,\beta} \underset{L \rightarrow \infty}{\sim} \sum_{x \in [a_1, a_2] \cap \frac{\mathbb{N}}{\sqrt{L}}} \frac{C_{\beta, q_L(x)}}{x^2 L} e^{x\sqrt{L}[\log \Gamma(\beta) - \tilde{\psi}(q_L(x))]} \quad (2.4.21)$$

By definition (see (2.2.7))  $\tilde{\psi}$  is  $\mathcal{C}^2$  on  $(0, \infty)$  and therefore, uniformly in  $x \in [a_1, a_2]$  we get

$$\tilde{\psi}(q_L(x)) = \tilde{\psi}(x^{-2}) - \tilde{\psi}'(x^{-2}) \frac{1}{x\sqrt{L}} + O\left(\frac{1}{L}\right) \quad (2.4.22)$$

such that (2.4.21) becomes

$$Q_{L,\beta} \underset{L \rightarrow \infty}{\sim} \frac{1}{L} \sum_{x \in [a_1, a_2] \cap \frac{\mathbb{N}}{\sqrt{L}}} \frac{C_{\beta, q_L(x)} e^{\tilde{\psi}'(x^{-2})}}{x^2} e^{\sqrt{L} \tilde{\mathcal{G}}(x)}, \quad (2.4.23)$$

with  $x \in (0, \infty) \mapsto \tilde{\mathcal{G}}(x) := x \log \Gamma(\beta) - x \tilde{\psi}(x^{-2})$  a function already used in [52, (1.27)], which is  $\mathcal{C}^2$ , negative, has negative second derivative (and therefore is strictly concave on  $(0, \infty)$ ), and reaches its unique maximum at  $a_\beta \in [a_1, a_2]$ .

At this stage we pick  $R > 0$  and we set  $\mathcal{T}_{R,L} := [a_\beta - \frac{R}{L^{1/4}}, a_\beta + \frac{R}{L^{1/4}}] \cap \frac{\mathbb{N}}{\sqrt{L}}$ , and

$$\tilde{Q}_{L,\beta}^{R,+} := \sum_{x \in \mathcal{T}_{R,L}} \frac{C_{\beta, q_L(x)} e^{\tilde{\psi}'(x^{-2})}}{x^2} e^{\sqrt{L} \tilde{\mathcal{G}}(x)}, \quad (2.4.24)$$

$$\tilde{Q}_{L,\beta}^{R,-} := \sum_{x \in [a_1, a_2] \setminus \mathcal{T}_{R,L} \cap \frac{\mathbb{N}}{\sqrt{L}}} \frac{C_{\beta, q_L(x)} e^{\tilde{\psi}'(x^{-2})}}{x^2} e^{\sqrt{L} \tilde{\mathcal{G}}(x)}.$$

We recall (2.2.8) and we note that for  $\beta > 0$ , the function  $q \in (0, \infty) \mapsto C_{\beta, q}$  is continuous since  $x \in (0, \beta/2) \mapsto \kappa(x)$  is continuous (see Lemma 2.5.13) as well as  $q \mapsto \tilde{h}^q$  (see (2.5.6) and (2.5.7)). Moreover  $q_L(x)$  converges to  $\frac{1}{a_\beta^2}$  uniformly in  $x \in \mathcal{T}_{R,L}$  and therefore

$$\lim_{L \rightarrow \infty} C_{\beta, q_L(x)} = C_{\beta, a_\beta^{-2}} \quad \text{uniformly in } x \in \mathcal{T}_{R,L} \quad (2.4.25)$$

so that we can rewrite

$$\tilde{Q}_{L,\beta}^{R,+} \underset{L \rightarrow \infty}{\sim} \frac{C_{\beta, a_\beta^{-2}} e^{\tilde{\psi}'(a_\beta^{-2})}}{a_\beta^2} \sum_{n=-RL^{1/4}}^{RL^{1/4}} e^{\sqrt{L} \tilde{\mathcal{G}}\left(a_\beta + \frac{n}{\sqrt{L}}\right)} \quad (2.4.26)$$

where we have changed the summation indices for computational convenience. Note that, in the argument of  $\tilde{\mathcal{G}}$  in (2.4.26), we should have considered  $\lfloor a_\beta \sqrt{L} \rfloor / \sqrt{L}$  rather than  $a_\beta$  which may not belong to  $\mathbb{N}/\sqrt{L}$ . However,  $\tilde{\mathcal{G}}$  being  $\mathcal{C}^1$  on  $[a_1, a_2]$  the equivalence in (2.4.26)

remains true. Since  $a_\beta$  is a maximum of  $\tilde{\mathcal{G}}$  that is  $\mathcal{C}^2$  we can write the following expansion of  $\tilde{\mathcal{G}}$ ,

$$\tilde{\mathcal{G}}\left(a_\beta + \frac{n}{\sqrt{L}}\right) = \tilde{\mathcal{G}}(a_\beta) + \tilde{\mathcal{G}}''(a_\beta) \frac{n^2}{2L} + O\left(\frac{n^3}{L^{3/2}}\right) \quad (2.4.27)$$

and use it in the last term in (2.4.26) to write

$$\begin{aligned} \sum_{n=-RL^{1/4}}^{RL^{1/4}} e^{\sqrt{L}\tilde{\mathcal{G}}(a_\beta + \frac{n}{\sqrt{L}})} &\underset{L \rightarrow \infty}{\sim} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \sum_{n=-RL^{1/4}}^{RL^{1/4}} e^{\frac{1}{2}\tilde{\mathcal{G}}''(a_\beta)\left(\frac{n}{L^{1/4}}\right)^2} \\ &\underset{L \rightarrow \infty}{\sim} L^{\frac{1}{4}} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \int_{-R}^R e^{\tilde{\mathcal{G}}''(a_\beta)\frac{x^2}{2}} dx \end{aligned} \quad (2.4.28)$$

where we have used a Riemann sum to obtain the last equality in (2.4.28) since  $\tilde{\mathcal{G}}''(a_\beta)$  is negative. It remains to combine (2.4.26) with (2.4.28) to obtain

$$Q_{L,\beta}^{R,+} \underset{L \rightarrow \infty}{\sim} \frac{C_{\beta, a_\beta^{-2}} e^{\tilde{\psi}'(a_\beta^{-2})}}{a_\beta^2} L^{\frac{1}{4}} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \int_{-R}^R e^{\tilde{\mathcal{G}}''(a_\beta)\frac{x^2}{2}} dx. \quad (2.4.29)$$

Let us now consider  $Q_{L,\beta}^{R,-}$ . To that aim, we bound it from above as

$$\tilde{Q}_{L,\beta}^{R,-} \leq M_{\beta,L} \sum_{x \in [a_1, a_2] \setminus \mathcal{T}_{R,L} \cap \frac{\mathbb{N}}{\sqrt{L}}} e^{\sqrt{L}\tilde{\mathcal{G}}(x)}. \quad (2.4.30)$$

with

$$M_{\beta,L} := \max_{x \in [a_1, a_2]} \frac{C_{\beta, q_L(x)} e^{\tilde{\psi}'(\frac{1}{x^2})}}{x^2}. \quad (2.4.31)$$

The continuity on  $(0, \infty)$  of both  $q \mapsto C_{\beta, q}$  and  $\tilde{\psi}'$  and the fact that, for  $L$  large enough,  $q_L(x) \in [1/(2a_2^2), 2/a_1^2]$  for  $x \in [a_1, a_2]$ , guarantees us that there exists a  $M > 0$  such that  $M_{\beta,L} \leq M$  for every  $L \geq 1$

Recalling that  $\tilde{\mathcal{G}}$  is  $\mathcal{C}^2$ , strictly concave and reaches its maximum at  $a_\beta$  there exists  $c > 0$  such that  $\tilde{\mathcal{G}}(x) \leq \tilde{\mathcal{G}}(a_\beta) - c(x - a_\beta)^2$  for  $x \in [a_1, a_2]$  and therefore the sum in the r.h.s. in (2.4.30) can be bounded above as

$$\sum_{x \in [a_1, a_2] \setminus \mathcal{T}_{R,L} \cap \frac{\mathbb{N}}{\sqrt{L}}} e^{\sqrt{L}\tilde{\mathcal{G}}(x)} \leq 2L^{1/4} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \frac{1}{L^{1/4}} \sum_{n=RL^{1/4}}^{\infty} e^{-c\left(\frac{n}{L^{1/4}}\right)^2} \quad (2.4.32)$$

The function  $x \mapsto e^{-cx^2}$  being non increasing on  $[0, \infty)$ , a standard comparison between sum and integral yields that for  $L$  large enough and every  $R \geq 2$ ,

$$\frac{1}{L^{1/4}} \sum_{n=RL^{1/4}}^{\infty} e^{-c\left(\frac{n}{L^{1/4}}\right)^2} \leq \int_{R-1}^{\infty} e^{-cx^2} dx. \quad (2.4.33)$$

It remains to use (2.4.30–2.4.33) to conclude that for  $L$  large enough and  $R \geq 2$ ,

$$\tilde{Q}_{L,\beta}^{R,-} \leq 2 M L^{1/4} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}} \int_{R-1}^{\infty} e^{-cx^2} dx. \quad (2.4.34)$$

Let us recall that  $\int_{\mathbb{R}} e^{\tilde{\mathcal{G}}''(a_\beta)\frac{x^2}{2}} dx = \sqrt{-2\pi/\tilde{\mathcal{G}}''(a_\beta)}$ . Therefore, we combine (2.4.23) with (2.4.24), (2.4.29) and (2.4.34) to claim that

$$Q_{L,\beta} \underset{L \rightarrow \infty}{\sim} \frac{\sqrt{2\pi} C_{\beta, a_\beta^{-2}} e^{\tilde{\psi}'(a_\beta^{-2})}}{a_\beta^2 |\tilde{\mathcal{G}}''(a_\beta)|^{1/2}} \frac{1}{L^{3/4}} e^{\tilde{\mathcal{G}}(a_\beta)\sqrt{L}}, \quad (2.4.35)$$

and it suffices to recall (2.4.19) to complete the proof of Proposition 2.4.1 with

$$K_\beta^\circ = 2 e^\beta \frac{\sqrt{2\pi} C_{\beta, a_\beta^{-2}} e^{\tilde{\psi}'(a_\beta^{-2})}}{a_\beta^2 |\tilde{\mathcal{G}}''(a_\beta)|^{1/2}}. \quad (2.4.36)$$

□

## 2.5 Proof of Proposition 2.4.5

Proposition 2.4.5 is a substantial improvement of [52, Prop. 2.5], since this latter proposition only allowed us to bound from below the quantity  $\mathbf{P}(\mathcal{V}_{n,qn^2})$  with a polynomial factor  $1/n^\gamma$  ( $\gamma > 2$ ) instead of  $1/n^2$  in the present Lemma. Let us first recall some results on the large deviation principle satisfied by the sequence of random vectors  $(\frac{1}{n}A_{n-1}, X_n)_{n \geq 1}$ . We then provide an outline of the proof of Proposition 2.4.5, which is divided into 4 steps, corresponding to Sections 2.5.1, 2.5.2, 2.5.3 and 2.5.4 respectively.

### Change of measure

Let  $X := (X_i)_{i \in \mathbb{N}}$  be a random walk starting from the origin and whose increments are iid with law  $\mathbf{P}_\beta$  (recall (2.2.1)). For  $|h| < \beta/2$ , recall the definitions of  $\mathcal{L}(h)$  and  $\tilde{\mathbf{P}}_h$  in (2.2.2), and (2.2.3). Recall also the definition of  $A_N(X)$  in (2.2.6).

Large deviations estimates for the random vector  $\Lambda_n := (\frac{1}{n}A_{n-1}(X), X_n)$  are displayed in [65]. Typically, one is interested in the probability of events as

$$\{\frac{1}{n}\Lambda_n = (q, p)\} \quad \text{with} \quad (q, p) \in \mathbb{R}^* \times \mathbb{R},$$

which requires to introduce tilted probability laws of the form

$$\frac{d\mathbf{P}_{n,h}}{d\mathbf{P}_{\beta,0}}(X) = e^{\mathbf{h} \cdot \Lambda_n - \mathcal{L}_{\Lambda_n}(\mathbf{h})} \quad \text{with} \quad \mathcal{L}_{\Lambda_n}(\mathbf{h}) := \log \mathbf{E}_{\beta,0}[e^{\mathbf{h} \cdot \Lambda_n}]. \quad (2.5.1)$$

where  $\mathbf{h} := (h_0, h_1) \in \mathcal{D}_{\beta,n}$  with

$$\mathcal{D}_{\beta,n} := \left\{ (h_0, h_1) \in \mathbb{R}^2; |h_1| < \beta/2, |(1 - \frac{1}{n})h_0 + h_1| < \beta/2 \right\}.$$

For  $(q, p) \in \mathbb{R}^* \times \mathbb{R}$ , the fact that  $\nabla[\frac{1}{n}\mathcal{L}_{\Lambda_n}]$  is a  $\mathcal{C}^1$  diffeomorphism from  $\mathcal{D}_{\beta,n}$  to  $\mathbb{R}^2$  (see [52, Lemma 5.4]), allows us to choose  $\mathbf{h} := \mathbf{h}_n(q, p)$  in (2.5.1) with  $\mathbf{h}_n(q, p)$  the unique solution (in  $\mathbf{h}$ ) of the equation

$$\mathbf{E}_{n,\mathbf{h}} \left[ \frac{1}{n} \Lambda_n \right] = \nabla \left[ \frac{1}{n} \mathcal{L}_{\Lambda_n} \right] (\mathbf{h}) = (q, p). \quad (2.5.2)$$

Under  $\mathbf{P}_{n,\mathbf{h}_n(q,p)}$  the event  $\{\frac{1}{n}\Lambda_n = (q, p)\}$  becomes typical and can be sharply estimated.

In the present context though we only consider events where the random walk returns to the origin after  $n$  steps, i.e.,  $\{\frac{1}{n}\Lambda_n = (q, 0)\}$ . Moreover, for straightforward symmetry reasons (stated e.g. in [52, Remark 5.5]) we have

$$\mathbf{h}_n(q, 0) = \left( h_n^q, -h_n^q \left( \frac{1}{2} - \frac{1}{2n} \right) \right), \quad q \in \mathbb{R}$$

where  $h_n^q$  is the unique solution in  $h$  of the equation  $\mathcal{G}'_n(h) = q$ , with

$$\begin{aligned} \mathcal{G}_n(h) &:= \frac{1}{n} \mathcal{L}_{\Lambda_n} \left[ h, -\frac{h}{2} \left( 1 - \frac{1}{n} \right) \right] \quad \text{for } h \in \left( -\frac{n\beta}{n-1}, \frac{n\beta}{n-1} \right) \\ &= \frac{1}{n} \sum_{i=1}^n \mathcal{L} \left[ \frac{h}{2} \left( 1 - \frac{2i-1}{n} \right) \right]. \end{aligned} \quad (2.5.3)$$

It is proven below in Lemma 2.5.2 that the function  $h \mapsto \mathcal{G}'_n(h)$  is a  $\mathcal{C}^1$  diffeomorphism from  $\left( -\frac{n\beta}{n-1}, \frac{n\beta}{n-1} \right)$  to  $\mathbb{R}$ , which justifies the existence and uniqueness of  $h_n^q$ . As a consequence, instead of those tilted probability laws in (2.5.1), we will rather use the probability laws  $\mathbf{P}_{n,h}$  that depend on the sole parameter  $h \in \left( -\frac{n\beta}{n-1}, \frac{n\beta}{n-1} \right)$ , i.e.,

$$\frac{d\mathbf{P}_{n,h}}{d\mathbf{P}_{\beta,0}}(X) = e^{\psi_{n,h}(A_{n-1}, X_n)}, \quad \text{with } \psi_{n,h}(a, x) := h \frac{a}{n} - \frac{h}{2} \left( 1 - \frac{1}{n} \right) x - n \mathcal{G}_n(h), \quad x, a \in \mathbb{Z}. \quad (2.5.4)$$

We note that, under  $\mathbf{P}_{n,h}$ , the increments  $(X_i - X_{i-1})_{i=1}^n$  are independent and  $X_i - X_{i-1}$  follows the law  $\tilde{\mathbf{P}}$  with parameter  $\frac{h}{2} \left( 1 - \frac{2i-1}{n} \right)$  (recall (2.2.3)).

It remains to define the continuous counterpart of  $\mathcal{L}_{\Lambda_n}$ .

$$\mathcal{D}_\beta := \left\{ (h_0, h_1) \in \mathbb{R}^2; |h_1| < \beta/2, |h_0 + h_1| < \beta/2 \right\}, \quad (2.5.5)$$

and for every  $\mathbf{h} = (h_0, h_1) \in \mathcal{D}_\beta$ ,

$$\mathcal{L}_\Lambda(\mathbf{h}) := \int_0^1 \mathcal{L}(h_0 x + h_1) dx. \quad (2.5.6)$$

As stated in [52, Lemma 5.3],  $\nabla \mathcal{L}_\Lambda(\mathbf{h})$  that can be written as

$$\begin{aligned}\nabla \mathcal{L}_\Lambda(\mathbf{h}) &= (\partial_{h_0} \mathcal{L}_\Lambda, \partial_{h_1} \mathcal{L}_\Lambda)(\mathbf{h}) \\ &= \left( \int_0^1 x \mathcal{L}'(xh_0 + h_1) dx, \int_0^1 \mathcal{L}'(xh_0 + h_1) dx \right),\end{aligned}\tag{2.5.7}$$

is a  $\mathcal{C}^1$  diffeomorphism from  $\mathcal{D}_\beta$  to  $\mathbb{R}^2$ . Thus, for  $(q, p) \in \mathbb{R}^2$  we let  $\tilde{\mathbf{h}}(q, p)$  be the unique solution in  $\mathbf{h} \in \mathcal{D}_\beta$  of the equation  $\nabla \mathcal{L}_\Lambda(\mathbf{h}) = (q, p)$ . As mentioned above for the discrete case, since we will only consider the case  $p = 0$ , the fact that  $\mathcal{L}'$  is an odd function combined with (2.5.7) ensures us that

$$\tilde{\mathbf{h}}(q, 0) = \left( \tilde{h}^q, -\frac{\tilde{h}^q}{2} \right), \quad q \in \mathbb{R}\tag{2.5.8}$$

where  $\tilde{h}^q$  is the unique solution in  $h$  of the equation  $\mathcal{G}'(h) = q$ , with

$$\begin{aligned}\mathcal{G}(h) &:= \mathcal{L}_\Lambda\left(h, -\frac{h}{2}\right) \quad \text{for } h \in (-\beta, \beta) \\ &= \int_0^1 \mathcal{L}\left(h\left(\frac{1}{2} - x\right)\right) dx.\end{aligned}\tag{2.5.9}$$

With Lemma 2.5.3 below, we prove that  $h \mapsto \mathcal{G}'(h)$  is a  $\mathcal{C}^1$  diffeomorphism from  $(-\beta, \beta)$  to  $\mathbb{R}$ . This justifies the existence and uniqueness of  $\tilde{h}^q$ .

**Outline of the proof of Proposition 2.4.5** With Step 1 below, we bound from above the difference between the finite size exponential decay rate of  $\mathbf{P}_\beta(\mathcal{V}_{n,qn^2})$  and its limit as  $n$  tends to  $\infty$ . In Step 2, we divide  $\mathbf{P}_\beta(\mathcal{V}_{N,qN^2})$  into a main term  $M_{N,q}$  and an error term  $E_{N,q}$ . The main term is obtained by adding to the definition of  $\mathcal{V}_{N,qN^2}$  some constraints concerning the possible values taken by  $X_{a_N}, A_{a_N}, X_{N-a_N}$  and  $A_{N-a_N-1}$  for an ad-hoc sequence  $(a_N)_{N \geq 1}$  satisfying both  $a_N = o(N)$  and  $\lim_{n \rightarrow \infty} a_N = \infty$ . The  $E_{N,q}$  term is bounded above in Step 3, while in Step 4 we provide a sharp estimate of  $M_{N,q}$ , and we conclude the proof in Step 4 by computing the pre-factors in the estimate of  $\mathbf{P}_\beta(\mathcal{V}_{N,qN^2})$ .

### 2.5.1 Step 1

The aim of this step is to prove the following Proposition, which is a strong improvement of [52, Proposition 2.3] since we bound from above the gap between discrete quantities and their continuous counterparts by  $n^{-2}$  instead of  $n^{-1}$ . As mentioned in Remark 2.4.6 above, our proof is close in spirit to that of [132, Theorem 2], in particular for Proposition 2.5.4 below. Recall the definition of  $\tilde{\psi}$  from (2.2.7).

**Proposition 2.5.1.** *For  $[q_1, q_2] \subset (0, \infty)$ , there exists a  $C > 0$  and an  $n_0 \in \mathbb{N}$  such that for every  $n \geq n_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{n^2}$*

$$\left| \frac{1}{n} \psi_{n, h_n^q}(qn^2, 0) - \tilde{\psi}(q) \right| \leq \frac{C}{n^2},\tag{2.5.10}$$



and

$$|h_n^q - \tilde{h}^q| \leq \frac{C}{n^2}. \quad (2.5.11)$$

*Proof.* We recall (2.5.3) and (2.5.9) where the definitions of  $\mathcal{G}_n$  and  $\mathcal{G}$  are displayed respectively. Let us start with two lemmas stating that  $\mathcal{G}'_n$  and  $\mathcal{G}'$  are  $\mathcal{C}^1$ -diffeomorphisms.

**Lemma 2.5.2.** *For every  $n \geq 2$ , the application  $\mathcal{G}_n$  is  $\mathcal{C}^\infty$ , strictly convex, even and satisfies  $\mathcal{G}'_n(h) \rightarrow +\infty$  as  $h \nearrow \frac{n\beta}{n-1}$ . Moreover, there exists  $R > 0$  such that  $(\mathcal{G}_n)''(h) \geq R$  for every  $n \geq 2$  and  $h \in (-\frac{n\beta}{n-1}, \frac{n\beta}{n-1})$ . As a consequence,  $(\mathcal{G}_n)'$  is an increasing  $\mathcal{C}^1$  diffeomorphism from  $(-\frac{n\beta}{n-1}, \frac{n\beta}{n-1})$  to  $\mathbb{R}$  and  $(\mathcal{G}_n)'(0) = 0$ .*

**Lemma 2.5.3.** *The application  $\mathcal{G}$  is  $\mathcal{C}^\infty$ , strictly convex, even and satisfies  $\mathcal{G}'(h) \rightarrow +\infty$  as  $h \nearrow \beta$ . Moreover, there exists an  $R > 0$  such that  $(\mathcal{G})''(h) \geq R$  for  $h \in (-\beta, \beta)$ . As a consequence,  $(\mathcal{G})'$  is an increasing  $\mathcal{C}^1$  diffeomorphism from  $(-\beta, \beta)$  to  $\mathbb{R}$  and  $(\mathcal{G})'(0) = 0$ .*

*Proof.* Lemma 2.5.2 being a discrete counterpart of Lemma 2.5.3, we will only prove the latter here. One easily observes (recall (2.2.2)) that  $\mathcal{L}$  is  $\mathcal{C}^\infty$ , strictly convex, even and that  $\mathcal{L}''$  is bounded below by a positive constant uniformly on  $(-\beta/2, \beta/2)$ . With the help of (2.5.9) we state that  $\mathcal{G}$  enjoys the very same properties on  $(-\beta, \beta)$  so that it simply remains to prove that  $\lim_{h \rightarrow \beta^-} \mathcal{G}'(h) = \infty$ . To that aim, we compute

$$\begin{aligned} \mathcal{G}'(h) &= 2 \int_0^{1/2} u \mathcal{L}'(hu) du \geq \frac{1}{2} \int_{1/4}^{1/2} \mathcal{L}'(hu) du \\ &\geq \frac{1}{2h} (\mathcal{L}(\frac{h}{2}) - \mathcal{L}(\frac{h}{4})) \end{aligned}$$

and it suffices to observe that  $\lim_{h \rightarrow \beta^-} \mathcal{L}(h/2) = \infty$  to complete the proof of the lemma.  $\square$

Let us resume the proof of Proposition 2.5.1 by recalling that  $h_n^q$  is the unique solution in  $h$  of  $(\mathcal{G}_n)'(h) = q$  and that  $\tilde{h}^q$  is the unique solution in  $h$  of  $\mathcal{G}'(h) = q$ . At this stage, we state a key result which substantially improves [52, Lemma 5.1]. Its proof is postponed to Appendix 2.B.

**Proposition 2.5.4.** *For every  $K \in (0, \beta)$ , there exist a  $C_K > 0$  and a  $n_K \in \mathbb{N}$  such that for  $j \in \{0, 1\}$*

$$\sup_{h \in [-\beta+K, \beta-K]} |\mathcal{G}_n^{(j)}(h) - \mathcal{G}^{(j)}(h)| \leq \frac{C_K}{n^2}, \quad \forall n \geq n_K. \quad (2.5.12)$$

With the help of Lemma 2.5.3 we can state the following corollary of Proposition 2.5.4.

**Corollary 2.5.5.** *For every  $M > 0$ , there exists a  $K > 0$  such that  $\mathcal{G}'(\beta - K) \geq 2M$  and  $\mathcal{G}'(K) \leq \frac{1}{2M}$ ; and there exists an  $n_0 \in \mathbb{N}$  such that  $\mathcal{G}'_n(\beta - K) \geq M$  and  $\mathcal{G}'_n(K) \leq \frac{1}{M}$  for every  $n \geq n_0$ .*

**Remark 2.5.6.** A straightforward consequence of Corollary (2.5.5) is that for  $[q_1, q_2] \subset (0, \infty)$  there exists a  $K > 0$  and an  $n_0 \in \mathbb{N}$  such that for every  $q \in [q_1, q_2]$  and every  $n \geq n_0$  we have  $h_n^q, \tilde{h}^q \in [K, \beta - K]$ .

We now have all the required tools in hand to prove Proposition 2.5.1. By using Lemma 2.5.3, Proposition 2.5.4 and Remark 2.5.6, we can state that there exist  $K > 0$ ,  $C > 0$  and  $R > 0$  such that for  $n$  large enough and  $q \in [q_1, q_2] \cap \mathbb{N}/n^2$  we have  $h_n^q, \tilde{h}^q \in [K, \beta - K]$ , so

$$\begin{aligned} |h_n^q - \tilde{h}^q| &\leq \frac{1}{R} \left| \int_{h_n^q}^{\tilde{h}^q} \mathcal{G}''(x) dx \right| = \frac{1}{R} \left| \mathcal{G}'(h_n^q) - \mathcal{G}'(\tilde{h}^q) \right| \\ &= \frac{1}{R} \left| \mathcal{G}'(h_n^q) - \mathcal{G}'_n(h_n^q) \right| \leq \frac{C}{R n^2} \end{aligned} \quad (2.5.13)$$

where we have used that  $q = \mathcal{G}'(\tilde{h}^q) = \mathcal{G}'_n(h_n^q)$  for the second equality in (2.5.13). The proof of (2.5.11) is therefore completed.

It remains to prove (2.5.10). To that aim we write

$$\left| h_n^q q - \mathcal{G}_n(h_n^q) - \tilde{h}^q q + \mathcal{G}(\tilde{h}^q) \right| \leq U_{n,q} + q_2 |h_n^q - \tilde{h}^q|, \quad (2.5.14)$$

where  $U_{n,q} := |\mathcal{G}_n(h_n^q) - \mathcal{G}(\tilde{h}^q)|$ . Proposition 2.5.4 also tells us that there exists a  $C > 0$  such that for  $n$  large enough and for every  $x \in [K, \beta - K]$  we have  $|\mathcal{G}_n(x) - \mathcal{G}(x)| \leq C/n^2$ . Thus, since  $\mathcal{G}$  is  $\mathcal{C}^1$  and recalling Remark 2.5.6 we can write

$$\begin{aligned} U_{n,q} &\leq \left| \mathcal{G}_n(h_n^q) - \mathcal{G}(h_n^q) \right| + \left| \mathcal{G}(h_n^q) - \mathcal{G}(\tilde{h}^q) \right| \\ &\leq \frac{C}{n^2} + \sup \left\{ |\mathcal{G}'(x)|, x \in [K, \beta - K] \right\} |h_n^q - \tilde{h}^q|. \end{aligned} \quad (2.5.15)$$

At this stage, (2.5.10) is obtained by combining (2.5.14) with (2.5.11) and (2.5.15). This completes the proof of Proposition (2.5.1).  $\square$

## 2.5.2 Step 2

We recall (2.4.17) and (2.5.4) and we set  $a_N := (\log N)^2$ . We define the two boxes

$$\begin{aligned} \mathcal{C}_N &:= \left[ \mathbf{E}_{N, h_N^q}(X_{a_N}) - (a_N)^{3/4}, \mathbf{E}_{N, h_N^q}(X_{a_N}) + (a_N)^{3/4} \right] \\ \mathcal{D}_N &:= \left[ \mathbf{E}_{N, h_N^q}(A_{a_N}) - (a_N)^{7/4}, \mathbf{E}_{N, h_N^q}(A_{a_N}) + (a_N)^{7/4} \right] \end{aligned} \quad (2.5.16)$$

and rewrite  $\mathbf{P}_\beta(\mathcal{V}_{N, qN^2}) = M_{N, q} + E_{N, q}$  where  $M_{N, q}$  is defined as

$$\mathbf{P}_\beta \left[ \mathcal{V}_{N, qN^2} \cap \left\{ X_{a_N} \in \mathcal{C}_N, A_{a_N} \in \mathcal{D}_N \right\} \cap \left\{ X_{N-a_N} \in \mathcal{C}_N, A_N - A_{N-a_N-1} \in \mathcal{D}_N \right\} \right]. \quad (2.5.17)$$

From now on,  $M_{N, q}$  will be referred to as the main term and  $E_{N, q} := \mathbf{P}_\beta(\mathcal{V}_{N, qN^2}) - M_{N, q}$  as the error term. The proof of Proposition 2.4.5 is a straightforward consequence of Lemmas 2.5.7 and 2.5.8 displayed below, which will be proven in Steps 3 and 4, respectively.

Let us start with the following Lemma, which allows us to control the error term uniformly in  $q$  belonging to any compact set of  $(0, \infty)$ .

**Lemma 2.5.7.** For  $[q_1, q_2] \subset (0, \infty)$ , there exists  $\varepsilon : \mathbb{N} \mapsto \mathbb{R}^+$  such that  $\lim_{N \rightarrow \infty} \varepsilon(N) = 0$  and for every  $N \in \mathbb{N}$  and  $q \in [q_1, q_2] \cap \frac{1}{N^2}$

$$E_{N,q} \leq \frac{\varepsilon(N)}{N^2} e^{-N \tilde{\psi}(q)}. \quad (2.5.18)$$

With the next Lemma, we estimate the main term uniformly in  $q$  belonging to any compact set of  $(0, \infty)$ . Recall (2.2.4) and (2.2.9) for the definitions of  $\kappa$  and  $\vartheta$ .

**Lemma 2.5.8.** For  $\beta > 0$  and  $[q_1, q_2] \subset (0, \infty)$ ,

$$M_{N,q} = \kappa \left( \frac{\tilde{h}^q}{2} \right)^2 \frac{(\vartheta(\tilde{h}^q))^{-\frac{1}{2}}}{2\pi N^2} e^{-N \tilde{\psi}(q)} (1 + o(1)) \quad (2.5.19)$$

where  $o(1)$  is a function that converges to 0 as  $N \rightarrow \infty$  uniformly in  $q \in [q_1, q_2] \cap \frac{1}{N^2}$ .

### 2.5.3 Step 3: proof of Lemma 2.5.7

Before starting the proof, let us make a quick remark about the time-reversibility of the random walk  $X$  of law  $\mathbf{P}_{n,h}$ .

**Remark 2.5.9** (Time-reversal property). Recall the definition of  $\tilde{\mathbf{P}}_h$  from (2.2.3). For  $h \in (-\beta/2, \beta/2)$ , one can easily check that if  $Z$  is a random variable of law  $\tilde{\mathbf{P}}_{-h}$  then  $-Z$  has law  $\tilde{\mathbf{P}}_h$ . Moreover, as explained below (2.5.4), if the  $n$ -step random walk  $X := (X_i)_{i=0}^n$  has law  $\mathbf{P}_{n,h}$ , then its increments  $(X_i - X_{i-1})_{i=1}^n$  are independent and for  $i \in \{1, \dots, n\}$  the law of  $X_i - X_{i-1}$  is  $\tilde{\mathbf{P}}$  with parameter  $\frac{h}{2}(1 - \frac{2i-1}{n})$ . A first consequence is that  $X$  is time-reversible, i.e.,

$$(X_i)_{i=0}^n \stackrel{\text{Law}}{=} (X_{n-i} - X_n)_{i=0}^n. \quad (2.5.20)$$

A second consequence is that, under  $\mathbf{P}_{n,h}$ , the random walk  $(X_i)_{i=0}^n$  is an inhomogeneous Markov chain which, for  $j \in \{1, \dots, N-1\}$ ,  $y \in \mathbb{Z}$  and  $\mathcal{O} \subset \mathbb{Z}^{n-j-1}$  satisfies that

$$\mathbf{P}_{n,h} \left( (X_{j+i})_{i=1}^{n-j-1} \in \mathcal{O}, X_n = 0 \mid X_j = y \right) = \mathbf{P}_{n,h} \left( (X_{n-j-i})_{i=1}^{n-j-1} \in \mathcal{O}, X_{n-j} = y \right). \quad (2.5.21)$$

Note finally that the case  $h = 0$  corresponds to the random walk  $X$  with i.i.d. increments of law  $\mathbf{P}_\beta$ .

A straightforward application of (2.5.20) with  $n = N$  and  $h = 0$  allows us to bound the error term from above as

$$E_{N,q} \leq 2 \mathbf{P}_\beta \left( \mathcal{V}_{N,qN^2} \cap \{X_{a_N} \notin \mathcal{C}_N\} \right) + 2 \mathbf{P}_\beta \left( \mathcal{V}_{N,qN^2} \cap \{A_{a_N} \notin \mathcal{D}_N\} \right). \quad (2.5.22)$$

Using (2.5.4) we obtain that for  $\mathcal{B} = \{X_{a_N} \notin \mathcal{C}_N\}$  or  $\mathcal{B} = \{A_{a_N} \notin \mathcal{D}_N\}$

$$\mathbf{P}_\beta \left( \mathcal{V}_{N,qN^2} \cap \mathcal{B} \right) \leq e^{-\psi_{N,h_N^q}(qN^2,0)} \mathbf{P}_{N,h_N^q} \left( \mathcal{V}_{N,qN^2} \cap \mathcal{B} \right). \quad (2.5.23)$$

Note that, the first inequality in Proposition 2.5.1 allows us to replace  $\psi_{N,h_N^q}(qN^2, 0)$  with  $N\tilde{\psi}(q)$  in the exponential of the r.h.s. in (2.5.23), at the cost of an at most constant factor. Therefore, the proof of Lemma 2.5.7 is completed by the following Claim.

**Claim 2.5.10.** *For  $[q_1, q_2] \subset (0, \infty)$ , there exists  $\varepsilon : \mathbb{N} \mapsto \mathbb{R}^+$  such that  $\lim_{N \rightarrow \infty} \varepsilon(N) = 0$  and for every  $N \in \mathbb{N}$  and  $q \in [q_1, q_2] \cap \frac{1}{N^2}$*

$$\mathbf{P}_{N,h_N^q}(\mathcal{V}_{N,qN^2} \cap \{X_{a_N} \notin \mathcal{C}_N\}) + \mathbf{P}_{N,h_N^q}(\mathcal{V}_{N,qN^2} \cap \{A_{a_N} \notin \mathcal{D}_N\}) \leq \frac{\varepsilon(N)}{N^2}. \quad (2.5.24)$$

*Proof.* Let us prove that (2.5.24) holds true for  $R_{N,q} := \mathbf{P}_{N,h_N^q}(\mathcal{V}_{N,qN^2} \cap \{X_{a_N} \notin \mathcal{C}_N\})$  and for  $S_{N,q} := \mathbf{P}_{N,h_N^q}(\mathcal{V}_{N,qN^2} \cap \{A_{a_N} \notin \mathcal{D}_N\})$ .

We set  $\{X_{[j,k]} > 0\} := \{X_i > 0, j \leq i \leq k\}$  for  $j \leq k \in \mathbb{N}$ . We develop  $R_{N,q}$  depending on the values  $y$  and  $z$  taken by  $X_{a_N}$  and  $A_{a_N}$  respectively. Then we use Markov property at time  $a_N$ , combined with the time reversal property (2.5.21) with  $n = N$ ,  $h = h_N^q$ ,  $j = a_N$  and

$$\mathcal{O} = \left\{ x \in \mathbb{N}^{N-j-1} : \sum_{i=1}^{N-j-1} x_i = qN^2 - z \right\},$$

on the time interval  $[a_N, N]$  to obtain

$$\begin{aligned} R_{N,q} &= \sum_{y \in \mathbb{N} \setminus \mathcal{C}_N} \sum_{z \in \mathbb{N}} \mathbf{P}_{N,h_N^q}(X_{a_N} = y, A_{a_N} = z, X_{[1,a_N]} > 0) \\ &\quad \times \mathbf{P}_{N,h_N^q}(X_{[1,N-a_N]} > 0, X_{N-a_N} = y, A_{N-a_N-1} = qN^2 - z). \end{aligned} \quad (2.5.25)$$

Let  $R_{N,q}^1 := R_{N,q}(|y| \leq N/a_N, |z| \leq N^2/a_N)$ ,  $R_{N,q}^2 := R_{N,q}(|y| > N/a_N)$  and  $R_{N,q}^3 := R_{N,q}(|z| > N^2/a_N)$ , where  $R_{N,q}(A)$  denotes the sum from (2.5.25) restricted to terms satisfying the condition  $A$ ; so that

$$R_{N,q} \leq R_{N,q}^1 + R_{N,q}^2 + R_{N,q}^3. \quad (2.5.26)$$

Let us prove the upper bound (2.5.24) for all three terms in the r.h.s., starting with  $R_{N,q}^1$ .

**Lemma 2.5.11.** *For  $[q_1, q_2] \subset (0, \infty)$ , there exists a  $C > 0$  such that for every  $N \geq 1$  and  $q \in [q_1, q_2] \cap \frac{1}{N^2}$  and  $(y, z) \in \mathbb{Z}^2$  with  $|y| \leq N/a_N$ ,  $|z| \leq N^2/a_N$ ,*

$$\mathbf{P}_{N,h_N^q}(X_{N-a_N} = y, A_{N-a_N-1} = qN^2 - z) \leq \frac{C}{N^2}. \quad (2.5.27)$$

We prove this Lemma afterwards. Plugging this in the development (2.5.25) and dropping the condition  $X_{[1,a_N]} > 0$ , we obtain

$$R_{N,q}^1 \leq \frac{C}{N^2} \mathbf{P}_{N,h_N^q}(X_{a_N} \notin \mathcal{C}_N). \quad (2.5.28)$$

Recalling (2.5.4), a straightforward computation gives us that

$$\text{Var}_{N,h_N^q}(X_{a_N}) = \sum_{i=1}^{a_N} \mathcal{L}'' \left( \frac{h_N^q}{2} \left( 1 - \frac{2i-1}{N} \right) \right), \quad (2.5.29)$$

so that, by using Tchebychev inequality we obtain

$$\begin{aligned}
 \mathbf{P}_{N,h_N^q}(X_{a_N} \notin \mathcal{C}_N) &= \mathbf{P}_{N,h_N^q}(|X_{a_N} - \mathbf{E}_{N,h_N^q}(X_{a_N})| > (a_N)^{3/4}) \\
 &\leq \frac{1}{a_N^{3/2}} \mathbf{Var}_{N,h_N^q}(X_{a_N}) \\
 &\leq \frac{1}{a_N^{1/2}} \sup_{x \in [0,1]} \left| \mathcal{L}''\left(h_N^q\left(\frac{1}{2} - x\right)\right) \right| \\
 &\leq \frac{C_1}{\sqrt{a_N}}
 \end{aligned} \tag{2.5.30}$$

for some uniform  $C_1 > 0$ , where the last inequality is a consequence of Remark 2.5.6 and the observation that  $\mathcal{L}$  is  $\mathcal{C}^\infty$  on  $(-\beta/2, \beta/2)$ . It remains to combine (2.5.28) and (2.5.30) to complete the proof for  $R_{N,q}^1$ .

Regarding  $R_{N,q}^2$  let  $(U_i)_{i \geq 1}$  be the increments of the process  $X$ ; recalling (2.5.4) and applying a Chernov inequality for some  $\lambda > 0$  small, we have

$$\begin{aligned}
 &\sum_{y > N/a_N} \mathbf{P}_{N,h_N^q}(X_{a_N} = y, A_{a_N} = z, X_{[1,a_N]} > 0) \\
 &\leq \mathbf{P}_{N,h_N^q}(X_{a_N} > N/a_N) \\
 &\leq \mathbf{E}_{\beta,0} \left[ \exp \left( \lambda \left( X_{a_N} - \frac{N}{a_N} \right) + \sum_{i=1}^{a_N} U_i \left( \frac{h_N^q}{2} \left( 1 - \frac{2i-1}{N} \right) \right) - \mathcal{L} \left( \frac{h_N^q}{2} \left( 1 - \frac{2i-1}{N} \right) \right) \right) \right] \\
 &\leq e^{-\lambda N/a_N} e^{c a_N} \leq C e^{-\frac{\lambda N}{2a_N}},
 \end{aligned} \tag{2.5.31}$$

for some  $c > 0, C > 0$ , which can be taken uniformly in  $q \in [q_1, q_2]$ . The same upper bound holds for the sum over  $y < N/a_N$ , hence

$$\begin{aligned}
 R_{N,q}^2 &\leq \sum_{z \in \mathbb{Z}} \mathbf{P}_{N,h_N^q}(A_{N-a_N-1} = qN^2 - z) \\
 &\quad \times \sum_{|y| > N/a_N} \mathbf{P}_{N,h_N^q}(X_{a_N} = y, A_{a_N} = z, X_{[1,a_N]} > 0) \\
 &\leq 2C e^{-\frac{\lambda N}{2a_N}}
 \end{aligned} \tag{2.5.32}$$

The last term  $R_{N,q}^3$  can be handled similarly, by noticing that  $A_{a_N} > N^2/a_N$  implies  $X_i > N^2/a_N^2$  for some  $1 \leq i \leq a_N$ , thereby

$$\begin{aligned}
 &\sum_{z > N^2/a_N} \mathbf{P}_{N,h_N^q}(X_{a_N} = y, A_{a_N} = z, X_{[1,a_N]} > 0) \\
 &\leq \mathbf{P}(A_{a_N} > N^2/a_N) \leq \sum_{i=1}^{a_N} \mathbf{P}(X_i > N^2/a_N^2) \leq C a_N e^{-\frac{\lambda N^2}{2a_N^2}},
 \end{aligned}$$

where we reproduced (2.5.31). Similarly to (2.5.32), we obtain an upper bound on  $R_{N,q}^3$ , and recollecting (2.5.26) with (2.5.28), (2.5.30) and (2.5.32), this finally proves (2.5.24) for  $R_{N,q}$ .

Let us now consider  $S_{N,q}$ , which we develop similarly to (2.5.25). Notice that the term  $S_{N,q}(y \notin \mathcal{C}_N)$  is already bounded from above by  $R_{N,q}$ , and that  $S_{N,q}(|z| > N^2/a_N)$  follows the same upper bound as  $R_{N,q}^3$  (we do not replicate the proof), thus we only have to prove (2.5.24) for  $S_{N,q}^1 := S_{N,q}(y \in \mathcal{C}_N, |z| \leq N^2/a_N)$  to complete the proof of Claim 2.5.10. Recalling Lemma 2.5.11, we have

$$S_{N,q}^1 \leq \frac{C}{N^2} \mathbf{P}_{N,h_N^q}(A_{a_N} \notin \mathcal{D}_N). \quad (2.5.33)$$

Similarly to (2.5.29), a direct computation gives

$$\mathrm{Var}_{N,h_N^q}(A_{a_N}) = \sum_{i=1}^{a_N} (a_N + 1 - i)^2 \mathcal{L}''\left(\frac{h_N^q}{2}\left(1 - \frac{2i-1}{N}\right)\right),$$

and using Tchebychev inequality, we obtain

$$\begin{aligned} \mathbf{P}_{N,h_N^q}(A_{a_N} \notin \mathcal{D}_N) &\leq \frac{1}{a_N^{7/2}} \mathrm{Var}_{N,h_N^q}(A_{a_N}) \\ &\leq \frac{1}{a_N^{1/2}} \sup_{x \in [0,1]} \left| \mathcal{L}''\left(h_N^q\left(\frac{1}{2} - x\right)\right) \right| \\ &\leq \frac{C_1}{\sqrt{a_N}} \end{aligned}$$

for some uniform  $C_1 > 0$ , where the last inequality is a consequence of Remark 2.5.6 and the observation that  $\mathcal{L}$  is  $C^\infty$  on  $(-\beta/2, \beta/2)$ . Combining this with (2.5.33), this concludes the proof of the Claim.  $\square$

*Proof of Lemma 2.5.11.* To lighten upcoming formulae in this proof, we will write  $N_1 := N - a_N$ . Recall (2.5.4), and that  $\mathbf{P}_{\beta,y}$  denotes the law of the random walk starting from  $y \in \mathbb{Z}$  with increments distributed as  $\mathbf{P}_\beta$ . Summing over possible values for  $(X_N, A_{N-1})$  and using Markov property, we write

$$\begin{aligned} &\mathbf{P}_{N,h_N^q}(X_{N_1} = y, A_{N_1-1} = qN^2 - z) \\ &= \sum_{(u,v) \in \mathbb{Z}^2} \mathbf{P}_{N,h_N^q}(X_{N_1} = y, A_{N_1-1} = qN^2 - z, X_N = u, A_{N-1} = v) \\ &= \sum_{(u,v) \in \mathbb{Z}^2} e^{\psi_{N,h_N^q}(v,u)} \mathbf{P}_{\beta,0}(X_{N_1} = y, A_{N_1-1} = qN^2 - z, X_N = u, A_{N-1} = v) \\ &= \sum_{(u,v) \in \mathbb{Z}^2} e^{\psi_{N,h_N^q}(v,u)} \mathbf{P}_{\beta,0}(X_{N_1} = y, A_{N_1-1} = qN^2 - z) \\ &\quad \times \mathbf{P}_{\beta,y}(X_{a_N} = u, A_{a_N-1} = v - qN^2 + z - y) \\ &= \mathbf{P}_{N_1,h_{N_1}^q}(X_{N_1} = y, A_{N_1-1} = qN^2 - z,) \\ &\quad \times \sum_{(u,v) \in \mathbb{Z}^2} \exp\left(\psi_{N,h_N^q}(v,u) - \psi_{N_1,h_{N_1}^q}(qN^2 - z, y)\right) \\ &\quad \times \mathbf{P}_{\beta,y}(X_{a_N} = u, A_{a_N-1} = v - qN^2 + z - y). \end{aligned} \quad (2.5.34)$$

A straightforward consequence of Proposition 2.A.1 and Remark 2.A.3 from Appendix 2.A is that the first factor is uniformly controlled by  $\frac{C}{N^2}$  for some uniform  $C > 0$ , so it remains to prove that the second factor is uniformly bounded. It can be written as

$$\mathbf{E}_{\beta,0} \left[ \exp \left( \psi_{N,h_N^q} \left( A_{a_N-1} + qN^2 - z + y a_N, X_{a_N} + y \right) - \psi_{N_1,h_{N_1}^q} \left( qN^2 - z, y \right) \right) \right],$$

and can be estimated with Proposition 2.5.1. Furthermore, we claim that for all  $q \in \mathbb{R}$ ,

$$\mathcal{L}(\tilde{h}^q/2) - q\tilde{h}^q - \mathcal{G}(\tilde{h}^q) = 0. \quad (2.5.35)$$

Indeed,

$$\begin{aligned} \mathcal{L}(\tilde{h}^q/2) - \mathcal{G}(\tilde{h}^q) &= - \int_0^1 \mathcal{L}(\tilde{h}^q(\frac{1}{2} - y)) dy + \mathcal{L}(\tilde{h}^q/2) \\ &= \int_0^1 \int_y^1 \tilde{h}^q \mathcal{L}'(\tilde{h}^q(u - \frac{1}{2})) du dy \\ &= \tilde{h}^q \int_0^1 u \mathcal{L}'(\tilde{h}^q(u - \frac{1}{2})) du \\ &= \tilde{h}^q \mathcal{G}'(\tilde{h}^q) = \tilde{h}^q q, \end{aligned}$$

where we have used that  $\mathcal{L}$  is even to obtain the second line and that  $\tilde{h}^q$  is solution in  $h$  of  $\mathcal{G}'(h) = q$  (recall Step 1) to obtain the last line.

We deduce from (2.5.34), Proposition 2.5.1 and (2.5.35) (we do not write the details here) that there exists  $C_1 > 0$  such that for  $n \in \mathbb{N}$ ,  $h \in ]-\beta, \beta[$  and  $|y| \leq N/a_N$ ,  $|z| \leq N^2/a_N$ ,

$$\begin{aligned} &\mathbf{P}_{N,h_N^q}(X_{N_1} = y, A_{N_1-1} = qN^2 - z) \\ &\leq \frac{C_1}{N^2} e^{-a_N \mathcal{L}(\tilde{h}^q/2)} \mathbf{E}_{\beta,0} \left[ \exp \left( 2\tilde{h}^q \frac{A_{a_N-1}}{N} - \frac{\tilde{h}^q}{2} X_{a_N} \left( 1 - \frac{2}{N} \right) \right) \right]. \end{aligned} \quad (2.5.36)$$

Let us now focus on that last factor, which we part into two terms. On the one hand, notice that

$$\begin{aligned} &\mathbf{E}_{\beta,0} \left[ \exp \left( 2\tilde{h}^q \frac{A_{a_N-1}}{N} - \frac{\tilde{h}^q}{2} X_{a_N} \left( 1 - \frac{2}{N} \right) \right) \mathbf{1}_{\{A_{a_N-1} \leq N\}} \right] \\ &\leq e^{2\tilde{h}^q} \mathbf{E}_{\beta,0} \left[ \exp \left( -\frac{\tilde{h}^q}{2} X_{a_N} \left( 1 - \frac{2}{N} \right) \right) \right] \\ &\leq e^{2\tilde{h}^q} \exp \left( a_N \mathcal{L} \left( \frac{\tilde{h}^q}{2} \left( 1 - \frac{2}{N} \right) \right) \right) \leq e^{2\tilde{h}^q + a_N \mathcal{L}(\tilde{h}^q/2)}, \end{aligned} \quad (2.5.37)$$

where we used that  $\mathcal{L}$  is symmetric, and non-decreasing on  $[0, +\infty)$ . On the other hand, the event  $A_{a_N-1} > N$  implies that  $U_i > N/a_N^2$  for some  $1 \leq i \leq a_N - 1$ , where  $(U_i)_{i \geq 1}$

denotes the increments of the process  $X$ . So

$$\begin{aligned}
 & \mathbf{E}_{\beta,0} \left[ \exp \left( 2\tilde{h}^q \frac{A_{a_N-1}}{N} - \frac{\tilde{h}^q}{2} X_{a_N} \left( 1 - \frac{2}{N} \right) \right) \mathbf{1}_{\{A_{a_N-1} > N\}} \right] \\
 & \leq \sum_{i=1}^{a_N-1} \mathbf{E}_{\beta,0} \left[ \mathbf{1}_{\{U_i > N/a_N^2\}} \exp \left( \sum_{j=1}^{a_N-1} U_j \tilde{h}^q \left( \frac{2(a_N-i)+1}{N} - \frac{1}{2} \right) \right) \right] \\
 & \leq \sum_{i=1}^{a_N-1} \left[ \prod_{j \neq i} e^{\mathcal{L} \left( \tilde{h}^q \left( \frac{2(a_N-i)+1}{N} - \frac{1}{2} \right) \right)} \times \sum_{k > N/a_N^2} \frac{e^{-\frac{\beta}{2}k}}{C_\beta} e^{\tilde{h}^q \left( \frac{2(a_N-k)+1}{N} - \frac{1}{2} \right)} \right] \\
 & \leq C (a_N - 1) e^{C_1(a_N-2)} e^{-\frac{\beta}{2} \frac{N}{a_N^2}} \leq C_2 e^{-\frac{\beta}{4} \frac{N}{a_N^2}}, \tag{2.5.38}
 \end{aligned}$$

for some  $C, C_1, C_2 > 0$  uniform in  $q \in [q_1, q_2]$ . Plugging (2.5.37) and (2.5.38) in (2.5.36), this concludes the proof of Lemma 2.5.11.  $\square$

## 2.5.4 Step 4: proof of lemma 2.5.8

In the following, we set  $\bar{x} = (x_1, x_2)$  and  $\bar{a} = (a_1, a_2)$ . We define also

$$\mathcal{H}_N := \{(\bar{x}, \bar{a}) \in \overline{\mathcal{C}}_N^2 \times \overline{\mathcal{D}}_N^2\}. \tag{2.5.39}$$

We use Markov property at time  $a_N$  and  $N - a_N$  and we apply time-reversibility on the part of the walk between time  $N - a_N$  and  $N$  (i.e., (2.5.20) with  $n = a_N$  and  $h = 0$ ) to obtain

$$M_{N,q} = \sum_{(\bar{x}, \bar{a}) \in \mathcal{H}_N} R_N(x_1, a_1) T_N(\bar{x}, \bar{a}) R_N(x_2, a_2), \tag{2.5.40}$$

with

$$R_N(x, a) := \mathbf{P}_\beta(X_{[1, a_N]} > 0, X_{a_N} = x, A_{a_N} = a) \tag{2.5.41}$$

and, after setting  $N_2 = N - 2a_N$ ,

$$T_N(\bar{x}, \bar{a}) := \mathbf{P}_\beta \left( X_{N_2} = x_2 - x_1, A_{N_2-1} = qN^2 - a_1 - a_2 - x_1(N_2 - 1), X_{[0, N_2]} > -x_1 \right) \tag{2.5.42}$$

We begin with  $R_N(x_1, a_1)$  and  $R_N(x_2, a_2)$ . For conciseness, we set  $h := \tilde{h}^q/2$ . Then, we recall (2.2.3) and we perform a change of measure by tilting every increment  $(X_{i+1} - X_i)_{i=0}^{a_N-1}$  with  $\tilde{\mathbf{P}}_h$ , that is for  $(x, a) \in \{(x_1, a_1), (x_2, a_2)\}$

$$R_N(x, a) := e^{-hx + a_N \mathcal{L}(h)} \tilde{\mathbf{P}}_h(X_i > 0, \forall i \leq a_N, X_{a_N} = x, A_{a_N} = a). \tag{2.5.43}$$

For the second term  $T_N(\bar{x}, \bar{a})$  we apply the tilting introduced in (2.5.4) and we obtain

$$T_N(\bar{x}, \bar{a}) = G_{N, \bar{x}, \bar{a}}^q e^{-h_{N_2}^q \left( \frac{qN^2 - a_1 - a_2}{N_2} - x_1 \left( 1 - \frac{1}{N_2} \right) \right)} e^{+\frac{h_{N_2}^q}{2} \left( 1 - \frac{1}{N_2} \right) (x_2 - x_1) + N_2 \mathcal{G}_{N_2}(h_{N_2}^q)} \tag{2.5.44}$$



with

$$G_{N,\bar{x},\bar{a}}^q := \mathbf{P}_{N_2, h_{N_2}^q} \left[ \Lambda_{N_2} = \left( \frac{qN^2 - a_1 - a_2}{N_2} - x_1 \left( 1 - \frac{1}{N_2} \right), x_2 - x_1 \right), X_{[0, N_2]} > -x_1 \right]. \quad (2.5.45)$$

Recollecting (2.5.40), (2.5.43), (2.5.44) and (2.5.45), we obtain

$$M_{N,q} = \sum_{(\bar{x}, \bar{a}) \in \mathcal{H}_N} e^{B_{N,q}^1(\bar{a}) + B_{N,q}^2(\bar{x})} \tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} = x_1, A_{a_N} = a_1) G_{N,\bar{x},\bar{a}}^q \quad (2.5.46)$$

$$\tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} = x_2, A_{a_N} = a_2)$$

with

$$B_{N,q}^1(\bar{a}) := N_2 \mathcal{G}_{N_2}(h_{N_2}^q) + 2a_N \mathcal{L}(h) - h_{N_2}^q \frac{qN^2 - a_1 - a_2}{N_2}, \quad (2.5.47)$$

$$B_{N,q}^2(\bar{x}) := (x_1 + x_2) \left[ \frac{h_{N_2}^q}{2} \left( 1 - \frac{1}{N_2} \right) - h \right].$$

Henceforth, we drop the  $\bar{a}$  and  $\bar{x}$  dependency of  $B_{N,q}^1(\bar{a})$  and  $B_{N,q}^2(\bar{x})$  for conciseness. We consider first  $B_{N,q}^2$ . We recall that  $h = \tilde{h}^q/2$ , thus

$$B_{N,q}^2 = \frac{x_1 + x_2}{2} \left( h_{N_2}^q - \tilde{h}^q - \frac{h_{N_2}^q}{N_2} \right), \quad (2.5.48)$$

which allows us to write the upper bound

$$|B_{N,q}^2| \leq \frac{|x_1| + |x_2|}{2} \left( |h_{N_2}^q - \tilde{h}^q| + \frac{h_{N_2}^q}{N_2} \right). \quad (2.5.49)$$

We recall (2.2.3) and observe that  $x \mapsto \tilde{\mathbf{E}}_x(X_1)$  is non decreasing. Thus, it suffices to apply Remark 2.5.6 to conclude that there exist a  $K > 0$  and a  $N_0 \in \mathbb{N}$  such that, for  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  the inequality  $\mathbf{E}_{N, h_{N_2}^q}(X_{a_N}) \leq a_N \tilde{\mathbf{E}}_{(\beta-K)/2}(X_1)$  holds true. Therefore, by recalling the definition of  $\mathcal{C}_N$  in (2.5.16) we can assert that every  $(x_1, x_2) \in (\mathcal{C}_N)^2$  satisfies  $|x_1| + |x_2| \leq (\text{const.})a_N$ , where (const.) is some positive (uniform) constant. It remains to use (2.5.11) to conclude that provided  $N_0 \in \mathbb{N}$  is large enough, for every  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  and  $N \geq N_0$

$$|B_{N,q}^2| \leq (\text{const.}) a_N \left( \frac{C}{(N_2)^2} + \frac{\beta}{N_2} \right)$$

$$\leq (\text{const.}) \frac{a_N}{N_2}, \quad (2.5.50)$$

and therefore,  $B_{N,q}^2$  converges to 0 as  $N \rightarrow \infty$  uniformly in  $q \in [q_1, q_2]$ .

Now, we consider  $B_{N,q}^1$ . We recall the definition of  $\mathcal{D}_N$  in (2.5.16) and similarly to what we did above, we can assert that provided  $N_0$  is chosen large enough, for  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  we have that every  $(a_1, a_2) \in (\mathcal{D}_N)^2$  satisfies  $|a_1| + |a_2| \leq (\text{const.})a_N^2$  for some constant. Moreover,

$$\frac{qN^2 - a_1 - a_2}{N_2} = qN + 2qa_N + O\left(\frac{(a_N)^2}{N}\right) \quad (2.5.51)$$

and therefore  $B_{N,q}^1 = \widehat{B}_N(q) + O\left(\frac{(a_N)^2}{N}\right)$  with

$$\widehat{B}_N(q) := N_2 \mathcal{G}_{N_2}(h_{N_2}^q) + 2a_N \mathcal{L}(\tilde{h}^q/2) - h_{N_2}^q q(N + 2a_N). \quad (2.5.52)$$

Our aim is to show that

$$\widehat{B}_N(q) = N\mathcal{G}(\tilde{h}^q) - N\tilde{h}^q + o(1) \quad (2.5.53)$$

and for that we rewrite

$$\begin{aligned} \widehat{B}_N(q) &= N_2 \left[ \mathcal{G}_{N_2}(h_{N_2}^q) - h_{N_2}^q q \right] + 2a_N \left[ \mathcal{L}(\tilde{h}^q/2) - 2h_{N_2}^q q \right] \\ &= N_2 \left[ \mathcal{G}(\tilde{h}^q) - \tilde{h}^q q + O\left(\frac{1}{(N_2)^2}\right) \right] + 2a_N \left[ \mathcal{L}(\tilde{h}^q/2) - 2\tilde{h}^q q + O\left(\frac{1}{(N_2)^2}\right) \right] \\ &= N \left[ \mathcal{G}(\tilde{h}^q) - \tilde{h}^q q \right] + 2a_N \left[ \mathcal{L}(\tilde{h}^q/2) - \tilde{h}^q q - \mathcal{G}(\tilde{h}^q) \right] + O\left(\frac{1}{N_2}\right) \end{aligned}$$

which, by (2.5.10) and (2.5.11), is true uniformly in  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$ , provided  $N_0$  is chosen large enough. Recalling (2.5.35), this completes the proof of (2.5.53).

We recall that to lighten notations we sometimes use  $h = \tilde{h}^q/2$ . At this stage, we deduce from (2.5.46), (2.5.50) and (2.5.53) that

$$M_{N,q} = (1 + o(1)) e^{N(\mathcal{G}(\tilde{h}^q) - \tilde{h}^q q)} \sum_{(\bar{x}, \bar{a}) \in \mathcal{H}_N} \tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} = x_1, A_{a_N} = a_1) \quad (2.5.54)$$

$$G_{N, \bar{x}, \bar{a}}^q \tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} = x_2, A_{a_N} = a_2).$$

We will complete the present step subject to Lemma 2.5.12 below. This Lemma will be proven in Appendix 2.A.

**Lemma 2.5.12.** For  $\beta > 0$ ,

$$G_{N, \bar{x}, \bar{a}}^q = \frac{(\vartheta(\tilde{h}^q))^{-\frac{1}{2}}}{2\pi N^2} (1 + o(1)) \quad (2.5.55)$$

where  $o(1)$  is a function that converges to 0 as  $N \rightarrow \infty$  uniformly in  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  and  $(\bar{x}, \bar{a}) \in \mathcal{H}_N$ .

By applying Lemma (2.5.12), we may rewrite (2.5.54) as

$$M_{N,q} = (1 + o(1)) \frac{(\vartheta(\tilde{h}^q))^{-\frac{1}{2}}}{2\pi N^2} e^{N(\mathcal{G}(\tilde{h}^q) - \tilde{h}^q q)} \widehat{M}_{N,q} \quad (2.5.56)$$

with

$$\begin{aligned} \widehat{M}_{N,q} &:= \sum_{(\bar{x}, \bar{a}) \in \mathcal{H}_N} \tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} = x_1, A_{a_N} = a_1) \\ &\quad \times \mathbf{P}_h(X_{[1, a_N]} > 0, X_{a_N} = x_2, A_{a_N} = a_2) \end{aligned} \quad (2.5.57)$$

$$= \left[ \tilde{\mathbf{P}}_h(X_{[1, a_N]} > 0, X_{a_N} \in \mathcal{C}_N, A_{a_N} \in \mathcal{D}_N) \right]^2, \quad (2.5.58)$$

from which we deduce that

$$\left| \left( \widehat{M}_{N,q} \right)^{1/2} - \widetilde{\mathbf{P}}_h(X_{[1,a_N]} > 0) \right| \leq \widetilde{\mathbf{P}}_h(X_{a_N} \notin \mathcal{C}_N) + \widetilde{\mathbf{P}}_h(A_{a_N} \notin \mathcal{D}_N). \quad (2.5.59)$$

Remark 2.5.6 guarantees us again that there exist a  $K > 0$  and a  $N_0 \in \mathbb{N}$  such that  $\widetilde{h}^q \in [K, \beta - K]$  for  $q \in [q_1, q_2]$  and that  $\widetilde{h}_N^q \in [K, \beta - K]$  for  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{N}{N^2}$ . Since  $\{X_{a_N} \notin \mathcal{C}_N\} = \{|X_{a_N} - \mathbf{E}_{N,h_N^q}(X_{a_N})| \geq (a_N)^{3/4}\}$  we set  $\alpha_q := \widetilde{\mathbf{E}}_h(X_1)$  so that  $\widetilde{\mathbf{E}}_h(X_{a_N}) = a_N \alpha_q$  (with  $h = \widetilde{h}^q/2$ ) and we compute

$$\begin{aligned} |a_N \alpha_q - \mathbf{E}_{N,h_N^q}(X_{a_N})| &= \left| \sum_{i=1}^{a_N} \widetilde{\mathbf{E}}_h(X_1) - \widetilde{\mathbf{E}}_{h_N^q(1-\frac{i}{N})-\frac{h_N^q}{2}(1-\frac{i}{N})}(X_1) \right| \\ &\leq \sum_{i=1}^{a_N} \left| \mathcal{L}'\left(\frac{\widetilde{h}^q}{2}\right) - \mathcal{L}'\left(\frac{h_N^q}{2} + \frac{h_N^q}{2}\left(\frac{1-2i}{N}\right)\right) \right| \\ &\leq \sum_{i=1}^{a_N} \max_{x \in \left[\frac{K}{2}, \frac{\beta-K}{2}\right]} |\mathcal{L}''(x)| \left( \left| \frac{\widetilde{h}^q - h_N^q}{2} \right| + \beta \frac{a_N}{N} \right) \\ &\leq (\text{const.}) \left( \frac{a_N}{N^2} + \frac{(a_N)^2}{N} \right). \end{aligned} \quad (2.5.60)$$

As a consequence  $\{X_{a_N} \notin \mathcal{C}_N\} \subset \{|X_{a_N} - a_N \alpha_q| \geq \frac{1}{2}(a_N)^{3/4}\}$  and by using Tchebychev inequality we conclude that

$$\begin{aligned} \widetilde{\mathbf{P}}_h(|X_{a_N} - a_N \alpha_q| \geq \frac{1}{2}(a_N)^{3/4}) &\leq (\text{const.}) \frac{1}{\sqrt{a_N}} \text{Var}_{\widetilde{h}^q/2}(X_1) \\ &\leq (\text{const.}) \frac{1}{\sqrt{a_N}}, \end{aligned} \quad (2.5.61)$$

where we have used that  $\mathcal{L}$  is  $\mathcal{C}^2$  and that  $\mathcal{L}''(x) = \text{Var}_x(X_1)$  for every  $x \in (-\beta/2, \beta/2)$ . This finally proves that  $\lim_{N \rightarrow \infty} \widetilde{\mathbf{P}}_h(X_{a_N} \notin \mathcal{C}_N) = 0$  and the same type of argument also gives that  $\lim_{N \rightarrow \infty} \widetilde{\mathbf{P}}_h(A_{a_N} \notin \mathcal{D}_N) = 0$ . Both convergences hold true uniformly in  $q \in [q_1, q_2]$ .

We come back to (2.5.59) and we can write that

$$\widehat{M}_{N,q} = (\widetilde{\mathbf{P}}_h(X_{[1,a_N]} > 0))^2 (1 + o(1)) \quad (2.5.62)$$

where  $o(1)$  is uniform in  $q \in [q_1, q_2] \cap \frac{N}{N^2}$ . We recall that  $\kappa(h) = \widetilde{\mathbf{P}}_h(X_{[1,\infty]} > 0)$  (see (2.2.4)). At this stage, it remains to prove the convergence of  $\widetilde{\mathbf{P}}_h(X_{[1,a_N]} > 0)$  towards  $\kappa(h)$  uniformly in  $h$  in any compact subset of  $(0, \beta/2)$ . This is the object of the following lemma.

**Lemma 2.5.13.** *For  $\beta > 0$ , the function  $x \in (0, \beta/2) \mapsto \kappa(x)$  is continuous and for  $[x_1, x_2] \subset (0, \beta/2)$*

$$\widetilde{\mathbf{P}}_x(V_{[1,k]} > 0) = \kappa(x) + o(1), \quad (2.5.63)$$

where  $o(1)$  is a function of  $x$  and  $k$  that converges to 0 as  $k \rightarrow \infty$  uniformly in  $x \in [x_1, x_2]$ . Moreover, for  $\beta > 0$  and  $x \in [0, \beta/2)$ , one has

$$\kappa(x) = \frac{e^{2x} - 1}{e^{x+\beta/2} - 1}. \quad (2.5.64)$$

*Proof.* We pick  $x \in [x_1, x_2]$ ,  $k \geq 1$  and we set  $\varepsilon := x_1/2$ , then

$$0 \leq \tilde{\mathbf{P}}_x(X_{[1,k]} > 0) - \kappa(x) \leq \sum_{j=k+1}^{\infty} \tilde{\mathbf{P}}_x(V_j \leq 0) \leq \sum_{j=k+1}^{\infty} e^{-(\mathcal{L}(x) - \mathcal{L}(x-\varepsilon))j} \quad (2.5.65)$$

where we have used a Markov exponential inequality to obtain the last inequality. The fact that  $\mathcal{L}$  is convex and non decreasing on  $[0, \beta/2)$  allows us to claim that for every  $x \in [x_1, x_2]$  we have  $\mathcal{L}(x) - \mathcal{L}(x - \varepsilon) \geq \tilde{r} := \mathcal{L}'(x_1 - \varepsilon)\varepsilon = \mathcal{L}'(\frac{x_1}{2})\frac{x_1}{2} > 0$  and therefore

$$0 \leq \tilde{\mathbf{P}}_x(X_{[1,k]} > 0) - \kappa(x) \leq \sum_{j=k+1}^{\infty} e^{-\tilde{r}j}. \quad (2.5.66)$$

At this stage, (2.5.63) is a straightforward consequence of (2.5.66). The continuity of  $x \mapsto \kappa(x)$  on  $(0, \beta/2)$  is a consequence of the fact that  $x \mapsto \tilde{\mathbf{P}}_x(X_{[1,k]} > 0)$  is continuous on  $(0, \beta/2)$  for every  $k \in \mathbb{N}$  and of (2.5.63) which guarantees us that the latter sequence of functions converges uniformly to  $x \mapsto \kappa(x)$  on every compact subset of  $(0, \beta/2)$ .

Let us now prove (2.5.64). Pick  $x \in [0, \beta/2)$  and define  $\rho = \inf\{i \geq 1, X_i \leq 0\}$ , hence

$$1 - \kappa(x) = \tilde{\mathbf{P}}_x(\rho < \infty) = \mathbf{E}_\beta[e^{xX_\rho - \mathcal{L}(x)\rho} 1_{\{\rho < \infty\}}] = \mathbf{E}_\beta[e^{xX_\rho - \mathcal{L}(x)\rho}],$$

where we used (2.2.3) and that  $\rho < \infty$   $\mathbf{P}_\beta$ -a.s.. As claimed in the proof of Lemma 2.3.2,  $\rho$  and  $X_\rho$  are independent, and  $-X_\rho$  follows a geometric law on  $\mathbb{N} \cup \{0\}$  with parameter  $1 - e^{-\beta/2}$ . Moreover,  $(e^{-xX_n - \mathcal{L}(x)n})_{n \geq 1}$  is a martingale under  $\mathbf{P}_\beta$  (recall that  $\mathcal{L}$  is symmetric), and it is uniformly integrable when stopped at time  $\rho$ ; thus a stopping time argument yields that

$$\mathbf{E}_\beta[e^{-\mathcal{L}(x)\rho}] = \mathbf{E}_\beta[e^{-xX_\rho}]^{-1},$$

(we do not write the details here). Thereby,

$$1 - \kappa(x) = \tilde{\mathbf{P}}_x(\rho < \infty) = \mathbf{E}_\beta[e^{xX_\rho}] \mathbf{E}_\beta[e^{-xX_\rho}]^{-1} = \frac{1 - e^{x-\beta/2}}{1 - e^{-x-\beta/2}},$$

which finally yields (2.5.64). □

We recall that  $h = \tilde{h}^q/2$ . Lemma 2.5.3 guarantees us that for every  $q \in [q_1, q_2]$  we have  $\tilde{h}^q \in [\tilde{h}^{q_1}, \tilde{h}^{q_2}] \subset (0, \beta)$ . Therefore, by using (2.5.56) combined with (2.5.62) and by applying Lemma 2.5.13 with  $[x_1, x_2] = [\frac{\tilde{h}^{q_1}}{2}, \frac{\tilde{h}^{q_2}}{2}]$  we complete the proof of Lemma 2.5.8.

## 2.6 Proof of Theorem 2.2.2

We recall (2.3.8–2.3.11) and Remark 2.3.1. For  $A \subset \Omega_L$  we will denote by  $Z_{L,\beta}(A)$  the partition function restricted to those trajectories in  $A$ , i.e.,

$$Z_{L,\beta}(A) = \sum_{\ell \in A} e^{H_{L,\beta}(\ell)}.$$

For  $L \in \mathbb{N}$ , the function  $k \mapsto \mathbf{P}_{L,\beta}(|I_{\max}(\ell)| \geq L - k)$  is non decreasing. Therefore, proving (2.2.15) with  $\mathbf{P}_{L,\beta}(|I_{\max}(\ell)| \geq L - 3k)$  will be sufficient. Pick  $\ell \in \Omega_L$ , and let  $\mathcal{J}_k := \max\{j \geq 0 : \mathfrak{X}_j \leq k\}$ . Note that, if  $\ell \in \Omega_L$  has no bead starting in  $\{k, \dots, L - k\}$  and if the last bead of  $\ell$  starting before  $k$  begins with at most  $k - 1$  horizontal steps, then the longest sequence of non-zero vertical stretches of alternating signs in  $\ell$  has a total length at least  $L - 3k$ , i.e.,

$$A_{L,k} \cap B_{L,k} \subset \{|I_{\max}(\ell)| \geq L - 3k\}. \quad (2.6.1)$$

with

$$\begin{aligned} A_{L,k} &:= \{\ell \in \Omega_L : \mathfrak{X} \cap \{k, \dots, L - k\} = \emptyset\} \\ B_{L,k} &:= \{\ell \in \Omega_L : \exists j \in \{1, \dots, k\} : \ell_{\tau_{\mathcal{J}_k} + j} \neq 0\}. \end{aligned} \quad (2.6.2)$$

Since  $k \mapsto \mathbf{P}_{L,\beta}(|I_{\max}(\ell)| \geq L - 3k)$  is non decreasing and since (2.6.1) yields

$$\{|I_{\max}(\ell)| \geq L - 3k\}^c \subset (A_{L,k})^c \cup (A_{L,k} \cap (B_{L,k})^c),$$

Theorem 2.2.2 will be proven once we show that for every  $\varepsilon > 0$  there exists a  $k_\varepsilon \in \mathbb{N}$  and a  $L_\varepsilon \in \mathbb{N}$  such that on one hand  $L_\varepsilon \geq 3k_\varepsilon$  and on the other hand, for  $L \geq L_\varepsilon$

$$\mathbf{P}_{L,\beta}((A_{L,k_\varepsilon})^c) \leq \varepsilon \quad \text{and} \quad \mathbf{P}_{L,\beta}(A_{L,k_\varepsilon} \cap (B_{L,k_\varepsilon})^c) \leq \varepsilon. \quad (2.6.3)$$

For simplicity we will write  $\alpha = -\tilde{\mathcal{G}}(a_\beta)$ . A straightforward consequence of Theorem 2.2.1 and Corollary 2.4.2 is that there exists  $0 < C_1 < C_2 < \infty$  such that for every  $L \in \mathbb{N}$ ,

$$\frac{C_1}{L^{3/4}} e^{\beta L - \alpha \sqrt{L}} \leq \hat{Z}_{L,\beta}^\circ \leq Z_{L,\beta} \leq \frac{C_2}{L^{3/4}} e^{\beta L - \alpha \sqrt{L}}. \quad (2.6.4)$$

Thus for  $k, L \in \mathbb{N}$  such that  $L \geq 3k$  we can write

$$\begin{aligned} \mathbf{P}_{L,\beta}((A_{L,k})^c) &\leq \sum_{j=k}^{L-k} \mathbf{P}_{L,\beta}(j \in \mathfrak{X}) = \frac{1}{Z_{L,\beta}} \sum_{j=k}^{L-k} Z_{j,\beta}^c Z_{L-j,\beta}(\ell_1 \geq 0) \\ &\leq \frac{1}{Z_{L,\beta}} \sum_{j=k}^{L-k} Z_{j,\beta} Z_{L-j,\beta} \\ &\leq (\text{const.}) e^{\alpha \sqrt{L}} L^{3/4} \sum_{j=k}^{L-k} \frac{1}{j^{3/4}} e^{-\alpha \sqrt{j}} \frac{1}{(L-j)^{3/4}} e^{-\alpha \sqrt{L-j}} \\ &\leq (\text{const.}) \sum_{j=k}^{L-k} \frac{L^{3/4}}{j^{3/4} (L-j)^{3/4}} e^{-\alpha(\sqrt{j} + \sqrt{L-j} - \sqrt{L})} \\ &\leq (\text{const.}) \sum_{j=k}^{L/2} \frac{1}{j^{3/4}} e^{-\frac{\alpha \sqrt{j}}{2}} \leq (\text{const.}) \sum_{j=k}^{\infty} \frac{1}{j^{3/4}} e^{-\frac{\alpha \sqrt{j}}{2}}, \end{aligned} \quad (2.6.5)$$

where in the second line we have used (2.6.4) and in the last line we have used the convex inequality:  $\sqrt{L} - \sqrt{L-j} \leq \frac{1}{2}\sqrt{j}$  for  $0 \leq j \leq L/2$ . Since the r.h.s. in (2.6.5) does not depend on  $L$  and vanishes as  $k \rightarrow \infty$ , the leftmost inequality in (2.6.3) is proven.

Let us now deal with the second inequality in (2.6.3). For  $L \geq 3k$ , we partition the set  $A_{L,k} \cap (B_{L,k})^c$  by recording  $r$  (respectively  $s$ ), the rightmost (resp. leftmost) point in  $\mathfrak{X}$  that is smaller than  $k$  (resp. larger than  $L-k$ ). Moreover, the fact that  $\ell \in (B_{L,k})^c$  implies that the bead which covers the interval  $\{k, \dots, L-k\}$  begins with  $k$  zero-length vertical stretches. Thus,

$$\begin{aligned} Z_{L,\beta}(A_{L,k} \cap (B_{L,k})^c) &= \sum_{r=0}^{k-1} \sum_{s=L-k+1}^L Z_{L,\beta}(\{\mathfrak{X} \cap \{r, \dots, s\} = \{r, s\}\} \cap (B_{L,k})^c) \\ &= \sum_{r=0}^{k-1} \sum_{s=L-k+1}^L Z_{r,\beta}^c \widehat{Z}_{s-r,\beta}^\circ(\ell_1 = \dots = \ell_k = 0) Z_{L-s,\beta}(\ell_1 \geq 0) \\ &\leq 2 \sum_{r=0}^{k-1} \sum_{s=L-k+1}^L Z_{r,\beta}^c \widehat{Z}_{s-r-k,\beta}^\circ Z_{L-s,\beta}(\ell_1 \geq 0) \\ &= 2 Z_{L-k,\beta}(\{\mathfrak{X} \cap \{r, \dots, s-k\} = \{r, s-k\}\}) \end{aligned} \quad (2.6.6)$$

where the third line in (2.6.6) is obtained by observing that  $\widehat{Z}_{j,\beta}^\circ(\ell_1 = \dots = \ell_k = 0) \leq 2\widehat{Z}_{j-k,\beta}^\circ$  for  $j \geq k+2$ . The factor 2 in the r.h.s. of the latter inequality comes from the fact that there is no constraint on the sign of the  $(k+1)$ -th stretch of  $\ell \in \widehat{\Omega}_j^\circ$  satisfying  $\ell_1 = \dots = \ell_k = 0$ . The r.h.s. in the fourth line of (2.6.6) is obviously bounded above by  $2Z_{L-k,\beta}$  and therefore,

$$\mathbf{P}_{L,\beta}(A_{L,k} \cap (B_{L,k})^c) = \frac{Z_{L,\beta}(A_{L,k} \cap (B_{L,k})^c)}{Z_{L,\beta}} \leq \frac{2 Z_{L-k,\beta}}{Z_{L,\beta}}. \quad (2.6.7)$$

It remains to use (2.6.4) so that (2.6.7) becomes

$$\begin{aligned} \mathbf{P}_{L,\beta}(A_{L,k} \cap (B_{L,k})^c) &\leq (\text{const.}) \frac{L^{3/4}}{(L-k)^{3/4}} e^{-\beta k} e^{-\alpha(\sqrt{L-k}-\sqrt{L})} \\ &\leq (\text{const.}) e^{-\beta k} e^{-\alpha(\sqrt{L-k}-\sqrt{L})}. \end{aligned} \quad (2.6.8)$$

For  $\varepsilon > 0$  we pick  $k_\varepsilon \in \mathbb{N}$  such that  $(\text{const.})^{-\beta k_\varepsilon} \leq \varepsilon/2$ . Moreover, a straightforward computation gives us that  $\lim_{L \rightarrow \infty} \sqrt{L-k_\varepsilon} - \sqrt{L} = 0$ . This completes the proof of the rightmost inequality in (2.6.3) and ends the proof of Theorem 2.2.2.

# Appendices to Chapter 2

## 2.A Local limit estimates, proof of Lemma 2.5.12

We will prove Lemma 2.5.12 subject to Proposition 2.A.1 and Lemma 2.A.2 that are stated below and which were proven in [54, Section 6]. To that aim, we recall (2.5.5–2.5.8) and we set

$$\mathbf{B}(\mathbf{h}) = \text{Hess } \mathcal{L}_\Lambda(\mathbf{h}), \quad \mathbf{h} \in \mathcal{D}, \quad (2.A.1)$$

and

$$f_{\mathbf{h}}(\bar{x}) = \frac{(\text{Det}(\mathbf{B}(\mathbf{h})))^{-1/2}}{2\pi} \exp\left(-\frac{1}{2}\langle \mathbf{B}(\mathbf{h})^{-1}\bar{x}, \bar{x} \rangle\right), \quad \bar{x} \in \mathbb{R}^2. \quad (2.A.2)$$

**Proposition 2.A.1.** [54, Proposition 6.1] For  $[q_1, q_2] \subset (0, \infty)$ , we have

$$\sup_{q \in [q_1, q_2] \cap \frac{\mathbb{N}}{n^2}} \sup_{y, z \in \mathbb{Z}} \left| n^2 \mathbf{P}_{n, h_n^q} \left( A_{n-1} = n^2 q + y, X_n = z \right) - f_{\tilde{\mathbf{h}}(q, 0)} \left( \frac{y}{n^{3/2}}, \frac{z}{\sqrt{n}} \right) \right| \xrightarrow{n \rightarrow \infty} 0.$$

**Lemma 2.A.2.** [54, Lemma 6.2] For  $[q_1, q_2] \subset (0, \infty)$ , there exist three positive constants  $C', C_1, \lambda$  such that for  $N$  large enough, the following bound holds true

$$\mathbf{E}_{N, h_N^q} \left[ e^{-\lambda X_j} \right] \leq C' e^{-C_1 j}, \quad \text{for } j \leq \frac{N}{2} \text{ and } q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}.$$

We recall (2.5.45) and we relax the constraint  $\{X_{[0, N_2]} > -x_1\}$  in  $G_{N, \bar{x}, \bar{a}}^q$  to define  $\bar{G}_{N, \bar{x}, \bar{a}}^q$ , i.e.,

$$\bar{G}_{N, \bar{x}, \bar{a}}^q := \mathbf{P}_{N_2, h_{N_2}^q} \left[ \Lambda_{N_2} = \left( \frac{qN^2 - a_1 - a_2}{N_2} - x_1 \left( 1 - \frac{1}{N_2} \right), x_2 - x_1 \right) \right]. \quad (2.A.3)$$

Remark 2.5.6 guarantees us that there exist a  $K > 0$  and a  $N_0 \in \mathbb{N}$  such that  $h_N^q \in [K, \beta - K]$  for  $N \geq N_0$  and  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$ . Therefore, there exists a  $c_1 > 0$  such that for  $N$  large enough,  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  and  $x_1 \in \mathcal{C}_N$  we have

$$x_1 \geq \mathbf{E}_{N, h_N^q} (X_{a_N}) - (a_N)^{3/4} \geq a_N \tilde{\mathbf{E}}_{\frac{K}{2}} (X_1) - (a_N)^{3/4} \geq c_1 a_N. \quad (2.A.4)$$

As a consequence, we can write

$$\bar{G}_{N,\bar{x},\bar{a}}^q \geq G_{N,\bar{x},\bar{a}}^q \geq \bar{G}_{N,\bar{x},\bar{a}}^q - \hat{G}_{N,\bar{x}}^q \quad (2.A.5)$$

with

$$\hat{G}_{N,\bar{x}}^q = \sum_{i=1}^{N_2-1} \mathbf{P}_{N_2, h_{N_2}^q} \left( X_{N_2} = x_2 - x_1, X_i \leq -c_1 a_N \right). \quad (2.A.6)$$

Until the end of the present section, every function  $o(1)$  converges to 0 as  $N \rightarrow \infty$  uniformly in  $q \in [q_1, q_2] \cap \frac{\mathbb{N}}{N^2}$  and  $(\bar{x}, \bar{a}) \in \mathcal{H}_N$  (we will omit to repeat this uniformity to lighten notations). Thanks to (2.A.5), the proof of Lemma 2.5.12 will be complete once we show that

$$N^2 \bar{G}_{N,\bar{x},\bar{a}}^q = \frac{\left( \vartheta(\tilde{h}^q) \right)^{-\frac{1}{2}}}{2\pi} + o(1), \quad (2.A.7)$$

and that

$$N^2 \hat{G}_{N,\bar{x}}^q = o(1). \quad (2.A.8)$$

Let us start with (2.A.7) by stating a remark about the Hessian matrices  $\mathbf{B}$  (see (2.A.1)).

**Remark 2.A.3.** From (2.2.2) we deduce easily that  $\mathcal{L}''(h) > 0$  for  $h \in (-\beta/2, \beta/2)$ . As a consequence, we can rule out the equality case when applying the Cauchy-Schwartz inequality to (2.2.9). Therefore,

$$\vartheta(\tilde{h}^q) = \text{Det}(\mathbf{B}(\tilde{\mathbf{h}}(q, 0))) > 0, \quad q \in \mathbb{R}. \quad (2.A.9)$$

Since  $\mathcal{L}_\Lambda$  is convex on  $\mathcal{D}$ , (2.A.9) is sufficient to assert that  $\mathbf{B}(\tilde{\mathbf{h}}(q, 0))$  is a symmetric positive-definite matrix for every  $q \in \mathbb{R}$ . Thus, the eigenvalues of  $\mathbf{B}(\tilde{\mathbf{h}}(q, 0))$  are positive and continuous in  $q \in \mathbb{R}$ . This yields that, for  $[q_1, q_2] \subset (0, \infty)$ , the eigenvalues of  $\mathbf{B}(\tilde{\mathbf{h}}(q, 0))$  are bounded above and below by positive constants that are uniform in  $q \in [q_1, q_2]$ . Therefore, there exists a compact subset  $\mathcal{K} \subset (0, \infty)$  such that  $\vartheta(\tilde{h}^q) \in \mathcal{K}$  for every  $q \in [q_1, q_2]$ . From (2.A.2), the latter implies that  $f_{\tilde{\mathbf{h}}(q, 0)}$  is bounded from above on  $\mathbb{R}^2$ , uniformly in  $q \in [q_1, q_2]$ .

We resume the proof of (2.A.7). By definition of  $N_2 = N - 2a_N$  we easily show that

$$N^2 \bar{G}_{N,\bar{x},\bar{a}}^q = (N_2)^2 \bar{G}_{N,\bar{x},\bar{a}}^q (1 + o(1)). \quad (2.A.10)$$

Thus, we apply Proposition 2.A.1 with  $n = N_2$ ,  $y = q(N^2 - (N_2)^2) - a_1 - a_2 - x_1 \frac{N_2-1}{N_2}$  and  $z = x_2 - x_1$ , and we obtain that

$$(N_2)^2 \bar{G}_{N,\bar{x},\bar{a}}^q = f_{\tilde{\mathbf{h}}(q, 0)} \left( \frac{y}{(N_2)^{3/2}}, \frac{z}{\sqrt{N_2}} \right) + o(1). \quad (2.A.11)$$



In Section 2.5.4, to derive (2.5.50) and (2.5.51), we have proven that there exists a  $C > 0$  (uniform in  $q \in [q_1, q_2]$ ) such that  $(\bar{x}, \bar{a}) \in \mathcal{H}_N$  implies  $|x_1| + |x_2| \leq C a_N$  and  $|a_1| + |a_2| \leq C a_N^2$ . Moreover, provided  $C$  is chosen large enough we also have  $N^2 - (N_2)^2 \leq C N a_N$ . Therefore, provided  $C$  is chosen large enough,

$$\frac{|y|}{(N_2)^{3/2}} \leq C \frac{a_N}{\sqrt{N}} \quad \text{and} \quad \frac{|z|}{\sqrt{N_2}} \leq C \frac{a_N}{\sqrt{N}}, \quad \text{for } (\bar{x}, \bar{a}) \in \mathcal{H}_N, \quad q \in [q_1, q_2]. \quad (2.A.12)$$

By remark 2.A.3, we know that the eigenvalues of  $\mathbf{B}(\tilde{\mathbf{h}}(q, 0))^{-1}$  are bounded above uniformly in  $q \in [q_1, q_2]$ , so that (2.A.12) allows us to write

$$\sup_{q \in [q_1, q_2]} \sup_{(\bar{x}, \bar{a}) \in \mathcal{H}_N} \left| f_{\tilde{\mathbf{h}}(q, 0)} \left( \frac{y}{(N_2)^{3/2}}, \frac{z}{\sqrt{N_2}} \right) - f_{\tilde{\mathbf{h}}(q, 0)}(0, 0) \right| \xrightarrow{N \rightarrow \infty} 0. \quad (2.A.13)$$

At this stage, we complete the proof of (2.A.7) by noticing that  $f_{\tilde{\mathbf{h}}(q, 0)}(0, 0) = \vartheta(\tilde{h}^q)^{-1/2} \frac{1}{2\pi}$  and by combining (2.A.10), (2.A.11) and (2.A.13) with the fact that  $f_{\tilde{\mathbf{h}}(q, 0)}$  is bounded from above on  $\mathbb{R}^2$ , uniformly in  $q \in [q_1, q_2]$ .

It remains to prove (2.A.8). We apply (2.5.20) with  $n = N_2$  and  $h = h_{N_2}^q$  to claim that for  $i \in \{1, \dots, N_2 - 1\}$ ,

$$\begin{aligned} \mathbf{P}_{N_2, h_{N_2}^q} (X_{N_2} = x_2 - x_1, X_i \leq -c_1 a_N) & \quad (2.A.14) \\ & = \mathbf{P}_{N_2, h_{N_2}^q} (X_{N_2} = x_1 - x_2, X_{N_2-i} \leq -c_1 a_N + x_1 - x_2). \end{aligned}$$

Since  $|x_1 - x_2| \leq 2(a_N)^{3/4}$  we have that, for  $N$  large enough,  $-c_1 a_N + x_1 - x_2 \leq -c_1 a_N/2$ . Therefore, by using (2.A.14) with  $i \in \{\frac{N_2}{2}, \dots, N_2 - 1\}$  we obtain that

$$\mathbf{P}_{N_2, h_{N_2}^q} (X_{N_2} = x_2 - x_1, X_i \leq -c_1 a_N) \leq \mathbf{P}_{N_2, h_{N_2}^q} (X_{N_2-i} \leq -\frac{c_1}{2} a_N). \quad (2.A.15)$$

Coming back to (2.A.6) and using (2.A.15) for  $i \geq N_2/2$  we can bound  $\widehat{G}_{N, \bar{x}}^q$  from above as

$$\begin{aligned} \widehat{G}_{N, \bar{x}}^q & \leq \sum_{i=1}^{N_2/2} \mathbf{P}_{N_2, h_{N_2}^q} (X_{N_2} = x_2 - x_1, X_i \leq -c_1 a_N) + \sum_{N_2/2}^{N_2-1} \mathbf{P}_{N_2, h_{N_2}^q} (X_{N-i} \leq -\frac{c_1}{2} a_N) \\ & \leq 2 \sum_{i=1}^{N_2/2} \mathbf{P}_{N_2, h_{N_2}^q} (X_i \leq -\frac{c_1}{2} a_N) \leq 2 \sum_{i=1}^{N_2/2} \mathbf{E}_{N_2, h_{N_2}^q} (e^{-\lambda X_i}) e^{-\frac{\lambda c_1 a_N}{2}} \quad (2.A.16) \end{aligned}$$

$$\leq 2 e^{-\frac{\lambda c_1 (\log N)^2}{2}} \sum_{i=1}^{\infty} C' e^{-C_1 j} = o(1) \frac{1}{N^2}, \quad (2.A.17)$$

where we have used a Chernov inequality to obtain the second inequality in the second line, and Lemma 2.A.2 for the last line. This completes the proof of Lemma (2.5.12).

## 2.B Functional estimates, proof of Proposition 2.5.4

Let us start with the case  $j = 0$ . We recall (2.5.3) and we set

$$f_{N,h}(x) := \mathcal{L}\left(\frac{h}{2}\left(1 + \frac{1}{N}\right) - h \frac{x}{N}\right) \quad (2.B.1)$$

and therefore  $N\mathcal{G}_N(h) = \sum_{i=1}^N f_{N,h}(i)$ . We apply the Euler-Maclaurin summation formula (see e.g. [146, Theorem 0.7]) and since  $f_{N,h}$  is  $\mathcal{C}^2$  we obtain that

$$N\mathcal{G}_n(h) = A(N, h) + B(N, h) \quad (2.B.2)$$

with

$$A(N, h) := \frac{f_{N,h}(1) + f_{N,h}(N)}{2} + \int_1^N f_{N,h}(x) dx \quad (2.B.3)$$

and

$$B(N, h) := \frac{1}{2} \int_1^N f_{N,h}''(x) \left( B_2(0) - B_2(x - \lfloor x \rfloor) \right) dx. \quad (2.B.4)$$

where  $B_2$  is seconde Bernoulli polynomial.

We start by considering  $A(N, h)$ . Recalling the definition of  $\mathcal{L}$  in (2.2.2) and the fact that  $\mathbf{P}_\beta$  is symmetric, we claim that  $\mathcal{L}$  is even and therefore

$$\frac{f_{N,h}(1) + f_{N,h}(N)}{2} = \mathcal{L}\left(\frac{h}{2} - \frac{h}{2N}\right). \quad (2.B.5)$$

We recall the definition of  $\mathcal{G}$  in (2.5.9) and a straightforward computation gives

$$\begin{aligned} \int_1^N f_{N,h}(x) dx &= N \int_{1/2N}^{1-1/2N} \mathcal{L}\left(h\left(\frac{1}{2} - y\right)\right) dy \\ &= N\mathcal{G}(h) - 2N \int_0^{1/2N} \mathcal{L}\left(h\left(\frac{1}{2} - z\right)\right) dz \end{aligned} \quad (2.B.6)$$

where we have used the change of variable  $y = x/N$  to get the first equality and the parity of  $\mathcal{L}$  combined with the change of variable  $z = 1 - y$  to obtain the second equality. Obviously, for  $N$  large enough and for every  $h \in [-\beta + K, \beta - K]$  we have that both  $\frac{h}{2}$  and  $\frac{h}{2}\left(1 - \frac{1}{N}\right)$  belong to  $\mathcal{R}_K := \left[-\frac{\beta}{2} + \frac{K}{2}, \frac{\beta}{2} - \frac{K}{2}\right]$ . Since  $\mathcal{L}$  is  $\mathcal{C}^1$  we set  $C'_K := \max\{|\mathcal{L}'(x)|, x \in \mathcal{R}_K\}$  and for  $N$  large enough

$$\left| \mathcal{L}\left(\frac{h}{2} - \frac{h}{2N}\right) - \mathcal{L}\left(\frac{h}{2}\right) \right| \leq C'_K \frac{|h|}{2N} \leq \beta C'_K \frac{1}{2N} \quad (2.B.7)$$

and

$$\left| 2N \int_0^{1/2N} \mathcal{L}\left(h\left(\frac{1}{2} - z\right)\right) dz - \mathcal{L}\left(\frac{h}{2}\right) \right| \leq 2N \int_0^{1/2N} C'_K |h| z dz \leq \beta C'_K \frac{1}{4N}. \quad (2.B.8)$$

Combining (2.B.5–2.B.8) we claim that, for  $N$  large enough

$$|A(N, h) - N\mathcal{G}(h)| \leq \beta C'_K \frac{1}{N}, \quad \forall h \in [-\beta + K, \beta - K]. \quad (2.B.9)$$

It remains to consider  $B(N, h)$  (recall (2.B.4)). To that aim, we compute the second derivative of  $f''_{N,h}$  and obtain

$$f''_{N,h}(x) = \frac{h^2}{N^2} \mathcal{L}'' \left[ \frac{h}{2} \left( 1 + \frac{1}{N} \right) - h \frac{x}{N} \right], \quad x \in [1, N]. \quad (2.B.10)$$

It turns out that for  $x \in [1, N]$  we have  $|\frac{h}{2}(1 + \frac{1}{N}) - h \frac{x}{N}| \leq \frac{|h|}{2}$ . Thus,  $h \in [-\beta + K, \beta - K]$  yields  $\frac{h}{2}(1 + \frac{1}{N}) - h \frac{x}{N} \in \mathcal{R}_K$  and since  $\mathcal{L}$  is  $\mathcal{C}^2$  we set  $C''_K := \max\{|\mathcal{L}''(x)|, x \in \mathcal{R}_K\}$  and we obtain  $|f''_{N,h}(x)| \leq C''_K \beta^2 / N^2$ . As a consequence, we can use (2.B.4) to write

$$|B(N, h)| \leq C''_K \frac{\beta^2}{N} \max\{|B_2(u)|, u \in [0, 1]\}. \quad (2.B.11)$$

At this stage, it remains to combine (2.B.2), (2.B.9) and (2.B.11) to complete the proof of Proposition 2.5.4 in the case  $j = 0$ .

The proof for the case  $j = 1$  being very close in spirit to that for  $j = 0$ , we will give less details in this case. We can repeat (2.B.1–2.B.4) with  $N\mathcal{G}'_N(h)$  instead of  $N\mathcal{G}_N(h)$  and after redefining  $f_{N,h}$  as

$$f_{N,h}(x) := \left( \frac{1}{2} \left( 1 + \frac{1}{N} \right) - \frac{x}{N} \right) \mathcal{L}' \left[ \frac{h}{2} \left( 1 + \frac{1}{N} \right) - h \frac{x}{N} \right]. \quad (2.B.12)$$

Using that  $\mathcal{L}'$  is odd, introducing the function  $g_h(u) := (\frac{1}{2} - u) \mathcal{L}'[h(\frac{1}{2} - u)]$  and computing  $\mathcal{G}'$  from (2.5.9), we obtain that in this case

$$\begin{aligned} A(N, h) &= N \int_{1/2N}^{1-1/2N} \left( \frac{1}{2} - y \right) \mathcal{L}'[h(\frac{1}{2} - y)] dy + \frac{1}{2} \left( 1 - \frac{1}{N} \right) \mathcal{L}'[h(\frac{1}{2} - \frac{1}{2N})] \\ &= N\mathcal{G}'(h) + g(\frac{1}{2N}) - 2N \int_0^{1/2N} g(u) du. \end{aligned} \quad (2.B.13)$$

Taking the derivative of  $g$  we obtain  $g'(u) = -\mathcal{L}'[h(\frac{1}{2} - u)] - h(\frac{1}{2} - u) \mathcal{L}''[h(\frac{1}{2} - u)]$  so that

$$\sup_{u \in [0, 1/2N]} \sup_{h \in [-\beta + K, \beta - K]} |g'(u)| \leq C'_K + \beta C''_K \quad (2.B.14)$$

Therefore,

$$\begin{aligned} \left| g(\frac{1}{2N}) - 2N \int_0^{1/2N} g(u) du \right| &\leq \left| g(\frac{1}{2N}) - g(0) \right| + 2N \int_0^{1/2N} |g(u) - g(0)| du \\ &\leq (C'_K + \beta C''_K) \frac{1}{N}. \end{aligned} \quad (2.B.15)$$

It remains to consider  $B(N, h)$  for which we need to compute  $f''_{N,h}$ , i.e., for  $x \in [1, N]$ ,

$$f''_{N,h}(x) = \frac{2h}{N^2} \mathcal{L}'' \left[ \frac{h}{2} \left( 1 + \frac{1}{N} \right) - h \frac{x}{N} \right] + \frac{h^2}{N^2} \left( \frac{1}{2} \left( 1 + \frac{1}{N} \right) - \frac{x}{N} \right) \mathcal{L}''' \left[ \frac{h}{2} \left( 1 + \frac{1}{N} \right) - h \frac{x}{N} \right]. \quad (2.B.16)$$

Thus, after defining  $C'''_K := \max\{|\mathcal{L}'''(x)|, x \in \mathcal{R}_K\}$  and by mimicking the former proof we obtain  $|f''_{N,h}(x)| \leq (2C''_K \beta + \beta^2 C'''_K) / N^2$  for  $x \in [1, N]$ . This is sufficient to claim (from (2.B.4)) that

$$|B(N, h)| \leq (2C''_K \beta + \beta^2 C'''_K) \frac{1}{N} \max\{|B_2(u)|, u \in [0, 1]\}. \quad (2.B.17)$$

We combine (2.B.13), (2.B.15) and (2.B.17) and it completes the proof of Proposition 2.5.4 in the case  $j = 1$ .



## **The generalized Poland-Scheraga model**



# The Generalized Poland-Scheraga model: (ir)-relevance of disorder

## 3.1 Introduction

The Poland-Scheraga (PS) model has been introduced in [135] to formally study the DNA denaturation phenomenon, that is the unbinding of two strands of DNA as temperature increases. It has proven to be relevant from a quantitative point of view (see e.g. [33, 32]) and has been subject to much interest from the mathematical, physical and biophysical communities (see e.g. [73, 84, 100, 110]). In the *homogeneous* version of the model, *i.e.* when bases in each strand are all the same (for instance AAA. . . and TTT. . .), an interesting feature is that the model is solvable: it is proven to undergo a denaturation (or delocalization) phase transition, and its critical behavior can be described precisely, cf. [84, Ch. 2].

In the PS model, it is assumed that the two strands are of equal length, and that only bases from each strand with the same index can be paired. To depict DNA denaturation more accurately, the generalized Poland-Scheraga (gPS) model has been introduced more recently, where those assumptions are relaxed, see [78, 77, 128]. From the mathematical point of view, the gPS model can be described as a pinning model based on a two-dimensional renewal process, see [87]. Interestingly, the homogeneous version of the model remains solvable, despite having a much more complex behavior—in particular it has (in general) other critical points, corresponding to “condensation” phase transitions, see [87] and [19].

The PS and gPS models can naturally embody the inhomogeneous character of DNA. In the PS model, one introduces a sequence of random variables—referred to as *disorder*

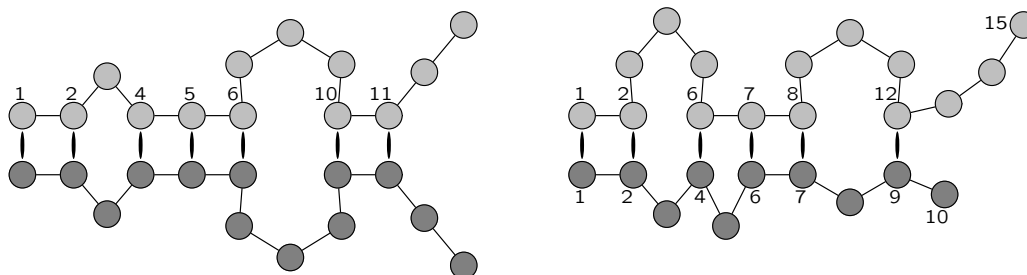
---

This chapter was supervised by Quentin Berger, and has been published in *Electronic Journal of Probability* [119].

in statistical mechanics— describing the inhomogeneous binding energies of successive pairs. A disordered version of the gPS model has been studied recently in [18], with the introduction of a two-dimensional disorder field: the random variable of index  $(i, j) \in \mathbb{N}^2$  corresponds to the binding energy of the  $i$ -th base of the first strand with the  $j$ -th base of the second strand. In [18], the authors chose the disorder field to be i.i.d.: this assumption is relevant when using the gPS model to portray the pinning of a polymer on an inhomogeneous surface, or a directed (stretched) polymer in a random environment (in the spirit of [58, 153]). However this choice is not satisfactory when describing the denaturation phenomenon between two inhomogeneous chains: the binding energy of a pair  $(i, j)$  should be a function of the  $i$ -th and  $j$ -th bases of each strand—in particular the binding energies of two pairs  $(i, j)$  and  $(i, k)$  are not independent because they share a common base.

The purpose of this paper is twofold:

- study the gPS model in a setting which portrays more faithfully the pinning of two inhomogeneous polymers, as in DNA denaturation;
- make progress on the understanding of disordered systems when disorder/random-ness is slightly elaborate, in particular not i.i.d.



(a) Poland-Scheraga.

(b) Generalized Poland-Scheraga.

Figure 3.1: Representations of the PS and gPS models. In the first figure, the two strands are of length 13 and are bounded symmetrically, on pairs 1, 2, 4, 5, 6, 10 and 11. In the other one, the two strands are of different lengths and are bound on pairs  $(1, 1)$ ,  $(2, 2)$ ,  $(6, 4)$ ,  $(7, 6)$ ,  $(8, 7)$  and  $(12, 9)$ .

*Some notation.* In the remainder of the paper, bold characters  $\mathbf{n}, \mathbf{i}, \mathbf{j}, \dots$  will denote elements of  $\mathbb{N}^2$  (or  $\mathbb{Z}^2$ ), and plain characters elements of  $\mathbb{N}$  or  $\mathbb{R}$ . In particular we denote  $\mathbf{1} := (1, 1)$ ,  $\mathbf{0} := (0, 0)$ . For  $r \in \{1, 2\}$ , the projection of any element  $\mathbf{n} \in \mathbb{N}^2$  on its  $r$ -th coordinate will be denoted  $\mathbf{n}^{(r)} \in \mathbb{N}$ .

### 3.1.1 The generalized Poland-Scheraga model

Let  $\tau = (\tau_i)_{i \geq 0}$  be a bivariate renewal process:  $\tau_0 = (0, 0)$ , and  $(\tau_i - \tau_{i-1})_{i \geq 1}$  are i.i.d.,  $\mathbb{N}^2$ -valued random variables. We denote  $\mathbf{P}$  its law, and we assume that the inter-arrival distribution satisfies for all  $a, b \in \mathbb{N}$ ,

$$\mathbf{P}(\tau_1 = (a, b)) = K(a + b) := \frac{L(a + b)}{(a + b)^{2+\alpha}}, \quad (3.1.1)$$

where  $\alpha > 0$ , and  $L(\cdot)$  is a slowly varying function (that is  $L(ux)/L(x) \rightarrow 1$  as  $x \rightarrow \infty$  for any  $u > 0$ , see [30]). We also assume that  $\tau$  is persistent, i.e.  $\sum_{n, m \geq 1} K(n + m) = 1$ . With a slight abuse of notation, we write  $\tau := \{\tau_0, \tau_1, \tau_2, \dots\} \subset \mathbb{N}^2$  the set of renewal points, and from now on we will omit the point  $\tau_0 = (0, 0)$ . Notice that  $\tau^{(r)} := \{\tau_1^{(r)}, \tau_2^{(r)}, \dots\}$  is a univariate renewal process with inter-arrival distribution  $\mathbf{P}(\tau_1^{(r)} = a) = \tilde{L}(a)a^{-(1+\alpha)}$ , with  $\tilde{L}(n) \sim (1 + \alpha)^{-1}L(n)$  some slowly varying function.

Let  $\omega = (\omega_i)_{i \in \mathbb{N}^2}$  be a field of real random variables indexed in  $\mathbb{N}^2$ , whose law is denoted  $\mathbb{P}$  ( $\omega_{(i,j)}$  represents the binding potential between the  $i$ -th and  $j$ -th bases of the first and second strand respectively). We assume that they all have the same law, and that there exists some  $\beta_0 \in (0, \infty]$  such that for all  $\beta \in [0, \beta_0)$ ,

$$\lambda(\beta) := \log \mathbb{E}[e^{\beta \omega_1}] < \infty, \quad (3.1.2)$$

(this is satisfied by bounded laws and by many unbounded laws, notably Gaussian or the product of two independent Gaussian variables).

For a fixed realization of  $\omega$  (*quenched disorder*), we define, for  $\beta \in [0, \beta_0)$  (the disorder strength) and  $h \in \mathbb{R}$  (the pinning potential), the following *polymer (Gibbs) measure*: for any renewal set  $\tau \subset \mathbb{N}^2$  and  $\mathbf{n} = (\mathbf{n}^{(1)}, \mathbf{n}^{(2)}) \in \mathbb{N}^2$ ,

$$\frac{d\mathbf{P}_{\mathbf{n}, h}^{\beta, \omega, \mathbf{q}}}{d\mathbf{P}}(\tau) := \frac{1}{Z_{\mathbf{n}, h}^{\beta, \omega, \mathbf{q}}} \exp \left( \sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{\mathbf{n} \in \tau\}}, \quad (3.1.3)$$

where  $Z_{\mathbf{n}, h}^{\beta, \omega, \mathbf{q}}$  is the *partition function with quenched disorder*,

$$Z_{\mathbf{n}, h}^{\beta, \omega, \mathbf{q}} := \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right], \quad (3.1.4)$$

and  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket$  denotes  $\llbracket 1, \mathbf{n}^{(1)} \rrbracket \times \llbracket 1, \mathbf{n}^{(2)} \rrbracket \subset \mathbb{N}^2$ . This represents the binding of two strands with respective lengths  $\mathbf{n}^{(1)}$  and  $\mathbf{n}^{(2)}$ , and  $\mathbf{i} \in \tau$  if and only if the base  $\mathbf{i}^{(1)}$  of the first strand is paired with the base  $\mathbf{i}^{(2)}$  of the second strand. The polymers are constrained to be bound on the last pair  $\mathbf{n}$ , and we give a reward (or a penalty if negative)  $\beta \omega_i - \lambda(\beta) + h$  for each bound pair  $\mathbf{i} \in \tau$ . Notice that the term  $-\lambda(\beta)$  in the reward is present only for renormalization purposes, see for instance (3.1.11) below. From now on, we will drop the superscript  $\omega$  in the quenched partition function and Gibbs measure to lighten notations, even though they are functions of  $\omega$ .



Let us now precise our choice of disorder field. In [18], the authors studied the gPS model under an i.i.d. disorder field  $\omega = (\omega_i)_{i \in \mathbb{N}^2}$ . In this paper we want the disorder field to depict the inhomogeneous composition of the two strands: we pick two independent sequences  $\hat{\omega} = (\hat{\omega}_{i_1})_{i_1 \in \mathbb{N}}$  and  $\bar{\omega} = (\bar{\omega}_{i_2})_{i_2 \in \mathbb{N}}$  of i.i.d. random variables, whose distributions are denoted  $\hat{\mathbb{P}}$  and  $\bar{\mathbb{P}}$  respectively. These random variables are thought as being charges attached to the two strands. For each  $\mathbf{i} \in \mathbb{N}^2$ , we fix

$$\omega_{\mathbf{i}} := f(\hat{\omega}_{i^{(1)}}, \bar{\omega}_{i^{(2)}}), \quad (3.1.5)$$

where  $f(\cdot, \cdot)$  is a function describing the interactions between the monomers. We will write  $\mathbb{P} := \hat{\mathbb{P}} \otimes \bar{\mathbb{P}}$  with an abuse of notation. We stress right away that  $\omega := (\omega_i)_{i \in \mathbb{N}^2}$  is a strongly correlated field, but that  $\omega_i$  and  $\omega_j$  are independent as soon as  $\mathbf{i}, \mathbf{j} \in \mathbb{N}^2$  are not aligned, *i.e.* are not on the same line or column:  $\mathbf{i}^{(1)} \neq \mathbf{j}^{(1)}$  and  $\mathbf{i}^{(2)} \neq \mathbf{j}^{(2)}$ .

### 3.1.2 The free energy and the denaturation transition

A physical quantity central to the study of the model is the free energy, defined in the following proposition.

**Proposition 3.1.1** (Free energy). *For all  $\gamma > 0$ ,  $h \in \mathbb{R}$ ,  $\beta \geq 0$  and every sequence  $\{m(n)\}_{n=1,2,\dots}$  such that  $\lim_{n \rightarrow \infty} m(n)/n = \gamma$ , the following limit exists:*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{\mathbf{n}, h}^{\beta, \mathbf{q}} = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \log Z_{\mathbf{n}, h}^{\beta, \mathbf{q}} =: F_{\gamma}(\beta, h), \quad (3.1.6)$$

where  $\mathbf{n} := (n, m(n))$ , both  $\mathbb{P}(d\omega)$ -almost surely and in  $L^1(\mathbb{P})$ . Also,  $(\beta, h) \mapsto F_{\gamma}(\beta, h + \lambda(\beta))$  is non-negative and convex (therefore continuous on  $(0, \infty) \times \mathbb{R}$ ),  $h \mapsto F_{\gamma}(\beta, h)$  and  $\beta \mapsto F_{\gamma}(\beta, h + \lambda(\beta))$  are non-decreasing, and  $\gamma \mapsto F_{\gamma}(\beta, h)$  is non-decreasing and continuous. Moreover, we have, for any  $0 < \gamma_1 \leq \gamma_2$ ,

$$F_{\gamma_1}(\beta, h) \leq F_{\gamma_2}(\beta, h) \leq \frac{\gamma_2}{\gamma_1} F_{\gamma_1}(\beta, h). \quad (3.1.7)$$

This proposition is analogous to [18, Thm. 1.1 and Prop 2.1]: the proof of these results is not affected by our choice of (correlated) disorder in any way whatsoever, because any trajectory of  $\tau$  contributing to  $Z_{\mathbf{n}, h}^{\beta, \mathbf{q}}$  only involves an i.i.d. subfamily of the field  $\omega$ : indeed, if  $\mathbf{i}, \mathbf{j} \in \tau$  and  $\mathbf{i} \neq \mathbf{j}$ , then necessarily they are not aligned because the inter-arrivals of  $\tau$  are in  $\mathbb{N}^2$ , hence  $\omega_i$  and  $\omega_j$  are independent. Therefore the proof of Proposition 3.1.1 is an immediate replica of that of [18, Thm. 1.1 and Prop 2.1].

Proposition 3.1.1 allows us to define the (*quenched*) *critical point*:

$$h_c^{\mathbf{q}}(\beta) = h_c(\beta) := \inf\{h : F_{\gamma}(\beta, h) > 0\}. \quad (3.1.8)$$

We stress that  $h_c(\beta)$  does not depend on  $\gamma > 0$ , thanks to (3.1.7).

The critical point  $h_c(\beta)$  marks the transition between a *localized* and a *delocalized* phase: this is the so-called *denaturation* (or *(de)-localization*) transition. Indeed, a standard

calculation gives that  $\partial_h \log Z_{n,h}^{\beta,q} = \mathbf{E}_{n,h}^{\beta,q} \left[ \sum_{i \in \llbracket 1,n \rrbracket} \mathbf{1}_{\{i \in \tau\}} \right]$ : by exploiting the convexity of the free energy and Proposition 3.1.1, we get that

$$\partial_h F_\gamma(\beta, h) = \lim_{\substack{n \rightarrow +\infty, \\ m(n)/n \rightarrow \gamma}} \mathbf{E}_{n,h}^{\beta,q} \left[ \frac{1}{n} \sum_{i \in \llbracket 1,n \rrbracket} \mathbf{1}_{\{i \in \tau\}} \right], \quad (3.1.9)$$

whenever  $\partial_h F_\gamma(\beta, h)$  exists. Therefore, for  $h > h_c(\beta)$  we have  $\partial_h F_\gamma(\beta, h) > 0$ , and in view of (3.1.9), there is a positive density of contacts between the two strands: they stick to each other. On the other hand, for  $h < h_c(\beta)$  we have  $\partial_h F_\gamma(\beta, h) = 0$ , and there is a zero density of contacts: the two strands wander away from one another.

### 3.1.3 The homogeneous and annealed models

The *homogeneous* model corresponds to the case when there is no disorder, *i.e.*  $\beta = 0$ . This model has been proven to be *exactly solvable*, and a fine analysis of  $F_\gamma(0, h)$  has been performed in [87].

**Theorem 3.1.2** (Thm. 1.2 in [87]). *For any  $\gamma \geq 1$ ,  $h_c(0) := \inf\{h : F_\gamma(0, h) > 0\} = 0$ . Moreover there are a slowly varying function  $L_\alpha(\cdot)$  and a constant  $c_{\alpha,\gamma}$  such that*

$$F_\gamma(0, h) \sim c_{\alpha,\gamma} L_\alpha(1/h) h^{1/\min(1,\alpha)}, \quad \text{as } h \searrow 0. \quad (3.1.10)$$

The exponent  $1/\min(1, \alpha)$  is often referred to as the *critical exponent*: it is the main quantification of the behavior of the model around its phase transition. Explicit expressions of  $L_\alpha$  are given in [87], in particular it is some constant if  $\alpha > 1$ .

The *annealed* model, on the other hand, corresponds to averaging the partition function over the disorder: the annealed partition function is, for  $\beta \in [0, \beta_0)$ ,

$$\begin{aligned} Z_{n,h}^{\beta,a} &:= \mathbb{E} \left[ Z_{n,h}^{\beta,q} \right] = \mathbb{E} \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket 1,n \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] \\ &= \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket 1,n \rrbracket} h \mathbf{1}_{\{i \in \tau\}} \right) \mathbf{1}_{\{n \in \tau\}} \right] = Z_{n,h}^{0,q}. \end{aligned} \quad (3.1.11)$$

Here we used that for any fixed trajectory of  $\tau$  the non-zero terms  $(\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}$  are independent and that  $\lambda(\beta) = \log \mathbb{E}[e^{\beta \omega_1}] < +\infty$  for  $\beta \in [0, \beta_0)$  (in particular this implies  $Z_{n,h}^{\beta,q} \in L^1(\mathbb{P})$ ).

Notice that the annealed model matches exactly the homogeneous model. Recalling Proposition 3.1.1, the annealed free energy is therefore

$$F_\gamma^a(\beta, h) := \lim_{\substack{n \rightarrow \infty, \\ m(n)/n \rightarrow \gamma}} \frac{1}{n} \log Z_{n,h}^{\beta,a} = F_\gamma(0, h). \quad (3.1.12)$$

We also directly have that the annealed critical point is  $h_c^a(\beta) := \min\{h : F_\gamma^a(\beta, h) > 0\} = 0$  (recall Theorem 3.1.2).

Now, a simple use of Jensen’s inequality in (3.1.6) gives that  $F_\gamma(\beta, h) \leq F_\gamma^a(\beta, h)$ . Moreover, we have that  $F_\gamma(0, h) \leq F_\gamma(\beta, h + \lambda(\beta))$  (recall that  $\beta \mapsto F_\gamma(\beta, h + \lambda(\beta))$  is non-decreasing). As a conclusion, we obtain the following bounds for the quenched critical point: for every  $\beta$  we have

$$0 = h_c^a(\beta) \leq h_c(\beta) \leq h_c(0) + \lambda(\beta) = \lambda(\beta). \quad (3.1.13)$$

An adaptation of the proof of [84, Th. 5.2] would easily give that the second inequality is strict for every  $\beta > 0$ . The first inequality may or may not be strict and this is an important issue which is directly linked to disorder relevance or irrelevance.

In the rest of the paper, we will work in the case  $\gamma = 1$ : recall that having  $\gamma \neq 1$  changes neither the value of the critical point  $h_c(\beta)$ , nor the homogeneous critical behavior (up to a constant factor, see inequality (3.1.7) and Theorem 3.1.2). To simplify notations, we will drop the dependence on  $\gamma$  in the free energy.

## 3.2 Presentation of the results: the question of disorder relevance

In general, going from a homogeneous model to a disordered one is a complex matter in statistical mechanics (even in the PS model, see [84, Ch. 5]). A first issue is whether the phase transition—in this paper we focus on the denaturation transition—remains when we introduce a small disorder; if so, at what critical value and with what critical behavior compared to the homogeneous model. If any disorder with any strength—parametrized by  $\beta$  in our setting—changes the critical behavior (notably the critical exponent) of the model from the homogeneous case, disorder is said to be *relevant*; if a disorder of small strength does not change the critical behavior, it is said to be *irrelevant*.

The physicist Harris [98] predicts that disorder (ir)-relevance for a  $d$ -dimensional system can be determined from the correlation length exponent  $\nu$  in the homogeneous model. If we admit that the correlation length is given by the reciprocal of the free energy, we obtain from Theorem 3.1.2 that  $\nu = 1/\min(1, \alpha)$ . Then Harris’ criterion predicts that when  $\nu > 2/d$  disorder should be irrelevant, and when  $\nu < 2/d$  it should be relevant (the case  $\nu = 2/d$ , dubbed *marginal*, is much harder to treat, even with heuristic methods).

Notice that in our setting, determining the dimension  $d$  of the system is a more delicate issue than it seems: even though the disorder field  $\omega$  is indexed in  $\mathbb{N}^2$ , it is constructed from two sequences  $\hat{\omega}$  and  $\bar{\omega}$ , therefore has a 1-dimensional degree of freedom. It is not obvious if one should pick  $d = 1$  or  $d = 2$  in Harris’ criterion. Actually we prove that there are two possible criteria for disorder (ir)-relevance depending on the law  $\mathbb{P}$ , which correspond to Harris’ prediction for each value  $d \in \{1, 2\}$ . We prove disorder irrelevance (same critical point and exponent), and disorder relevance (shift of the critical point and smoothing of the phase transition) for small disorder intensity in both cases.

We will study disorder (ir)-relevance in the case of  $\hat{\omega}$  and  $\bar{\omega}$  having the same distribution

$\widehat{\mathbb{P}} = \bar{\mathbb{P}}$ , and with a product interaction function  $f(\cdot, \cdot)$ :

$$\omega_i := f(\widehat{\omega}_{i(1)}, \bar{\omega}_{i(2)}) = \widehat{\omega}_{i(1)} \times \bar{\omega}_{i(2)}. \quad (3.2.1)$$

We assume that  $\widehat{\omega}_1, \bar{\omega}_1$  are not constant a.s. (otherwise the model is homogeneous). The condition that  $\mathbb{E}[e^{\beta\omega_1}]$  is finite for  $\beta < \beta_0$  can be guaranteed simply by asking that  $\mathbb{E}[e^{\frac{1}{2}\beta\widehat{\omega}_1^2}]$  is finite for  $\beta < \beta_0$ , using that  $xy \leq (x^2 + y^2)/2$ . This is verified for example when  $\widehat{\omega}, \bar{\omega}$  are sequences of Gaussian variables, or when  $\widehat{\omega}, \bar{\omega}$  are bounded. Let us denote the moments of  $\widehat{\omega}_1$  by  $m_k := \mathbb{E}[\widehat{\omega}_1^k] = \mathbb{E}[\bar{\omega}_1^k]$  for all  $k \in \mathbb{N}$ . In particular  $\mathbb{E}[\omega_1^k] = m_k^2$  for all  $k \in \mathbb{N}$ .

### 3.2.1 Main results I: disorder irrelevance

Our first result is the following theorem, showing disorder irrelevance for  $\alpha < 1/2$ , regardless of the law  $\mathbb{P}$ .

**Theorem 3.2.1** (Irrelevance of disorder). *Let  $\tau$  and  $\tau'$  be two independent copies of a renewal process with law  $\mathbf{P}$ . If  $\tau^{(1)} \cap \tau'^{(1)}$  (or equivalently  $\tau^{(2)} \cap \tau'^{(2)}$ ) is terminating (in particular if  $\alpha < 1/2$ ), then there exists  $\beta_1 > 0$  such that for every  $\beta \in [0, \beta_1)$ , one has: (i)  $h_c(\beta) = h_c^a(\beta) = 0$ ; (ii) for any  $h \in [0, 1]$ ,*

$$L_1(1/h) h^{1/\alpha} \leq \mathbf{F}(\beta, h) \leq L_\alpha(1/h) h^{1/\alpha}, \quad (3.2.2)$$

for some (explicit) slowly varying functions  $L_1, L_\alpha$ .

Recall that  $\tau^{(1)}$  is a univariate renewal process such that  $\mathbf{P}(\tau^{(1)} = a) = \tilde{L}(a)a^{-(1+\alpha)}$ , where  $\tilde{L} \sim (1 + \alpha)^{-1}L$ . Proposition 3.B.4 gives a necessary and sufficient condition for  $\tau^{(1)} \cap \tau'^{(1)}$  to be terminating—in particular Theorem 3.2.1 holds for  $\alpha < 1/2$ . The upper bound in (3.2.2) is a direct consequence of Jensen's inequality, (3.1.12) and Theorem 3.1.2: so the interesting features are (i) and the lower bound in (3.2.2). This result implies that, provided that  $\beta$  is small enough, the quenched critical point and the quenched critical exponent are the same as those given by Theorem 3.1.2 for the homogeneous and annealed models—which means that disorder is irrelevant.

When  $\alpha > 1/2$ , we will state below that disorder is relevant for *almost* all disorder laws. Indeed, let us define for any  $x > 0$  the distribution

$$\mathbb{P}_{\pm x}(\widehat{\omega}_1 = x) = \mathbb{P}_{\pm x}(\widehat{\omega}_1 = -x) = 1/2, \quad (3.2.3)$$

(and  $\bar{\omega}_1$  has same law). Note that this family of distributions  $(\mathbb{P}_{\pm x})_{x>0}$  is characterized by the identities  $\mathbb{E}_{\pm x}[\widehat{\omega}_1] = m_1 = 0$ , and  $\text{Var}_{\pm x}(\widehat{\omega}_1^2) = m_4 - m_2^2 = 0$ —that is  $\widehat{\omega}_1^2 = x^2$  a.s..

**Theorem 3.2.2** (Irrelevance of disorder for distributions  $\mathbb{P}_{\pm x}$ ). *Assume that  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ . Then the results of Theorem 3.2.1 hold as soon as  $\tau \cap \tau'$  is terminating (in particular if  $\alpha < 1$ ). That is, there exists  $\beta_1 > 0$  such that for every  $\beta \in [0, \beta_1)$ , one has (i)  $h_c(\beta) = 0$ ; (ii) (3.2.2) holds.*

The fact that  $\alpha < 1$  is a sufficient condition for having  $\tau \cap \tau'$  terminating (while  $\alpha \leq 1$  is necessary) is ensured by Proposition 3.B.4 below. In particular, Theorem 3.2.2 shows that disorder with distribution  $\mathbb{P}_{\pm x}$  is irrelevant for all  $\alpha \in (0, 1)$ .

### 3.2.2 Main results II: disorder relevance

In the two previous theorems, we stated that if  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ , disorder is irrelevant when  $\alpha < 1/2$ ; and if  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , it is irrelevant when  $\alpha < 1$ . We now prove disorder relevance in each case for  $\alpha > 1/2$  and  $\alpha > 1$  respectively.

We first focus on the shift of the critical point, starting with the case  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ .

**Theorem 3.2.3** (Shift of the critical point, lower bound). *Assume  $\alpha > 1/2$  and  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ . Then for any fixed  $\varepsilon > 0$  (small), there exists  $\beta_\varepsilon > 0$  such that for every  $\beta \in (0, \beta_\varepsilon)$ , one has*

$$h_c(\beta) \geq \begin{cases} \beta^{\max(\frac{2\alpha}{2\alpha-1}, 2) + \varepsilon} & \text{if } m_1 \neq 0, \\ \beta^{\max(\frac{4\alpha}{2\alpha-1}, 4) + \varepsilon} & \text{if } m_1 = 0, m_4 > m_2^2. \end{cases} \quad (3.2.4)$$

In particular  $h_c(\beta) > h_c^a(\beta) = 0$  for all  $\beta \in (0, \beta_\varepsilon)$ .

Notice that all (non-constant) distributions other than  $\mathbb{P}_{\pm x}$ ,  $x > 0$  are covered by the assumptions  $m_1 \neq 0$  or  $m_4 > m_2^2$ . To complete this result, we provide an upper bound on  $h_c(\beta)$ .

**Proposition 3.2.4** (Shift of the critical point, upper bound). *Assume  $\alpha > 1/2$  and  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ . Then there exist  $\beta_1 > 0$  and some slowly varying function  $L_2$ , such that for any  $\beta \in [0, \beta_1)$ ,*

$$h_c(\beta) \leq L_2(1/\beta) \beta^{\max(\frac{2\alpha}{2\alpha-1}, 2)}. \quad (3.2.5)$$

Let us stress that this upper bound (almost) matches the lower bound found in Theorem 3.2.3 when  $m_1 \neq 0$ ; however it is not satisfactory when  $m_1 = 0$ . We will discuss in Section 3.3.3 why we strongly believe that this upper bound can be improved to (almost) match the lower bound when  $m_1 = 0$ , see Remark 3.3.5 below —actually, when  $\hat{\omega}, \bar{\omega}$  are two sequences of i.i.d., centered Gaussian variables, computations of Section 3.3.2 can be carried out exactly and give an upper bound on the shift of order  $L_2(1/\beta) \beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)}$ , which (almost) matches the lower bound from Theorem 3.2.3.

When  $m_1 = 0$ ,  $m_4 = m_2^2$  —that is  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ — and  $\alpha > 1$ , we prove (almost) optimal bounds on the critical point shift.

**Theorem 3.2.5** (Shift for distributions  $\mathbb{P}_{\pm x}$ ). *Assume  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , and  $\alpha > 1$ . Then for every  $\varepsilon > 0$ , there exist  $\beta_\varepsilon > 0$  such that for any  $\beta \in [0, \beta_\varepsilon)$ ,*

$$\beta^{\max(\frac{2\alpha}{\alpha-1}, 4) + \varepsilon} \leq h_c(\beta) \leq L_3(1/\beta) \beta^{\max(\frac{2\alpha}{\alpha-1}, 4)}, \quad (3.2.6)$$

with  $L_3$  a slowly varying function. In particular  $h_c(\beta) > 0$  for all  $\beta \in (0, \beta_\varepsilon)$ .

This fully covers the shift of the critical point for all disorder laws (except for the marginal cases:  $\alpha = 1/2$  when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  and  $\alpha = 1$  when  $\mathbb{P} = \mathbb{P}_{\pm x}$ ; we will discuss them at the end of this section).

With regards to the critical exponent, in the case  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  and  $\alpha > 1/2$  we prove that the phase transition is smoother in the disordered model than in the homogeneous one. This smoothing phenomenon has been first highlighted for the disordered pinning model by [90], and our proof follows the same lines. Unfortunately we could not prove a smoothing phenomena in the case  $\mathbb{P} = \mathbb{P}_{\pm x}$ —this will be further discussed in Section 3.2.3.

As in [90] we need an additional assumption on the disorder law  $\mathbb{P}$  (mostly for technical reasons).

**Assumption 3.2.6.** *Let  $\tilde{\mathbb{P}}_\delta$  denote the law of  $(1 + \delta)\hat{\omega}_1$  for any  $\delta \in \mathbb{R}$ . There are  $c > 0$  and  $\delta_0 > 0$  such that for all  $\delta \in (-\delta_0, \delta_0)$ ,*

$$H(\tilde{\mathbb{P}}_\delta | \mathbb{P}) := \tilde{\mathbb{E}}_\delta \left[ \log \frac{d\tilde{\mathbb{P}}_\delta}{d\mathbb{P}} \right] \leq c \delta^2, \quad (3.2.7)$$

(Recall that  $H(\tilde{\mathbb{P}}_\delta | \mathbb{P})$  is well-defined as soon as  $\frac{d\tilde{\mathbb{P}}_\delta}{d\mathbb{P}}$  exists, and it is non-negative).

Of course we make the same assumption regarding  $\bar{\omega}$  (we assumed  $\hat{\mathbb{P}} = \bar{\mathbb{P}}$ ). Notice that Assumption 3.2.6 is verified when  $\hat{\omega}, \bar{\omega}$  are Gaussian sequences, and for many unbounded laws; however it does not hold for bounded disorder, in particular it does not hold for  $\mathbb{P}_{\pm x}$ .

**Theorem 3.2.7** (Smoothing of the phase transition). *Suppose that Assumption 3.2.6 holds for  $\mathbb{P}$  (in particular  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ ). Then for any  $\beta \in (0, \beta_0)$ , there are constants  $c_\beta$  and  $t_\beta > 0$  such that for any  $t \in (0, t_\beta)$ , one has*

$$F(\beta, h_c(\beta) + t) \leq c_\beta t^2. \quad (3.2.8)$$

When  $\alpha > 1/2$ , Theorem 3.2.7 shows that disorder has a smoothing effect on the phase transition. Indeed Theorem 3.1.2 claims that the homogeneous model has a critical exponent  $1/\min(1, \alpha)$ , which is strictly smaller than 2 if  $\alpha > 1/2$ . Notice also that Theorem 3.2.7 is not restricted to small values of  $\beta$  (as opposed to other results in this paper).

### 3.2.3 Some comments on the results and the techniques of the proofs

**Criteria for disorder (ir)-relevance: dependence on  $\mathbb{P}$ .**

An interesting feature of our setting is that, unlike the PS model or gPS with i.i.d. disorder, criteria on  $\mathbb{P}$  for disorder (ir)-relevance are not the same for all disorder distributions  $\mathbb{P}$ . We may foresee this peculiarity by looking at the correlation between rewards given by two different indices, that is  $\mathbb{E}[e^{\beta(\omega_i + \omega_j)}] - \mathbb{E}[e^{\beta\omega_i}]\mathbb{E}[e^{\beta\omega_j}]$ ,  $\mathbf{i}, \mathbf{j} \in \mathbb{N}^2$  (in particular those correlations appear in the proof of Theorems 3.2.1 and 3.2.2 in Section 3.3.2, when computing the second moment of the partition function). It is obviously 0 if  $\mathbf{i}, \mathbf{j}$  are on different lines and columns, and it is greater than 0 if  $\mathbf{i} = \mathbf{j}$ . However if  $\mathbf{i} \neq \mathbf{j}$  are on the same line or column (that is  $\mathbf{i}^{(1)} = \mathbf{j}^{(1)}$  or  $\mathbf{i}^{(2)} = \mathbf{j}^{(2)}$ ) then this correlation is 0 *if and only if*

$\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$  —otherwise it is positive for  $\beta$  small (this follows from a Taylor expansion).

Therefore, when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ , the field  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$  of rewards given by the disorder has strong correlations on each line and column, and the criteria for disorder (ir)-relevance in that case correspond to Harris’ prediction for one-dimensional systems, *i.e.* the marginal regime is at the value  $\alpha = 1/2$ . Whereas when  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , that field is much less correlated (all its two-point correlations are 0), and it matches Harris’ prediction for two-dimensional systems, *i.e.* the marginal regime is at  $\alpha = 1$ .

### Comparison with the existing literature and other models.

Let us compare our results to [18], where the authors studied the gPS model with an i.i.d. disorder field  $(\omega_i)_{i \in \mathbb{N}^2}$ . In this setting, they prove that disorder is irrelevant as soon as  $\tau \cap \tau'$  terminates, and that the critical point is shifted for all  $\alpha > 1$  by  $\beta^{\max(\frac{2\alpha}{\alpha-1}, 4) + \varepsilon}$ . Those results are the same as our Theorems 3.2.2 and 3.2.5. This supports the idea that in our setting, when  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , the field  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$  is poorly correlated (even though it is not i.i.d.) and has the same influence on the system as an i.i.d. field. However, when disorder has another distribution, the field becomes highly correlated and the comparison with i.i.d. disorder falls apart —in particular the criteria for disorder (ir)-relevance are not the same.

Another interesting comparison is to the standard PS model (or the pinning model). The question of disorder relevance has been studied extensively in that context, see [6, 20, 64, 84, 90, 147] among others. For that model, it has been proven that disorder is irrelevant if and only if the process  $\tau \cap \tau'$  terminates, with  $\tau, \tau'$  two independent copies of the univariate renewal process —in particular it is irrelevant for all  $\alpha < 1/2$  and relevant for all  $\alpha > 1/2$ , where we assume that the inter-arrival distribution is  $\mathbf{P}(\tau_1 = a) = L(a)a^{-(1+\alpha)}$ . With regards to Theorems 3.2.1 and 3.2.3, and noticing that the processes  $\tau^{(r)}$  in our setting and  $\tau$  in the PS model are very similar, we observe analogous criteria for disorder (ir)-relevance in both systems —the standard PS model and our setting with  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ — with a marginal value  $\alpha = 1/2$ .

Finally, the order of the shift of the critical point when  $m_1 \neq 0$  in our setting matches the standard PS model —that is  $\beta^{\max(\frac{2\alpha}{2\alpha-1}, 2)}$  — but not when  $m_1 = 0$ . Rewriting the disorder field  $\omega_{(i,j)} = \zeta_i \xi_j + m_1 \zeta_i + m_1 \xi_j + m_1^2$  with  $\zeta_i, \xi_j$  independent, centered variables, this seems to imply that when  $m_1 \neq 0$ , the main contribution to the disorder are the terms  $m_1 \zeta_i + m_1 \xi_j$ , which are very similar to the disorder of the PS model ; whereas if  $m_1 = 0$ , the remaining term  $\zeta_i \xi_j$  carries a lower disorder intensity, leading to a shift of the critical point of smaller order.

### About the marginal cases.

In our paper we did not fully treat the marginal cases, that is  $\alpha = 1/2$  when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  and  $\alpha = 1$  when  $\mathbb{P} = \mathbb{P}_{\pm x}$ . By comparing our setting to the PS model, it is expected that the assumptions of Theorems 3.2.1 and 3.2.2 are optimal, and that a shift of the critical point can be proven whenever  $\tau^{(r)} \cap \tau'^{(r)}$  is persistent for  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ , respectively whenever  $\tau \cap \tau'$

is persistent for  $\mathbb{P} = \mathbb{P}_{\pm x}$  (notice that in [18] the authors make the same conjecture for the gPS model with i.i.d. disorder). This is only conjectural for now, and we expect that a great amount of technical work is needed in both cases (for instance, bivariate renewal estimates are very complex in the case  $\alpha = 1$ , see [17]).

### About higher disorder intensities.

To answer the issue of disorder (ir)-relevance, we only had to consider small disorder intensities  $\beta$  (similarly to [18]). It would still be interesting to study the disordered model for any  $\beta \in (0, \beta_0)$ , yet it seems particularly challenging. First, there is no clear reason for  $h_c(\beta)$  to be non-decreasing in  $\beta$ , hence for  $h_c(\beta)$  to be positive for all  $\beta > 0$  under the assumptions of Theorems 3.2.3 or 3.2.5 (the proof from [88] for the PS model does not apply here due to the correlations of  $\omega$ ). Second, apart for the smoothing of the phase transition, our methods rely heavily on taking small values of  $\beta$ : regarding disorder irrelevance in Section 3.3, we require that  $\lambda(2\beta) - 2\lambda(\beta)$  is sufficiently small to ensure that some exponential moments are finite; and regarding the shift of the critical point, we require an upper bound like (3.5.37) in the change of measure procedure, which we obtained with a Taylor expansion argument. Both those estimates are not guaranteed for larger values of  $\beta$ , and may even depend heavily on the choice of disorder distribution  $\mathbb{P}$ .

Nonetheless when  $\beta$  is very large (close to  $\beta_0$ ), one could adapt [149, Theorem 2.1] to prove that there is a strong disorder regime even when  $\alpha < 1/2$  (see [149, Corollary 3.2, Remark 3.3]): this is based on a quite straightforward fractional moment estimate.

### About the proof of disorder relevance

Let us make some technical comments on our results for disorder relevance, starting with the shift of the critical point. Our lower bounds on the critical point shift are obtained with a coarse graining procedure, together with estimates on the fractional moments of the partition function obtained via a change of measure argument. This method has first been applied in [64] for the PS model, and was adapted to the i.i.d. gPS model in [18]. In this paper we use the same coarse-graining procedure as [18], but the change of measure argument—which in [18] relies on an i.i.d. tilt of the field  $\omega$ —cannot be replicated straightforwardly in the general case, because of the non-independent structure of  $\omega$  (we only do it for distributions  $\mathbb{P}_{\pm x}$  in Section 3.6, when proving the lower bound of Theorem 3.2.5). In the case  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ , we introduce another change of measure—a simultaneous tilt of both sequences  $\hat{\omega}$  and  $\bar{\omega}$ —to prove Theorem 3.2.3. Interestingly enough, this change of measure relies on the correlated structure of the disorder, and does not lead to pertinent estimates in the case  $\mathbb{P} = \mathbb{P}_{\pm x}$ . Moreover, it is less costly than the tilt of the field  $\omega$ —for a system of size  $\mathbf{n} = (n, n)$ , we tilt  $2n$  variables instead of  $n^2$ . In comparison to [18] or the case  $\mathbb{P} = \mathbb{P}_{\pm x}$ , this induces the appearance of a shift of the critical point when  $\alpha \in (1/2, 1]$ , and a greater shift when  $\alpha \in (1, 2]$ .

Let us stress that we have not pursued optimal upper and lower bounds on the critical point shift: the  $\beta^\varepsilon$  in Theorems 3.2.3-3.2.5 can certainly be replaced by slowly varying



functions with a more sophisticated coarse graining technique, (see [20, 50] and [83, Ch. 6] for the PS model). However the proofs of Theorems 3.2.3 and 3.2.5 are already rather technical, and getting sharper bounds would have been even more laborious; so we decided to stick with these “almost” optimal lower bounds for the sake of clarity.

As far as Theorem 3.2.7 is concerned, a small disappointment in our result is that our proof relies heavily on Assumption 3.2.6, whereas one could expect to prove the same smoothing inequality for any disorder law other than  $\mathbb{P}_{\pm x}$ —the same way smoothing inequalities have been more recently proven for all laws in some disordered polymer models (including PS and copolymer models, see [43]). Because the disorder we consider is highly correlated, many technical difficulties appear when trying to prove a smoothing inequality, even under strong technical assumptions (*e.g.* bounded and symmetric disorder). We do not try to expand Theorem 3.2.7 under other assumptions in this paper, because we could not handle the computations without very restrictive assumptions, and even so the computations remain extremely cumbersome.

In the case  $\mathbb{P} = \mathbb{P}_{\pm x}$ , we do not have any result on the smoothing of the phase transition when  $\alpha > 1$ . Actually, proving a smoothing inequality is also an issue in the gPS model with i.i.d. disorder. It is conjectured in [18, Conj. 1.5] that the gPS model with (say Gaussian) i.i.d. disorder undergoes a smoothing phenomenon with exponent  $\min(\frac{2\alpha}{\alpha+1}, \frac{4}{3})$  when  $\alpha > 1$ . Regarding our previous comments about the similarity between the i.i.d. gPS model and our setting with  $\mathbb{P} = \mathbb{P}_{\pm x}$ , and comments in [18] leading the authors to their conjecture, it is reasonable to expect that a similar smoothing phenomenon occurs in our setting for disorder with distribution  $\mathbb{P}_{\pm x}$ , but it is purely conjectural for now.

### Open questions and perspectives

Let us list here some questions on this model that are not answered yet, and perspectives for further study:

(i) Improving some of our results which are not fully satisfactory, namely: proving a smoothing inequality for  $\mathbb{P} = \mathbb{P}_{\pm x}$  (this matter seems highly related to the gPS model with i.i.d. disorder), proving a smoothing without additional assumptions when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  (which seems very technical), improving the upper bound from Proposition 3.2.4 to match Theorem 3.2.3 when  $m_1 = 0$ .

(ii) Dealing with the marginal cases, in particular  $\mathbb{P} = \mathbb{P}_{\pm x}$  and  $\alpha = 1$ , because bivariate renewal processes with  $\alpha = 1$  are not extensively understood yet.

(iii) We also omitted the case  $\alpha = 0$ , for two reasons: the case  $\alpha = 0$  should be “strongly irrelevant”, in the sens that the quenched and annealed critical points should *always* be equal (in the same spirit as in [9]), so it should somehow be easier to prove disorder irrelevance; from a technical point of view, there are no estimates for bivariate renewals in the case  $\alpha = 0$  (there is no  $\alpha$ -stable domain of attraction), and this should therefore require a separate analysis.

(iv) Dealing with higher values of  $\beta$ . As mentioned above, proving a strong disorder regime when  $\beta$  approaches  $\beta_0$  should be feasible even when  $\alpha < 1/2$  by adapting [149, The-

orem 2.1]. However handling intermediate values of  $\beta$  may require much more technicalities or even different methods.

(v) As stated in [87], the gPS model undergoes other phase transitions. What is the effect of disorder on those? Do they still occur in the disordered model and are their features modified? These questions are addressed in [128], but only via heuristic and numerical arguments. More recently, in [86], the authors treat generalized (one-dimensional) pinning models whose homogeneous versions exhibit a *big-jump* phase transition —analogous to the other phase transitions in the homogeneous gPS model, see [19]. They show that the presence of disorder is incompatible with the presence of a big jump, and therefore completely smears out the *big-jump* phase transition (contradicting the conclusions of [128]). This suggests in our case that only the localization/delocalization phase transition survives in the disordered gPS model.

(vi) About the choice of disorder : with regards to the interaction function  $f$ , we chose a product function in (3.2.1), but we can conjecture that a generic (symmetric) function should lead to the same criteria for disorder (ir)-relevance, depending on the correlations of the field  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$ . However we assumed the two sequences  $\hat{\omega}, \bar{\omega}$  (*i.e.* the two strands) to be independent, while it is known that two DNA strands have a strong symmetry (an A-base on one strand faces a T-base on the other, same for C- and G-bases). Therefore, a more pertinent choice of disorder field would be

$$\omega_i := f(\hat{\omega}_{i^{(1)}}, \hat{\omega}_{i^{(2)}}), \quad (3.2.9)$$

with  $\hat{\omega}$  an (i.i.d.) sequence of random variables. This setting represents even more faithfully the denaturation of DNA, and it has been considered in numerical studies (see [78, 77]). But it is much more difficult to handle than the one in this paper: the argument of Proposition 3.1.1 does not apply (even if the sequence  $\hat{\omega}$  is still assumed i.i.d.), hence even the convergence of the quenched free energy is not clear; moreover the annealed model does not match the homogeneous one anymore, and its analysis seems very challenging.

### 3.2.4 Notation and organisation of the paper

Let us introduce some notations for the subsequent sections.  $C_1, C_2, \dots$  will denote some constants, and  $L_1, L_2, \dots$  some slowly varying functions. Unless otherwise specified, they may depend on  $\alpha$  but will not depend on any other parameter  $n, h, \beta, \dots$ . Moreover  $L$  will always denote the function of the inter-arrival distribution (see (3.1.1)), and  $L_\alpha$  the function introduced in Theorem 3.1.2 for the critical behavior of the homogeneous model.

For any  $\mathbf{i}, \mathbf{j} \in \mathbb{N}^2$ , we will note  $\mathbf{i} \leftrightarrow \mathbf{j}$  if  $\mathbf{i} = \mathbf{j}$  or  $\mathbf{i}$  and  $\mathbf{j}$  are on the same line or column (*i.e.* if  $\mathbf{i}^{(r)} = \mathbf{j}^{(r)}$  for some  $r \in \{1, 2\}$ ), recall the notation  $\mathbf{i} = (\mathbf{i}^{(1)}, \mathbf{i}^{(2)})$  for all  $\mathbf{i} \in \mathbb{N}^2$ . We also introduce an order on  $\mathbb{N}^2$ :

$$\begin{aligned} \mathbf{i} \prec \mathbf{j} & \text{ if } \mathbf{i}^{(1)} < \mathbf{j}^{(1)}, \mathbf{i}^{(2)} < \mathbf{j}^{(2)}, \\ \mathbf{i} \preceq \mathbf{j} & \text{ if } \mathbf{i}^{(1)} \leq \mathbf{j}^{(1)}, \mathbf{i}^{(2)} \leq \mathbf{j}^{(2)}. \end{aligned} \quad (3.2.10)$$

For all  $a < b \in \mathbb{N}$ , we define  $\llbracket a, b \rrbracket := [a, b] \cap \mathbb{N}$ . For all  $\mathbf{a} \preccurlyeq \mathbf{b} \in \mathbb{N}^2$ ,  $\llbracket \mathbf{a}, \mathbf{b} \rrbracket$  denotes the rectangle in  $\mathbb{N}^2$  with bottom-left corner  $\mathbf{a}$  and top-right corner  $\mathbf{b}$ :  $\llbracket \mathbf{a}, \mathbf{b} \rrbracket := \{\mathbf{m} \in \mathbb{N}^2; \mathbf{a} \preccurlyeq \mathbf{m} \preccurlyeq \mathbf{b}\}$ . Finally we will note  $\|\cdot\|$  the  $L^1$  norm on  $\mathbb{N}^2$ :  $\|\mathbf{i}\| := \mathbf{i}^{(1)} + \mathbf{i}^{(2)}$  for any  $\mathbf{i} \in \mathbb{N}^2$ .

The remainder of the paper is organized into four sections. In Section 3.3 we show how to use a second moment method together with second moment estimates to prove both disorder irrelevance (Theorems 3.2.1 and 3.2.2) and upper bounds on the shift of the critical point (Proposition 3.2.4 and right-hand-side inequality in Theorem 3.2.5). In Section 3.4 we prove the smoothing inequality of Theorem 3.2.7 under Assumption 3.2.6, via a rare-stretch strategy. In Section 3.5, we display the coarse-graining method, and compute estimates on the fractional moments to obtain lower bounds on the shift of the critical point when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$  (Theorem 3.2.3) Finally in Section 3.6, we adapt the change of measure argument from [18] to distributions  $\mathbb{P}_{\pm x}$  to prove the same estimates on the fractional moments, thereby proving the left-hand side inequality in Theorem 3.2.5. Section 3.6 relies on the coarse-graining procedure introduced in Section 3.5, the other sections are independent.

Additionally, we provide in Appendix 3.A some computations on the partition function of the homogeneous gPS model that are needed in Section 3.5, in Appendix 3.B we collect useful estimates on bivariate renewals, and in Appendix 3.C we compute the upper bound claimed below Proposition 3.2.4 for Gaussian disorder.

### 3.3 Disorder irrelevance

In this section, we choose  $\mathbf{n} = (n, n)$  where  $n \in \mathbb{N}$  (recall that we assume  $\gamma = 1$ ). Let us define the *free* version of the gPS model, where the constraint  $1_{\{\mathbf{n} \in \tau\}}$  is removed in (3.1.3), *i.e.* the endpoints are free: its partition function is defined as

$$Z_{\mathbf{n},h}^{\beta,q,free} := \mathbf{E} \left[ \exp \left( \sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) 1_{\{i \in \tau\}} \right) \right]. \quad (3.3.1)$$

We claim that the constrained and free partition functions are comparable: more precisely for any  $\alpha_+ > \alpha$  and  $\mathbf{m} \in \mathbb{N}^2$ ,

$$Z_{\mathbf{m},h}^{\beta,q} \leq Z_{\mathbf{m},h}^{\beta,q,free} \leq Z_{\mathbf{m},h}^{\beta,q} \left( 1 + C_1 \|\mathbf{m}\|^{3+\alpha_+} \sup_{l \leq \mathbf{m}, l \leftrightarrow \mathbf{m}} (e^{\beta(\omega_l - \omega_{\mathbf{m}})}) \right), \quad (3.3.2)$$

where  $C_1$  is a uniform constant (see [18, Lem. 2.2]: here again the proof is not altered by our setting of disorder). In particular, the free energy is not modified with respect to (3.1.6): we have that  $F(\beta, h) = \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{\mathbf{n},h}^{\beta,q,free}$ , both  $\mathbb{P}(d\omega)$ -a.s. and in  $L^1(\mathbb{P})$ .

Note that, as in [19], we defined the free partition function simply by removing the constraint  $\{\mathbf{n} \in \tau\}$ , but let us mention that in the literature the free ends may be assigned some entropic cost: for instance see [87], where the authors consider the (homogeneous) free partition function  $\sum_{i,j \leq n} K_f(i) K_f(j) Z_{\mathbf{n},h}^{0,q}$ , with  $K_f$  some regularly varying function. However this would not affect our results, since it has no effect at the level of the free energy.

### 3.3.1 Second moment method

Our proof of disorder irrelevance relies on the idea that, if  $\sup_{n \in \mathbb{N}} \mathbb{E}[(Z_{n,0}^{\beta,q,free})^2] < +\infty$ , then  $Z_{n,h}^{\beta,q,free}$  should remain concentrated around its mean, at least in a certain regime for  $n, h$ ; in particular the quenched and annealed free energy should remain close to each other (this idea has been exploited and turned into a proof in [6, 147] for the PS model). In [115] (for the PS model again), it is roughly showed that as long as  $\mathbb{E}[(Z_{n,0}^{\beta,q,free})^2]$  is of order 1, the measure  $\mathbf{P}_{n,0}^{\beta,q,free}$  does not differ much from  $\mathbf{P}$ —this has also been exploited in [20]. We use this idea to obtain the following statement.

**Proposition 3.3.1.** *Fix some constant  $C > 1$ , and define*

$$n_\beta := \sup \left\{ n \in \mathbb{N}; \mathbb{E}[(Z_{n,0}^{\beta,q,free})^2] \leq C \right\}. \quad (3.3.3)$$

*Then there is some (explicit) slowly varying function  $L_1$  such that the critical point satisfies*

$$0 \leq h_c(\beta) \leq L_1(n) n^{-\min(\alpha,1)}, \quad (3.3.4)$$

*for any  $n \leq n_\beta$ .*

*If  $(Z_{n,0}^{\beta,q,free})_{n \in \mathbb{N}}$  is bounded in  $L^2(\mathbb{P})$ , then  $n_\beta = +\infty$  (provided that  $C$  had been fixed large), so  $h_c(\beta) = 0$ ; moreover there exists a slowly varying function  $L_2$  such that for all  $h \in (0, 1)$ ,*

$$\mathbf{F}(\beta, h) \geq L_2(1/h) h^{1/\min(1,\alpha)}. \quad (3.3.5)$$

Before we prove this proposition, notice that it fully implies the non-relevance of the disorder as soon as  $(Z_{n,0}^{\beta,q,free})_{n \in \mathbb{N}}$  is bounded in  $L^2(\mathbb{P})$ . We show in Section 3.3.2 that this holds under the assumptions of either Theorem 3.2.1 or 3.2.2 with  $\beta$  small, proving both theorems. Otherwise one can use (3.3.4) and an estimate of  $n_\beta$  to obtain an upper bound for the shift of the critical point, which we do in Section 3.3.3.

(Note that it is not self-evident that  $\mathbb{E}[(Z_{n,0}^{\beta,q,free})^2] < \infty$  for any  $n \in \mathbb{N}$ . We actually prove it in Section 3.3.2 for  $\beta < \beta_0/2$ ).

*Proof.* We already stated the left inequality of (3.3.4) with Jensen's inequality in (3.1.13). Let us prove the right inequality. Fix  $h \in \mathbb{R}$ . We note that the sequence  $(\mathbb{E} \log Z_{n,h}^{\beta,q})_{n \in \mathbb{N}}$  is super-additive, so (3.1.6) and (3.3.2) give for any  $n \in \mathbb{N}$ ,

$$\mathbf{F}(\beta, h) = \sup_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E} \log Z_{n,h}^{\beta,q} \geq \frac{1}{n} \mathbb{E} \log Z_{n,h}^{\beta,q,free} - C_2 \frac{\log n}{n}. \quad (3.3.6)$$

Moreover we have:

$$\begin{aligned} \mathbb{E} \log Z_{n,h}^{\beta,q,free} &\geq h \mathbb{E} \left[ \partial_h \log Z_{n,h}^{\beta,q,free} \Big|_{h=0} \right] + \mathbb{E} \left[ \log Z_{n,0}^{\beta,q,free} \right] \\ &\geq h \mathbb{E} \mathbf{E}_{n,0}^{\beta,q,free} \left[ \sum_{i \in \llbracket 1, n \rrbracket} 1_{\{i \in \tau\}} \right] - C_3 \log n, \end{aligned} \quad (3.3.7)$$

where we used the convexity of  $h \mapsto \log Z_{n,h}^{\beta,q,free}$ , and the obvious bound  $Z_{n,h}^{\beta,q,free} \geq \mathbf{P}(\|\tau_1\| > 2n)$ . Here we have to estimate the contact fraction of the renewal process under  $\mathbf{P}_{n,0}^{\beta,q,free}$ . We first do it under  $\mathbf{P}$ .

**Lemma 3.3.2.** For any  $\varepsilon > 0$ , there exist  $C_4 > 0$  and  $n_0$  such that for any  $n \geq n_0$ ,

$$\mathbf{P}\left(|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket| \geq C_4 f(n)\right) \geq 1 - \varepsilon, \quad (3.3.8)$$

where  $f(n) := n^\alpha/L(n)$  if  $\alpha \in (0, 1)$  and  $f(n) := n/\mu(n)$  if  $\alpha \geq 1$ , with  $\mu(n) = \mathbf{E}[\tau_1^{(1)}]$  if  $\mathbf{E}[\tau_1^{(1)}] < +\infty$  and  $\mu(n) := \mathbf{E}[\tau_1^{(1)} \mathbf{1}_{\{\tau_1^{(1)} \leq n\}}]$  if  $\mathbf{E}[\tau_1^{(1)}] = +\infty$ .

*Proof.* When  $\alpha \in (0, 1)$ , this result is already proved in [18, Lem. A.2]. When  $\alpha \geq 1$ , we have that

$$\frac{|\tau^{(1)} \cap [1, n]|}{f(n)} \rightarrow 1 \quad \text{in probability.} \quad (3.3.9)$$

Indeed, when  $\mathbf{E}[\tau_1^{(1)}] < +\infty$  then this is simply the law of large numbers, whereas when  $\alpha = 1$  and  $\mathbf{E}[\tau_1^{(1)}] = +\infty$ , this is for example in [95]. Hence, provided that  $C_4 < 1$ , then  $\mathbf{P}\left(|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket| < C_4 f(n)\right) \rightarrow 0$  as  $n \rightarrow \infty$ , which proves the lemma.  $\square$

Let us denote  $A_n := \{|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket| \geq C_4 f(n)\}$ . We can estimate the contact fraction under  $\mathbf{P}_{n,0}^{\beta, \mathbf{q}, \text{free}}$  with the simple inequality:

$$\mathbf{P}_{n,0}^{\beta, \mathbf{q}, \text{free}}(A_n^c) \leq \mathbf{1}_{\{Z_{n,0}^{\beta, \mathbf{q}, \text{free}} \leq 1/2\}} + 2\mathbf{E}\left[\mathbf{1}_{A_n^c} e^{\sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta)) \mathbf{1}_{\{i \in \tau\}}}\right], \quad (3.3.10)$$

so we obtain for any  $\varepsilon > 0$  and  $n \geq n_0$ ,

$$\begin{aligned} \mathbb{E}\left[\mathbf{P}_{n,0}^{\beta, \mathbf{q}, \text{free}}(A_n^c)\right] &\leq \mathbb{P}\left(Z_{n,0}^{\beta, \mathbf{q}, \text{free}} \leq 1/2\right) + 2\mathbf{P}(A_n^c) \\ &\leq 1 - \frac{1}{4\mathbb{E}\left[(Z_{n,0}^{\beta, \mathbf{q}, \text{free}})^2\right]} + 2\varepsilon, \end{aligned} \quad (3.3.11)$$

where we used Paley-Zygmund inequality and Lemma 3.3.2. If we choose  $\varepsilon$  small enough (more precisely  $\varepsilon < (8C)^{-1}$ ), and  $n \leq n_\beta$ , this implies

$$\mathbb{E}\mathbf{E}_{n,0}^{\beta, \mathbf{q}, \text{free}}\left[|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|\right] \geq C_4 f(n) \mathbb{E}\left[\mathbf{P}_{n,0}^{\beta, \mathbf{q}, \text{free}}(A_n)\right] \geq C_5 f(n), \quad (3.3.12)$$

(where we get rid of the condition  $n \geq n_0$  by adjusting  $C_5$  for finitely many terms). Going back to (3.3.6) and (3.3.7), we finally obtain the following lower bound for any  $n \leq n_\beta$ :

$$\mathbf{F}(\beta, h) \geq \frac{C_6}{n} \left(hf(n) - C_7 \log n\right), \quad (3.3.13)$$

where we recall  $f(n) := n/\mu(n)$  if  $\alpha \geq 1$  (with  $\mu(n)$  either a constant, or going to  $+\infty$  as a slowly varying function in the case  $\alpha = 1$ ,  $\mathbf{E}[\tau_1^{(1)}] = +\infty$ .), and  $f(n) := n^\alpha/L(n)$  if  $\alpha \in (0, 1)$ . If we take  $h = h_c(\beta)$  in (3.3.13), then  $\mathbf{F}(\beta, h_c(\beta)) = 0$ , so we get

$$h_c(\beta) \leq C_8 \frac{\log n}{f(n)}, \quad (3.3.14)$$

which concludes the proof of (3.3.4), with  $L_1(n) = C_8 L(n) \log(n)$  if  $\alpha \in (0, 1)$ ;  $L_1(n) = C_8 \mu(n) \log n$  if  $\alpha \geq 1$ .

If  $(Z_{n,0}^{\beta, \mathfrak{q}, \text{free}})_{n \in \mathbb{N}}$  is bounded in  $L^2(\mathbb{P})$ , then we can choose  $C$  such that  $n_\beta = \infty$ , so (3.3.4) holds for any  $n \in \mathbb{N}$  and we obviously have  $h_c(\beta) = 0$ . Furthermore (3.3.13) also holds for any  $n \in \mathbb{N}$ , so if we take  $h > 0$  and  $n = C_9 V_\alpha(1/h)$  where  $V_\alpha$  is the asymptotical inverse of  $b \mapsto b^{\min(\alpha, 1)} L_1(b)^{-1}$  as  $b \rightarrow \infty$  and  $C_9$  is a suitable constant, we finally obtain (3.3.5) from (3.3.13).  $\square$

### 3.3.2 Second moment estimates

With regard to Proposition 3.3.1, it suffices to estimate  $\mathbb{E}[(Z_{n,0}^{\beta, \mathfrak{q}, \text{free}})^2]$  to prove disorder irrelevance, or at least an upper bound on the shift of the critical point. To do so we introduce two independent copies  $\tau, \tau'$  of a renewal process with law  $\mathbb{P}$ .

**Proposition 3.3.3.** *For any  $\beta \in [0, \beta_0/2)$  one has  $Z_{n,0}^{\beta, \mathfrak{q}, \text{free}} \in L^2(\mathbb{P})$  for all  $n \in \mathbb{N}$  and:*

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[ \left( Z_{n,0}^{\beta, \mathfrak{q}, \text{free}} \right)^2 \right] \leq \mathbf{E}_{(\tau, \tau')} \left[ e^{\frac{3}{2} (\lambda(2\beta) - 2\lambda(\beta)) (|\tau^{(1)} \cap \tau'^{(1)}| + |\tau^{(2)} \cap \tau'^{(2)}|)} \right], \quad (3.3.15)$$

When  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , one has for any  $\beta \in \mathbb{R}_+$ :

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\pm x} \left[ \left( Z_{n,0}^{\beta, \mathfrak{q}, \text{free}} \right)^2 \right] = \mathbf{E}_{(\tau, \tau')} \left[ e^{(\lambda(2\beta) - 2\lambda(\beta)) |\tau \cap \tau'|} \right]. \quad (3.3.16)$$

Note that this proposition also applies for any sequence of indices  $\mathbf{n} \in \mathbb{N}^2$  such that  $\mathbf{n}^{(1)}, \mathbf{n}^{(2)} \rightarrow \infty$ . Plugging those estimates into Proposition 3.3.1, this straightforwardly proves Theorems 3.2.1 and 3.2.2.

*Proof of Theorems 3.2.1-3.2.2.* In the general case we use the upper bound (3.3.15). If  $\tau^{(1)} \cap \tau'^{(1)}$  is terminating, then  $|\tau^{(1)} \cap \tau'^{(1)}|$  follows a geometric distribution (recall that  $\tau^{(1)}$  and  $\tau'^{(1)}$  are independent univariate renewal processes), so it has some finite exponential moment. The same holds for  $\tau^{(2)} \cap \tau'^{(2)}$ , and for their sum. Because  $\lambda(2\beta) - 2\lambda(\beta) \sim c\beta^2$  for some  $c > 0$  when  $\beta \rightarrow 0$ , there exists  $\beta_1$  such that the right hand side of (3.3.15) is finite for any  $\beta \in [0, \beta_1)$ , so  $(Z_{n,0}^{\beta, \mathfrak{q}, \text{free}})_{n \in \mathbb{N}}$  is bounded in  $L^2(\mathbb{P})$ . Applying Proposition 3.3.1, this proves Theorem 3.2.1.

When  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ , the same argument applies with (3.3.16) and  $\tau \cap \tau'$ —if  $\tau \cap \tau'$  is terminating then  $|\tau \cap \tau'|$  also follows a geometric distribution—so this proves Theorem 3.2.2.  $\square$

Before proving Proposition 3.3.3, we need to introduce some new notations. Using a replica trick and Fubini-Tonelli theorem, we can write the second moment of the partition function as

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta, \mathfrak{q}, \text{free}} \right)^2 \right] = \mathbf{E}_{(\tau, \tau')} \mathbb{E} \left[ \exp \left( \sum_{i \in [1, n]} (\beta \omega_i - \lambda(\beta)) (1_{\{i \in \tau\}} + 1_{\{i \in \tau'\}}) \right) \right]. \quad (3.3.17)$$

For any trajectories  $\tau, \tau' \subset \mathbb{N}^2$ ,  $\tau$  can be written as a sequence  $\tau = \{\tau_1, \tau_2, \dots\}$  with  $\tau_k^{(r)} < \tau_{k+1}^{(r)}$  for every  $r \in \{1, 2\}, k \in \mathbb{N}$ , and the same holds for  $\tau'$ . One can rewrite (3.3.17) by taking the sum on  $\mathbf{i} \in (\tau \cup \tau') \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket$  (other terms are 0). We claim that there are three kind of points in  $\tau \cup \tau'$  contributing to this sum:

- $\mathbf{i} \in \tau \cap \tau'$ . We will call such  $\mathbf{i}$  *double points* of  $\tau \cup \tau'$ .
- $\mathbf{i} \in \tau \cup \tau' \setminus \tau \cap \tau'$ , and  $\mathbf{i}^{(1)} \notin \tau^{(1)} \cap \tau'^{(1)}$  and  $\mathbf{i}^{(2)} \notin \tau^{(2)} \cap \tau'^{(2)}$ : that is,  $\mathbf{i}$  is in either  $\tau$  or  $\tau'$ , and no other point from  $\tau \cup \tau'$  is on the same line or column than  $\mathbf{i}$ . We will call those *isolated points* of  $\tau \cup \tau'$ .
- $\mathbf{i} \in \tau \cup \tau' \setminus \tau \cap \tau'$ , and  $\mathbf{i}^{(1)} \in \tau^{(1)} \cap \tau'^{(1)}$  or  $\mathbf{i}^{(2)} \in \tau^{(2)} \cap \tau'^{(2)}$ . We will call those *chained points* of  $\tau \cup \tau'$ .

Let us explain the denomination *chained points*. Let  $\mathbf{i}_1 \in \tau \cup \tau'$  be the first chained point for the lexical order on  $\mathbb{N}^2$ . Assume  $\mathbf{i}_1 \in \tau \setminus \tau'$  without loss of generality. Then there exists  $\mathbf{i}_2 \in \tau' \setminus \tau$  such that  $\mathbf{i}_1 \leftrightarrow \mathbf{i}_2$  and  $\mathbf{i}_1^{(r)} = \mathbf{i}_2^{(r)}, \mathbf{i}_1^{(3-r)} < \mathbf{i}_2^{(3-r)}$  for one  $r \in \{1, 2\}$  (because  $\mathbf{i}_1$  comes first in the lexical order). Moreover there is no other point  $\mathbf{i} \in \tau \cup \tau', \mathbf{i} \neq \mathbf{i}_2$  such that  $\mathbf{i} \leftrightarrow \mathbf{i}_1, \mathbf{i} \neq \mathbf{i}_1$ : indeed, it would imply  $\mathbf{i} \in \tau' \setminus \tau$ , which is impossible since  $\tau'$  is strictly increasing on each coordinate.

Let us now assume that there exists another  $\mathbf{i}_3 \in \tau \cup \tau', \mathbf{i}_3 \notin \{\mathbf{i}_1, \mathbf{i}_2\}$  such that  $\mathbf{i}_3 \leftrightarrow \mathbf{i}_2$ . Then we obviously have  $\mathbf{i}_3 \in \tau \setminus \tau'$  (because  $\mathbf{i}_2 \in \tau'$  and the sequence  $\tau'$  has strictly increasing coordinates), and  $\mathbf{i}_2^{(3-r)} = \mathbf{i}_3^{(3-r)}, \mathbf{i}_2^{(r)} < \mathbf{i}_3^{(r)}$  with  $r \in \{1, 2\}$  given above, since  $\tau$  has strictly increasing coordinates. Moreover, this  $\mathbf{i}_3$  is unique (if it exists), because of the same argument that proved  $\mathbf{i}_2$  is unique.

By repeating this process until there is no more  $\mathbf{i} \in \tau \cup \tau'$  satisfying  $\mathbf{i} \leftrightarrow \mathbf{i}_k, \mathbf{i} \notin \{\mathbf{i}_1, \dots, \mathbf{i}_k\}$ , we define a sequence  $(\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_k)$  of points in  $\tau \cup \tau'$  with  $k \geq 2$ , such that:

- $\mathbf{i}_1, \mathbf{i}_3, \dots \in \tau \setminus \tau'$  and  $\mathbf{i}_2, \mathbf{i}_4, \dots \in \tau' \setminus \tau$  (or the other way around).
- There is  $r \in \{1, 2\}$  such that for any  $1 \leq l \leq k-1$ , one has  $\mathbf{i}_l^{(r)} = \mathbf{i}_{l+1}^{(r)}, \mathbf{i}_l^{(3-r)} < \mathbf{i}_{l+1}^{(3-r)}$  if  $l$  is odd, and  $\mathbf{i}_l^{(3-r)} = \mathbf{i}_{l+1}^{(3-r)}, \mathbf{i}_l^{(r)} < \mathbf{i}_{l+1}^{(r)}$  if  $l$  is even. In particular,  $\mathbf{i}_l \leftrightarrow \mathbf{i}_{l+1}$  for any  $1 \leq l \leq k-1$ .
- There is no  $\mathbf{i} \in \tau \cup \tau', \mathbf{i} \notin \{\mathbf{i}_1, \dots, \mathbf{i}_k\}$  such that  $\mathbf{i} \leftrightarrow \mathbf{i}_l$  for some  $1 \leq l \leq k$ .

We call  $\sigma_1 = (\mathbf{i}_1, \dots, \mathbf{i}_k)$  a *chain of points* in  $\tau \cup \tau'$ . Note that  $\mathbf{i}_1, \dots, \mathbf{i}_k$  are all *chained points* as defined previously. Note also that this construction may lead to an infinite chain in  $\mathbb{N}^2$ , but is always finite if we restrict  $\tau \cup \tau'$  to  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket$ .

Furthermore, if we apply the same construction process to  $\tau \cup \tau' \setminus \sigma_1$ , we can define another chain of points  $\sigma_2$  satisfying the same properties ; and by repeating it again, we obtain a sequence  $(\sigma_1, \sigma_2, \dots)$  of chains of points in  $\tau \cup \tau'$  satisfying the same properties (once again this sequence is finite for  $\tau \cup \tau'$  restricted to  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket$ , and may be infinite in  $\mathbb{N}^2$ ). Moreover, this sequence covers all chained points (indeed, any chained point  $\mathbf{i} \in \tau \cup \tau'$  is only preceded by a finite number of points in  $\tau \cup \tau'$  for the lexical order, so the construction process always reaches  $\mathbf{i}$ ).

Using this construction, we can partition  $\tau \cup \tau'$  as stated in the following proposition.

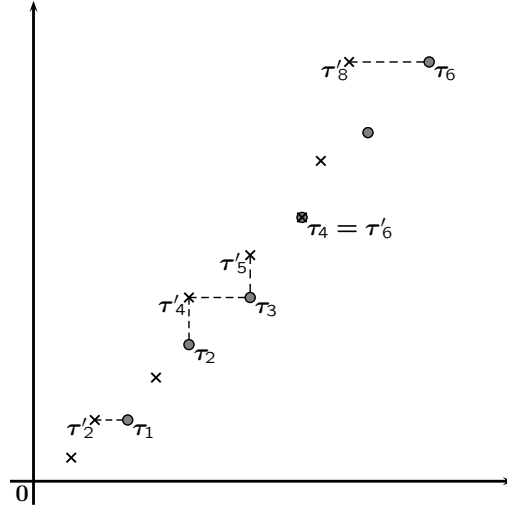


Figure 3.1: Representation of the union of two renewal sets  $\tau, \tau' \subset \mathbb{N}^2$ . It has one double point  $\tau_4 = \tau'_6$ , several isolated points ( $\tau'_1, \tau'_3, \tau'_7$  and  $\tau_5$  in lexical order), and three chains of points  $(\tau'_2, \tau_1)$ ,  $(\tau_2, \tau_4, \tau_3, \tau_5)$  and  $(\tau'_8, \tau_6)$ .

**Proposition 3.3.4.** *Let  $\tau$  and  $\tau'$  be two subsets of  $\mathbb{N}^2$  such that they can be written as sequences  $\tau = \{\tau_1, \tau_2, \dots\}$  with  $\tau_k^{(r)} < \tau_{k+1}^{(r)}$  for every  $r \in \{1, 2\}, k \in \mathbb{N}$  (same for  $\tau'$ ). Then the subset  $\tau \cup \tau'$  can be partitioned like this:*

$$\tau \cup \tau' = \nu \cup \rho \cup \left( \bigcup_{m \in \mathbb{N}} \sigma_m \right), \quad (3.3.18)$$

where  $\nu := \tau \cap \tau'$ ,  $\rho$  is the set of isolated points of  $\tau \cup \tau'$ ,  $\mathfrak{S} := \bigcup_{m \in \mathbb{N}} \sigma_m$  is the set of chained points of  $\tau \cup \tau'$ , and  $\sigma_m, m \in \mathbb{N}$  are chains of points. All the sets  $\nu, \rho$  and  $\sigma_m, m \in \mathbb{N}$  are separated.

Moreover, if  $\mathbf{i}, \mathbf{j} \in \tau \cup \tau', \mathbf{i} \neq \mathbf{j}$  and  $\mathbf{i} \leftrightarrow \mathbf{j}$ , then there exists  $m \in \mathbb{N}$  such that  $\mathbf{i}, \mathbf{j} \in \sigma_m$ .

The decomposition in (3.3.18) comes from the construction above. The latter claim is obvious: if  $\mathbf{i} \leftrightarrow \mathbf{j}$  and  $\mathbf{i} \neq \mathbf{j}$ , then neither  $\mathbf{i}$  or  $\mathbf{j}$  can be a double point or isolated. So they are chained points, and they must be in the same chain (because of the last property of chains). This claim also ensures that disorder values  $(\omega_i)_{i \in \tau \cup \tau'}$  are independent from one set of the partition to another. We now have all the needed tools to prove Proposition 3.3.3.

*Proof of Proposition 3.3.3.* We partition  $(\tau \cup \tau') \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket$  into  $\nu_n, \rho_n, \sigma_{n,m}, m \in \mathbb{N}$  using Proposition 3.3.4, and we use it to separate the right hand side of (3.3.17) in independent



products:

$$\begin{aligned} \mathbb{E} \left[ \left( Z_{n,0}^{\beta, \text{q,free}} \right)^2 \right] &= \mathbf{E}_{(\tau, \tau')} \mathbb{E} \left[ \prod_{i \in \nu_n} e^{2(\beta\omega_i - \lambda(\beta))} \prod_{i \in \rho_n} e^{\beta\omega_i - \lambda(\beta)} \prod_{m \in \mathbb{N}} \prod_{i \in \sigma_{n,m}} e^{\beta\omega_i - \lambda(\beta)} \right] \\ &= \mathbf{E}_{(\tau, \tau')} \left[ \mathbb{E} \left[ \prod_{i \in \nu_n} e^{2(\beta\omega_i - \lambda(\beta))} \right] \mathbb{E} \left[ \prod_{i \in \rho_n} e^{\beta\omega_i - \lambda(\beta)} \right] \prod_{m \in \mathbb{N}} \mathbb{E} \left[ \prod_{i \in \sigma_{n,m}} e^{\beta\omega_i - \lambda(\beta)} \right] \right]. \end{aligned} \quad (3.3.19)$$

We used that  $1_{\{i \in \tau\}} + 1_{\{i \in \tau'\}} = 2$  if and only if  $i \in \tau \cap \tau'$ , and  $1_{\{i \in \tau\}} + 1_{\{i \in \tau'\}} = 1$  for any other  $i \in \tau \cup \tau' \setminus \tau \cap \tau'$ . Note that an isolated point in  $(\tau \cup \tau') \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket$  can be chained for a higher  $n$ , thus we write the dependance on  $n$  of the decomposition in (3.3.19) with a subscript.

Now we have to compute those expectations. Note that  $(\omega_i)_{i \in \nu_n}$  is an independent family, and so is  $(\omega_i)_{i \in \rho_n}$ . Thus  $\mathbb{E} \left[ \prod_{i \in \rho_n} e^{\beta\omega_i - \lambda(\beta)} \right] = \prod_{i \in \rho_n} \mathbb{E} [e^{\beta\omega_i - \lambda(\beta)}] = 1$  (recall that  $e^{\lambda(\beta)} = \mathbb{E} [e^{\beta\omega_1}]$  for any  $\beta \in [0, \beta_0)$ ), and  $\mathbb{E} \left[ \prod_{i \in \nu_n} e^{2(\beta\omega_i - \lambda(\beta))} \right] = e^{|\nu_n|(\lambda(2\beta) - 2\lambda(\beta))}$ .

Let  $\sigma = \{\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_k\}$ ,  $k \in \mathbb{N}$  be a (finite) chain of  $(\tau \cup \tau') \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket$  as defined previously, and let us estimate  $\mathbb{E} \left[ \prod_{i \in \sigma} e^{\beta\omega_i - \lambda(\beta)} \right]$ .

**General case** This expectation cannot be computed in the general case, but we can obtain an upper bound with Cauchy-Schwarz inequality. In Remark 3.3.5 we will discuss why we think this upper bound is not sharp and can be improved, but it is fully sufficient to prove disorder irrelevance.

Let us note  $\sigma^{\text{odd}} = \{\mathbf{i}_1, \mathbf{i}_3, \dots\}$  and  $\sigma^{\text{even}} = \{\mathbf{i}_2, \mathbf{i}_4, \dots\}$ . Notice that  $\sigma^{\text{odd}} \subset \tau$  and  $\sigma^{\text{even}} \subset \tau'$  (or the other way around), and recall that  $\tau$  and  $\tau'$  are strictly increasing on each coordinate. Hence  $(\omega_i)_{i \in \sigma^{\text{even}}}$  is a family of independent variables, and so is  $(\omega_i)_{i \in \sigma^{\text{odd}}}$ . Applying Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathbb{E} \left[ \prod_{i \in \sigma} e^{\beta\omega_i - \lambda(\beta)} \right] &\leq \mathbb{E} \left[ \prod_{i \in \sigma^{\text{odd}}} e^{2(\beta\omega_i - \lambda(\beta))} \right]^{\frac{1}{2}} \mathbb{E} \left[ \prod_{i \in \sigma^{\text{even}}} e^{2(\beta\omega_i - \lambda(\beta))} \right]^{\frac{1}{2}} \\ &\leq \prod_{i \in \sigma^{\text{odd}}} \mathbb{E} \left[ e^{2(\beta\omega_i - \lambda(\beta))} \right]^{\frac{1}{2}} \prod_{i \in \sigma^{\text{even}}} \mathbb{E} \left[ e^{2(\beta\omega_i - \lambda(\beta))} \right]^{\frac{1}{2}} \\ &\leq e^{(\lambda(2\beta) - 2\lambda(\beta)) \frac{|\sigma^{\text{odd}}| + |\sigma^{\text{even}}|}{2}} = e^{(\lambda(2\beta) - 2\lambda(\beta)) \frac{|\sigma|}{2}}, \end{aligned} \quad (3.3.20)$$

where we recall the definition of  $\lambda$  (see (3.1.2)), and that those expectations are finite if  $\beta < \beta_0/2$ . Therefore we obtain the following upper bound on (3.3.19):

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta, \text{q,free}} \right)^2 \right] \leq \mathbf{E}_{(\tau, \tau')} \left[ e^{(\lambda(2\beta) - 2\lambda(\beta)) \left( |\nu_n| + \frac{|\mathfrak{S}_n|}{2} \right)} \right]. \quad (3.3.21)$$

(Recall that  $\mathfrak{S}_n := \bigcup_{m \in \mathbb{N}} \sigma_{n,m}$ ). In particular this proves that  $Z_{n,0}^{\beta, \text{q,free}} \in L^2(\mathbb{P})$  for any  $n \in \mathbb{N}$  and  $\beta < \beta_0/2$ .

Let us note  $\tau_{\preccurlyeq n} := \tau \cap [\mathbf{1}, \mathbf{n}]$  (resp.  $\tau'_{\preccurlyeq n} := \tau' \cap [\mathbf{1}, \mathbf{n}]$ ). Using the decomposition of  $\tau_{\preccurlyeq n} \cup \tau'_{\preccurlyeq n}$  from Proposition 3.3.4, one can prove by induction:

$$\frac{1}{2} (|\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}|) \leq |\nu_n| + |\mathfrak{S}_n| \leq 2 (|\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}|). \quad (3.3.22)$$

(The constants are optimal: the left hand size is an equality if  $|\nu_n| > 0$ ,  $\mathfrak{S}_n = \emptyset$ , and the right hand size is an equality if  $\nu_n = \emptyset$  and  $\mathfrak{S}_n$  only contains chains of length 2). So (3.3.21) becomes

$$\begin{aligned} \mathbb{E} \left[ \left( Z_{n,0}^{\beta, \text{q,free}} \right)^2 \right] &\leq \mathbf{E}_{(\tau, \tau')} \left[ e^{\left( \lambda(2\beta) - 2\lambda(\beta) \right) \left( \frac{|\nu_n|}{2} + |\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}| \right)} \right] \\ &\leq \mathbf{E}_{(\tau, \tau')} \left[ e^{\frac{3}{2} \left( \lambda(2\beta) - 2\lambda(\beta) \right) \left( |\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}| \right)} \right], \end{aligned} \quad (3.3.23)$$

where we used  $|\nu_n| \leq |\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}|$ , and  $\lambda(2\beta) - 2\lambda(\beta) \geq 0$  for any  $\beta \geq 0$ . Moreover  $|\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}| + |\tau_{\preccurlyeq n}^{(2)} \cap \tau'_{\preccurlyeq n}{}^{(2)}|$  is non-decreasing in  $\mathbf{n}$ , and converges to  $|\tau^{(1)} \cap \tau'^{(1)}| + |\tau^{(2)} \cap \tau'^{(2)}|$ . Using the monotone convergence theorem, this finishes the proof of (3.3.15).

**Case**  $\omega \sim \mathbb{P}_{\pm x}$ ,  $x > 0$  In that case we can compute  $\mathbb{E}_{\pm x} \left[ \prod_{i \in \sigma} e^{\beta \omega_i - \lambda(\beta)} \right]$  exactly for any finite chain  $\sigma = \{\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_k\}$ ,  $k \in \mathbb{N}$ . Indeed, we can write without loss of generality

$$\omega_{\mathbf{i}_1} = \widehat{\omega}_1 \bar{\omega}_1, \omega_{\mathbf{i}_2} = \widehat{\omega}_1 \bar{\omega}_2, \omega_{\mathbf{i}_3} = \widehat{\omega}_2 \bar{\omega}_2, \dots \quad (3.3.24)$$

and so on, with  $(\widehat{\omega}_1, \bar{\omega}_1, \widehat{\omega}_2, \dots)$  a family of i.i.d. random variables. Then we can compute

$$\begin{aligned} \mathbb{E}_{\pm x} \left[ \prod_{i \in \sigma} e^{\beta \omega_i - \lambda(\beta)} \right] &= e^{-k\lambda(\beta)} \mathbb{E}_{\pm x} \left[ \mathbb{E}_{\pm x} \left[ e^{\beta \widehat{\omega}_1 \bar{\omega}_1} \mid (\widehat{\omega}_j)_{j \geq 1}, (\bar{\omega}_j)_{j \geq 2} \right] \prod_{l=2}^k e^{\beta \omega_{\mathbf{i}_l}} \right] \\ &= e^{-k\lambda(\beta)} \mathbb{E}_{\pm x} \left[ \frac{e^{\beta x \widehat{\omega}_1} + e^{-\beta x \widehat{\omega}_1}}{2} \prod_{l=2}^k e^{\beta \omega_{\mathbf{i}_l}} \right] = e^{-(k-1)\lambda(\beta)} \mathbb{E}_{\pm x} \left[ \prod_{l=2}^k e^{\beta \omega_{\mathbf{i}_l}} \right], \end{aligned}$$

where we used that  $\frac{1}{2}(e^{\beta x \widehat{\omega}_1} + e^{-\beta x \widehat{\omega}_1}) = \cosh(\beta x \widehat{\omega}_1) = \cosh(\beta x^2) = e^{\lambda(\beta)}$  (because  $|\widehat{\omega}_1| = x$   $\mathbb{P}_{\pm x}$ -a.s. and  $\cosh$  is symmetric). By iteration, we finally obtain for any finite chain  $\sigma$  and  $\beta \in \mathbb{R}_+$ ,

$$\mathbb{E}_{\pm x} \left[ \prod_{i \in \sigma} e^{\beta \omega_i - \lambda(\beta)} \right] = 1. \quad (3.3.25)$$

Recalling (3.3.19), we get the exact formula:

$$\mathbb{E}_{\pm x} \left[ \left( Z_{n,0}^{\beta, \text{q,free}} \right)^2 \right] = \mathbf{E}_{(\tau, \tau')} \left[ e^{\left( \lambda(2\beta) - 2\lambda(\beta) \right) |\nu_n|} \right]. \quad (3.3.26)$$

And we obtain (3.3.16) by applying the monotone convergence theorem.  $\square$

Notice that the particular behavior of distributions  $\mathbb{P}_{\pm x}$  relies solely on (3.3.25). When  $\sigma$  is a chain of length two, it matches the correlations of the field  $(e^{\beta \omega_i})_{i \in \mathbb{N}^2}$ , as discussed in Section 3.2.3.

### 3.3.3 Upper bounds for the shift of the critical point

When we cannot bound the second moment of the partition function for all  $n \in \mathbb{N}$ , we can still estimate  $n_\beta$  with the finite volume equivalent of Proposition 3.3.3 (that is inequality (3.3.23) when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ , and identity (3.3.26) when  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ ): thanks to Proposition 3.3.1, we are able to obtain an upper bound for the shift of the critical point.

#### Case $\mathbb{P} \neq \mathbb{P}_{\pm x}$ for all $x > 0$ : Proposition 3.2.4

Let us consider (3.3.23): we can split the two intersections of the projections with a Cauchy-Schwarz inequality

$$\mathbb{E} \left[ \left( Z_{\mathbf{n},0}^{\beta,\mathfrak{q},free} \right)^2 \right] \leq \mathbf{E}_{(\tau,\tau')} \left[ e^{3(\lambda(2\beta) - 2\lambda(\beta))|\tau_{\leq n}^{(1)} \cap \tau'_{\leq n}|} \right], \quad (3.3.27)$$

where we also used  $\mathbf{n} = (n, n)$  and that both projections have the same law. Recall that  $\tau^{(1)}, \tau'^{(1)}$  are independent univariate renewal processes with inter-arrival distribution  $\mathbf{P}(\tau^{(r)} = a) = \tilde{L}(a)a^{-(1+\alpha)}$ , and recall that  $\lambda(2\beta) - 2\lambda(\beta) \sim c\beta^2$  as  $\beta \searrow 0$  for some  $c > 0$ . In particular this upper bound has already been studied in [6, 147] for  $\alpha \in (1/2, 1)$ , and gives the estimate  $n_\beta \geq L_3(1/\beta)\beta^{\frac{2}{2\alpha-1}}$  (we do not write the details here). In the case  $\alpha \geq 1$ , one easily gets from (3.3.27) that the second moment is bounded by  $\exp(c\beta^2 n)$ , giving the estimate  $n_\beta \geq c'\beta^{-2}$ . Recollecting (3.3.4), this fully proves Proposition 3.2.4 (notice that when  $m_1 = 0$  and  $\alpha \geq 1$ , this does not lead to a better bound for the critical point than the trivial one  $h_c(\beta) \leq \lambda(\beta) = \beta m_1^2 + \beta^2 m_2^2 / 2(1 + o(1))$ ).

**Remark 3.3.5.** As far as Theorem 3.2.3 is concerned, Proposition 3.2.4 yields a satisfactory upper bound when  $m_1 \neq 0$ , but not when  $m_1 = 0, m_4 > m_2^2$ . We strongly believe that our lower bound is of the right order (i.e.  $\beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)}$  when  $m_1 = 0, m_4 > m_2^2$ ), and that our upper bound is too rough when we estimate the second moment—even though it is sufficient to prove disorder irrelevance and give a satisfactory upper bound when  $m_1 \neq 0$ . Furthermore, a work in progress with Q. Berger [21]<sup>†</sup> is leading us to the following claim:

*Suppose  $\alpha \in (1/2, 1)$ ,  $m_1 = 0$  and  $m_4 > m_2^2$ . Let  $\beta > 0$  and  $\mathbf{m}_\beta := (m_\beta, m_\beta)$  with  $m_\beta \sim \beta^{\frac{-4}{2\alpha-1}} L(1/\beta)^{\frac{2}{2\alpha-1}}$  as  $\beta \searrow 0$ . Then  $Z_{\mathbf{m}_\beta,0}^{\beta,\mathfrak{q},free}$  converges in distribution to some non-trivial random variable  $\mathbf{Z}^{free}$  as  $\beta \searrow 0$ . This convergence also holds in  $L^2$  under some appropriate coupling.*

In particular this claim implies that  $n_\beta \sim L_4(1/\beta)\beta^{-4/(2\alpha-1)}$  when  $\alpha \in (1/2, 1)$ : once this result is proven, it will directly give an upper bound  $h_c(\beta) \leq L_5(1/\beta)\beta^{4\alpha/(2\alpha-1)}$ , which is fully satisfactory with regards to Theorem 3.2.3 in the case  $m_1 = 0, m_4 > m_2^2$ .

<sup>†</sup>. This work is displayed in Chapter 4.

Moreover, when  $\widehat{\omega}, \bar{\omega}$  are two sequences of centered, i.i.d. Gaussian variables, one can fully compute the contribution of chains of points to the second moment of the partition function, which eventually leads to (for  $\beta$  sufficiently small)

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta,q,free} \right)^2 \right] \leq 2 \mathbb{E}_{(\tau,\tau')} \left[ e^{C_{10}\beta^4 |\tau_{\preccurlyeq n}^{(1)} \cap \tau'_{\preccurlyeq n}{}^{(1)}|} \right]. \quad (3.3.28)$$

This is the same as (3.3.27) with an exponent of order  $\beta^4$  instead of  $\lambda(2\beta) - 2\lambda(\beta) \sim c\beta^2$ . In particular it gives the expected upper bound  $h_c(\beta) \leq L_6(1/\beta)\beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)}$  on the shift of the critical point, supporting again the idea that Theorem 3.2.3 gives the correct order of the lower bound. We carry out the full computation in Appendix 3.C.

**Case  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$**

In that case we have an exact computation of the second moment (3.3.26), where the right hand side is the same as in the proof of [18, Prop. 3.3] for the gPS model with i.i.d. disorder. Thus we obtain the same estimate as in [18], that is  $n_\beta \sim L_7(1/\beta)\beta^{-\max(\frac{2\alpha}{\alpha-1}, 4)}$  for any  $\alpha > 1$  (see the proof of [18, Prop. 3.3] for the details). Therefore Proposition 3.3.1 immediately gives an upper bound on  $h_c(\beta)$  of order  $L_8(1/\beta)\beta^{\max(\frac{2\alpha}{\alpha-1}, 4)}$  which proves the right inequality in Theorem 3.2.5.

### 3.4 Disorder relevance: smoothing of the phase transition when $\mathbb{P} \neq \mathbb{P}_{\pm x}$

The inequality of Theorem 3.2.7 is proven via a rare-stretch strategy, as done in [90] (or more recently [43]). We introduce some notations that we use in this section and the next one to lighten upcoming formulae: we write  $Z_n^q := Z_{n,h}^{\beta,q}$  for the (constrained) partition function with quenched disorder (the choice of parameters  $\beta$  and  $h$  will always be explicit), and  $Z_{n,h} := Z_{n,h}^{0,q}$  for the homogeneous (or annealed) partition function. Moreover we will denote by  $Z_{a,b}^q$  the partition function conditioned to start from  $\mathbf{a}$  and constrained to end in  $\mathbf{b}$ . More precisely for any  $\mathbf{0} \preccurlyeq \mathbf{a} \prec \mathbf{b}$ ,

$$Z_{a,b}^q := \mathbb{E} \left[ \exp \left( \sum_{i \in [\mathbf{a}+1, \mathbf{b}]} (\beta\omega_i - \lambda(\beta) + h) 1_{\{i \in \tau\}} \right) 1_{\{\mathbf{b} \in \tau\}} \middle| \mathbf{a} \in \tau \right], \quad (3.4.1)$$

and  $Z_{a,b}^q := 0$  if  $\mathbf{a} \not\prec \mathbf{b}$ . Note that  $Z_{a,b}^q$  has same law as  $Z_{b-a}^q$ , and for any  $\mathbf{a} \preccurlyeq \mathbf{b} \preccurlyeq \mathbf{c} \preccurlyeq \mathbf{d}$ ,  $Z_{a,b}^q$  and  $Z_{c,d}^q$  are independent.

#### 3.4.1 Rare-stretch strategy

The rare-stretch strategy consists in obtaining a lower bound on the partition function by considering the contribution of only one type of trajectories, which target favorable (but sparse) regions in the environment.

Fix  $\beta > 0$  and  $h \in \mathbb{R}$ , and let  $\mathcal{A}_l \subset \mathbb{R}^{l^2}$ ,  $l \in \mathbb{N}$  be a sequence of Borel sets. We will write  $\mathbf{l} := (l, l)$ , and assume that there is  $\mathcal{G} \geq 0$ ,  $\mathcal{C} \geq 0$  such that for any  $l \in \mathbb{N}$  (or at least infinitely many  $l \in \mathbb{N}$ ):

$$\begin{aligned} & \bullet \frac{1}{l} \log Z_{\mathbf{l}}^q \geq \mathcal{G}, \quad \text{for any } \omega \text{ such that } \omega_{\llbracket \mathbf{l}, \mathbf{l} \rrbracket} := (\omega_i)_{i \in \llbracket \mathbf{l}, \mathbf{l} \rrbracket} \in \mathcal{A}_l, \\ & \bullet \frac{1}{l} \log \mathbb{P}(\omega_{\llbracket \mathbf{l}, \mathbf{l} \rrbracket} \in \mathcal{A}_l) \geq -\mathcal{C}. \end{aligned} \tag{3.4.2}$$

Here  $\mathcal{G}$  stands for *gain*, and  $\mathcal{C}$  for *cost*.

**Lemma 3.4.1.** *For any  $\beta > 0$  and  $h \in \mathbb{R}$ , if  $\mathcal{G}$ ,  $\mathcal{C}$  are defined as in (3.4.2) (for some sequence of  $(\mathcal{A}_l)_{l \in \mathbb{N}}$ ), then the following holds:*

$$\mathcal{G} - (2 + \alpha)\mathcal{C} > 0 \implies \mathbb{F}(\beta, h) > 0. \tag{3.4.3}$$

*Proof.* We replicate here the proof of [43], but with our disorder indexed by  $\mathbb{N}^2$ . Fix  $l$  such that the above conditions hold, and let  $T_i(\omega)$ ,  $i \in \mathbb{N}$  be the indices of blocks of size  $l \times l$  on the diagonal satisfying the event  $\mathcal{A}_l$ . More precisely, let  $T_0(\omega) := 0$ , and

$$T_i(\omega) := \inf\{n > T_{i-1}(\omega) ; \omega_{\llbracket (n-1)l+1, nl \rrbracket} \in \mathcal{A}_l\}, \tag{3.4.4}$$

for any  $i \geq 1$ . Note that  $(T_{i+1}(\omega) - T_i(\omega))_{i \in \mathbb{N}}$  is an i.i.d. sequence with law  $\text{Geom}(\mathbb{P}(\mathcal{A}_l))$ , so  $\mathbb{E}[T_1] = \mathbb{P}(\mathcal{A}_l)^{-1} \leq e^{cl}$ . We can give a lower bound of the partition function on the block  $T_k \mathbf{l}$ ,

$$Z_{T_k \mathbf{l}}^q \geq \prod_{i=1}^k Z_{T_{i-1} \mathbf{l}, (T_i-1) \mathbf{l}}^q Z_{(T_i-1) \mathbf{l}, T_i \mathbf{l}}^q. \tag{3.4.5}$$

Note that  $\omega_{\llbracket (T_i-1)l, T_i l \rrbracket} \in \mathcal{A}_l$  for all  $i$  by definition of  $T_i$ , so that  $Z_{T_{i-1} \mathbf{l}, (T_i-1) \mathbf{l}}^q \geq e^{gl}$  by definition of  $\mathcal{G}$ . We also have the obvious bound  $Z_{\mathbf{a}, \mathbf{b}}^q \geq K(\|\mathbf{b} - \mathbf{a}\|) e^{\beta \omega_{\mathbf{b}} - \lambda(\beta)h}$  for any  $\mathbf{a} \prec \mathbf{b}$ . Therefore we have

$$\begin{aligned} Z_{T_k \mathbf{l}}^q & \geq \prod_{i=1}^k \left( e^{\beta \omega_{(T_i-1)l} - \lambda(\beta)h} K(\|(T_i - 1 - T_{i-1}) \mathbf{l}\|) \times e^{lg} \right) \\ & \geq e^{kl\mathcal{G}} \prod_{i=1}^k \frac{c_{\alpha_+} e^{\beta \omega_{(T_i-1)l} - \lambda(\beta)h}}{(l(T_i - T_{i-1}))^{2+\alpha_+}}, \end{aligned} \tag{3.4.6}$$

where the last inequality holds for any  $\alpha_+ > \alpha$  and a convenient  $c_{\alpha_+} > 0$ , by Potter's bound (cf. [30]). We can now estimate from below the free energy, using the strong law of large

numbers:

$$\begin{aligned}
 F(\beta, h) &= \lim_{k \rightarrow \infty} \frac{1}{l T_k} \log Z_{T_k l}^q \\
 &\geq \lim_{k \rightarrow \infty} \frac{k}{T_k} \frac{1}{kl} \left( kl \mathcal{G} + \sum_{i=1}^k \left[ c_{\alpha_+, \beta, h} + \beta \omega_{(T_i - T_{i-1})\mathbf{1}} - (2 + \alpha_+) \log \left( l(T_i - T_{i-1}) \right) \right] \right) \\
 &\geq \frac{1}{\mathbb{E}[T_1]} \left( \mathcal{G} + \frac{c_{\alpha_+, \beta, h} + \beta \mathbb{E}[\omega_1]}{l} - (2 + \alpha_+) \frac{\log l + \mathbb{E}[\log(T_1)]}{l} \right) \\
 &\geq e^{-cl} \left( \mathcal{G} - (2 + \alpha_+) \mathcal{C} + \frac{c'_{\alpha_+, \beta, h} - (2 + \alpha_+) \log l}{l} \right),
 \end{aligned} \tag{3.4.7}$$

where we set  $c_{\alpha_+, \beta, h} := \log c_{\alpha_+} - \lambda(\beta) + h$ . For the last inequality, we used Jensen's inequality to get that  $\mathbb{E}[\log(T_1)] \leq \log(1/\mathbb{P}(\mathcal{A}_l)) \leq Cl$ . Finally, if  $\mathcal{G} - (2 + \alpha)\mathcal{C} > 0$ , then it also holds for some  $\alpha_+ > \alpha$ , and the right hand side is strictly positive for  $l$  large enough, which implies  $F(\beta, h) > 0$  and concludes the proof.  $\square$

### 3.4.2 Smoothing of the phase transition in $\beta$

Let us discuss the strategy of the proof first. For the PS model (with i.i.d. disorder), the method used in [90] to prove a smoothing is to fix  $h = h_c(\beta)$ , so that  $F(\beta, h_c(\beta)) = 0$  and Lemma 3.4.1 implies  $\mathcal{G} \leq (2 + \alpha)\mathcal{C}$ . Then one chooses a gain  $\mathcal{G}$  close to  $F(\beta, h_c(\beta) + u)$ —which matches the free energy of the model with a *shifted* disorder, *i.e.* with  $\omega$  replaced by  $\omega + u/\beta$ —and expresses the corresponding cost  $\mathcal{C}$  with the cost of the change of measure from  $\omega$  to  $\omega + u/\beta$  (which can be estimated via a relative entropy inequality), therefore obtaining an upper bound on the free energy near the critical point.

However this method doesn't apply well to our model, mostly because of the dimension of the field  $\omega$  (this is also discussed in [18] for an i.i.d. disorder). A direct shift of the disorder field  $\omega$  is too costly (we shift  $n^2$  variables in a model of size  $n$ ). On the other hand an i.i.d. shift of the sequences  $\hat{\omega}$  and  $\bar{\omega}$  by  $u/\sqrt{\beta}$ —although involving only  $2n$  variables—is not easily related to a free energy with different parameters. Therefore, we needed to adapt this method. We first prove a smoothing inequality *with respect to  $\beta$*  instead of  $h$ , using a *dilation* of the disorder instead of a shift, *i.e.* we replace the sequence  $\hat{\omega}$  (resp.  $\bar{\omega}$ ) by  $(1 + \delta)\hat{\omega}$  (resp.  $(1 + \delta)\bar{\omega}$ ). This change of measure matches the same model with disorder intensity  $\beta(1 + \delta)^2 \approx \beta(1 + 2\delta)$  instead of  $\beta$ , and is not too costly (we change the law of  $2n$  variables in a system of size  $n$ ).

Let us introduce the “shifted” free energy  $\tilde{F}(\beta, h) := F(\beta, h + \lambda(\beta))$  (*i.e.* if we omit the term  $-\lambda(\beta)$  in the definition of the partition function): in view of Proposition 3.1.1, we get that  $(\beta, h) \mapsto \tilde{F}(\beta, h)$  is convex, so the critical line  $\beta \mapsto \tilde{h}_c(\beta) := \inf\{h: \tilde{F}(\beta, h) > 0\} = h_c(\beta) - \lambda(\beta)$  is concave. Actually it is decreasing and continuous (recall that the upper bound in (3.1.13) is strict for  $\beta > 0$ ), so one can consider the inverse map of  $\beta \mapsto \tilde{h}_c(\beta)$ , that we denote  $h \mapsto \tilde{\beta}_c(h)$ : for each  $h > 0$ , the value  $\tilde{\beta}_c(h)$  is the critical value for the map  $\beta \mapsto \tilde{F}(\beta, h)$  corresponding to the localization transition. One can therefore consider the

transition as  $\beta$  varies, and the next proposition tells this phase transition is smooth, *i.e.* for fixed  $h$ , the growth of  $\beta \mapsto \tilde{F}(\beta, h)$  close to  $\tilde{\beta}_c(h)$  is at most quadratic.

**Proposition 3.4.2.** *Under Assumption 3.2.6, for any  $h > 0$ , there exists  $c_h > 0$  such that for any  $\delta \in (0, 1)$ , one has*

$$\tilde{F}\left((1 + \delta)\tilde{\beta}_c(h), h\right) \leq c_h \delta^2. \quad (3.4.8)$$

*Proof.* For any  $\delta > 0$ , we define  $\tilde{\mathbb{P}}_{l,\delta}$  the law of the disorder in  $[[\mathbf{1}, l]]$  *dilated* by  $(1 + \delta)$  on each coordinate (*i.e.*  $\hat{\omega}_{[[\mathbf{1}, l]]}$  is replaced with  $(1 + \delta)\hat{\omega}_{[[\mathbf{1}, l]]}$ , same for  $\bar{\omega}_{[[\mathbf{1}, l]]}$ ). We denote  $\tilde{\mathbb{P}}_\delta$  the infinite product law. Note that  $H(\tilde{\mathbb{P}}_{l,\delta}|\mathbb{P}) = lH(\tilde{\mathbb{P}}_{1,\delta}|\mathbb{P})$ , and that there is some  $c > 0$  such that  $H(\tilde{\mathbb{P}}_{1,\delta}|\mathbb{P}) \leq c\delta^2$  by Assumption 3.2.6. For any  $\beta > 0$ ,  $h \in \mathbb{R}$  and  $l \in \mathbb{N}$ , let us define

$$\mathcal{A}_l := \left\{ \omega_{[[\mathbf{1}, l]]} \in \mathbb{R}^{l^2} ; \frac{1}{l} \log \tilde{Z}_l^q \geq \frac{1}{2} \tilde{F}(\beta(1 + \delta)^2, h) \right\}, \quad (3.4.9)$$

where  $\tilde{Z}_l^q := Z_{l, h + \lambda(\beta)}^{\beta, q}$ , so that  $\lim_{l \rightarrow +\infty} \frac{1}{l} \log \tilde{Z}_l^q := \tilde{F}(\beta, h)$ .

The set  $\mathcal{A}_l$  is chosen so that the *gain* (as defined for Lemma 3.4.1) is  $\mathcal{G} = \frac{1}{2} \tilde{F}(\beta(1 + \delta)^2, h)$ , where  $\tilde{F}(\beta(1 + \delta)^2, h)$  is exactly the free energy of the gPS model with partition function  $\tilde{Z}_l^q$ , where we changed the disorder law from  $\mathbb{P}$  to  $\tilde{\mathbb{P}}_\delta$  (because multiplying  $\hat{\omega}$  and  $\bar{\omega}$  by  $(1 + \delta)$  is the same thing as multiplying  $\beta$  by  $(1 + \delta)^2$ ). Thus  $\frac{1}{l} \log \tilde{Z}_l^q$  converges  $\tilde{\mathbb{P}}_\delta$ -a.s. to  $F(\beta(1 + \delta)^2, h)$  when  $l \rightarrow \infty$ , so we obviously have that

$$\tilde{\mathbb{P}}_{l,\delta}(\mathcal{A}_l) \xrightarrow[l \rightarrow \infty]{} 1. \quad (3.4.10)$$

Now we can estimate  $\mathbb{P}(\mathcal{A}_l)$  via a standard relative entropy inequality, which gives

$$\mathbb{P}(\mathcal{A}_l) \geq \tilde{\mathbb{P}}_{l,\delta}(\mathcal{A}_l) \exp\left(-\frac{1}{\tilde{\mathbb{P}}_{l,\delta}(\mathcal{A}_l)} \left(H(\tilde{\mathbb{P}}_{l,\delta}|\mathbb{P}) + e^{-1}\right)\right). \quad (3.4.11)$$

We therefore get that  $\mathbb{P}(\mathcal{A}_l) \geq \frac{e^{-2/e}}{2} e^{-2c\delta^2 l}$ , for  $l$  large enough (so that  $\tilde{\mathbb{P}}_{l,\delta}(\mathcal{A}_l) \geq 1/2$ ). Therefore, for  $l$  large (how large depends on  $\delta$ ), we get that

$$\frac{1}{l} \log \mathbb{P}(\mathcal{A}_l) \geq -3c\delta^2. \quad (3.4.12)$$

Now, for any  $h > 0$ , fix  $\beta = \tilde{\beta}_c(h)$ , so that  $h = \tilde{h}_c(\beta)$ . Then  $\tilde{F}(\beta, h_c(\beta)) = \tilde{F}(\tilde{\beta}_c(h), h) = 0$ , so Lemma 3.4.1 implies  $\mathcal{G} - (2 + \alpha)\mathcal{C} \leq 0$ , thereby

$$\frac{1}{2} \tilde{F}\left((1 + \delta)^2 \tilde{\beta}_c(h), h\right) - 3c(2 + \alpha)\delta^2 \leq 0, \quad (3.4.13)$$

which gives  $\tilde{F}\left((1 + \delta)^2 \tilde{\beta}_c(h), h\right) \leq c'\delta^2$ . This concludes the proof by simply recalling that  $\tilde{F}$  is non-decreasing in its first coordinate, so  $\tilde{F}\left((1 + \delta)\tilde{\beta}_c(h), h\right) \leq \tilde{F}\left((1 + \delta)^2 \tilde{\beta}_c(h), h\right)$ .  $\square$

### 3.4.3 Conclusion: smoothing of the phase transition in $h$

Once we have the smoothing inequality with respect to  $\beta$  of Proposition 3.4.2, we are able to transcribe it to a smoothing in  $h$  using the convexity properties of  $\tilde{F}(\beta, h)$ . We now conclude the proof of Theorem 3.2.7, thanks to Proposition 3.4.2.

*Proof of Theorem 3.2.7.* We fix  $\beta > 0$ , and we take  $t > 0$  small enough such that  $\tilde{h}_c(\beta) + t < 0$ . Recall that  $\tilde{h}_c(\cdot)$  is a concave, non-increasing, continuous function: there exists  $\beta_t \in (0, \beta)$  such that  $\tilde{h}_c(\beta_t) = \tilde{h}_c(\beta) + t$ . Hence,

$$F(\beta, h_c(\beta) + t) = \tilde{F}(\beta, \tilde{h}_c(\beta) + t) = \tilde{F}(\beta_t + u, \tilde{h}_c(\beta_t)), \quad (3.4.14)$$

where  $u = \beta - \beta_t > 0$ . Using Proposition 3.4.2 and the fact that  $\tilde{\beta}_c(\tilde{h}_c(\beta_t)) = \beta_t$ , we have

$$F(\beta, h_c(\beta) + t) = \tilde{F}(\beta_t(1 + u/\beta_t), \tilde{h}_c(\beta_t)) \leq c_{\tilde{h}_c(\beta_t)}(u/\beta_t)^2. \quad (3.4.15)$$

Note that  $\beta_t \nearrow \beta > 0$  as  $t \searrow 0$ , so the factor  $c_{\tilde{h}_c(\beta_t)}\beta_t^{-2}$  is bounded by a constant that depends only on  $\beta$ , uniformly in  $t$  sufficiently small. Using that  $\tilde{h}_c(\cdot)$  is concave, we also have

$$-\frac{t}{u} = \frac{\tilde{h}_c(\beta) - \tilde{h}_c(\beta_t)}{\beta - \beta_t} \leq \tilde{h}'_c(\beta_t^+) \leq \tilde{h}'_c((\beta/2)^+) < 0, \quad (3.4.16)$$

where  $\tilde{h}'_c(\cdot^+)$  is the right derivative of  $\tilde{h}_c$ , and the second inequality holds for any  $t$  sufficiently small (because  $\tilde{h}'_c(\beta_t^+) < 0$  as soon as  $\beta_t > 0$ , and it decreases as  $t \searrow 0$ ). We deduce that  $t \geq c_\beta u$  for some  $c_\beta > 0$ , and plugging it in (3.4.15), we finally obtain

$$F(\beta, h_c(\beta) + t) \leq c'_\beta t^2, \quad (3.4.17)$$

where  $c'_\beta > 0$  depends only on  $\beta$ . This concludes the proof of Theorem 3.2.7.  $\square$

## 3.5 Disorder relevance: shift of the critical point when

$$\mathbb{P} \neq \mathbb{P}_{\pm x}$$

In this section, we prove Theorem 3.2.3, that is a the lower bound for the shift of the critical point when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ . We will discuss the case  $\mathbb{P} = \mathbb{P}_{\pm x}$  in Section 3.6.

### 3.5.1 Coarse-graining and fractional moment method

Our proof is based on a fractional moment method, introduced in [64] for the original PS model, and slightly adapted to the gPS model with independent disorder in [18]. The first part of the proof (the coarse-graining procedure) is identical to that of [18], but the different estimates are specific to our setting.



Recall the notation  $Z_n^q := Z_{n,h}^{\beta,q}$  from Section 3.4. Let us define for any  $\mathbf{n} \in \mathbb{N}^2$  (note that we do not assume  $\mathbf{n} = (n, n)$  in this section) the *fractional moment* of the partition function:

$$A_n := \mathbb{E}\left[\left(Z_n^q\right)^\eta\right], \quad (3.5.1)$$

where  $\eta \in (0, 1)$  is a constant we will fix later on (notice that  $A_n < \infty$  because  $Z_{n,h}^{\beta,q} \in L^1(\mathbb{P})$ ). For any  $k \in \mathbb{N}$  we set  $\mathbf{k} := (k, k)$ , and for  $\mathbf{n} \succ \mathbf{k}$ , we decompose the partition function  $Z_n^q$  according to the first point  $\mathbf{n} - \mathbf{i}$  of  $\tau$  which lies in the square  $[\mathbf{n} - \mathbf{k} + \mathbf{1}, \mathbf{n}]$  (in particular  $\mathbf{0} \preccurlyeq \mathbf{i} \prec \mathbf{k}$ ), and the point  $\mathbf{n} - \mathbf{i} - \mathbf{j}$  before: in particular it is the last point of  $\tau$  which isn't in the previous square, so it can only be in one of the three boxes  $[[0, \mathbf{n}^{(1)} - k] \times [0, \mathbf{n}^{(2)} - k], [\mathbf{n}^{(1)} - k, \mathbf{n}^{(1)}] \times [0, \mathbf{n}^{(2)} - k]$  or  $[[0, \mathbf{n}^{(1)} - k] \times [\mathbf{n}^{(2)} - k, \mathbf{n}^{(2)}]]$ . This decomposition gives us for any  $\mathbf{n} \succ \mathbf{k}$ ,

$$Z_n^q = Z_n^{q,1} + Z_n^{q,2} + Z_n^{q,3}, \quad (3.5.2)$$

where we write for any  $s \in \{1, 2, 3\}$ ,

$$Z_n^{q,s} := \sum_{(\mathbf{i}, \mathbf{j}) \in D_{\mathbf{k}, \mathbf{n}}^s} Z_{\mathbf{n}-\mathbf{i}-\mathbf{j}}^\omega \times \mathbf{P}(\tau_1 = \mathbf{j}) e^{\beta\omega_{\mathbf{n}-\mathbf{i}} - \lambda(\beta)h} \times Z_{\mathbf{n}-\mathbf{i}, \mathbf{n}}^\omega, \quad (3.5.3)$$

where  $Z_{\mathbf{n}-\mathbf{i}, \mathbf{n}}^\omega$  is defined in (3.4.1), and the sets  $D_{\mathbf{k}, \mathbf{n}}^s$ ,  $s \in \{1, 2, 3\}$  are defined as follow (with  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ ):

$$\begin{aligned} D_{\mathbf{k}, \mathbf{n}}^1 &:= \{(\mathbf{i}, \mathbf{j}) \in \mathbb{N}_0^2 \times \mathbb{N}^2 ; \mathbf{i} \prec \mathbf{k} \preccurlyeq \mathbf{i} + \mathbf{j} \preccurlyeq \mathbf{n}\}, \\ D_{\mathbf{k}, \mathbf{n}}^2 &:= \{(\mathbf{i}, \mathbf{j}) \in \mathbb{N}_0^2 \times \mathbb{N}^2 ; \mathbf{i} \prec \mathbf{k}, \mathbf{i}^{(1)} + \mathbf{j}^{(1)} < k, \mathbf{i}^{(2)} + \mathbf{j}^{(2)} \geq k, \mathbf{i} + \mathbf{j} \preccurlyeq \mathbf{n}\}, \\ D_{\mathbf{k}, \mathbf{n}}^3 &:= \{(\mathbf{i}, \mathbf{j}) \in \mathbb{N}_0^2 \times \mathbb{N}^2 ; \mathbf{i} \prec \mathbf{k}, \mathbf{i}^{(1)} + \mathbf{j}^{(1)} \geq k, \mathbf{i}^{(2)} + \mathbf{j}^{(2)} < k, \mathbf{i} + \mathbf{j} \preccurlyeq \mathbf{n}\}. \end{aligned} \quad (3.5.4)$$

We also define  $D_{\mathbf{k}, \infty}^s = \cup_{\mathbf{n} \in \mathbb{N}^2} D_{\mathbf{k}, \mathbf{n}}^s$  (which means we drop the condition  $\mathbf{i} + \mathbf{j} \preccurlyeq \mathbf{n}$ ). Recall that  $Z_{\mathbf{n}-\mathbf{i}, \mathbf{n}}^q$  and  $Z_{\mathbf{i}}^q$  have same law, and that  $Z_{\mathbf{n}-\mathbf{i}-\mathbf{j}}^q$ ,  $e^{\beta\omega_{\mathbf{n}-\mathbf{i}} - \lambda(\beta)h}$  and  $Z_{\mathbf{n}-\mathbf{i}, \mathbf{n}}^q$  are independent. Using this together with the definition of  $A_n$  and the standard inequality  $(\sum_i a_i)^\eta \leq \sum_i a_i^\eta$  for any  $\eta \in (0, 1)$  and  $a_i \geq 0, i \in \mathbb{N}$ , we obtain

$$A_n \leq A_n^1 + A_n^2 + A_n^3, \quad (3.5.5)$$

where for any  $s \in \{1, 2, 3\}$ ,

$$A_n^s := c_{\beta, h, \eta} \sum_{(\mathbf{i}, \mathbf{j}) \in D_{\mathbf{k}, \mathbf{n}}^s} A_{\mathbf{n}-\mathbf{i}-\mathbf{j}} \times K(\|\mathbf{j}\|)^\eta \times A_{\mathbf{i}}, \quad (3.5.6)$$

and  $c_{\beta, h, \eta} := \mathbb{E}[e^{(\beta\omega_1 - \lambda(\beta) + h)\eta}] = e^{\lambda(\eta\beta) - \eta\lambda(\beta) + \eta h}$ . Note that  $c_{\beta, h, \eta}$  is uniformly bounded for  $\beta, h$  small and  $0 \leq \eta \leq 1$ , so we can bound it by a constant  $C_1$ .

As in [18], the proof of Theorem 3.2.3 relies on the following claim.

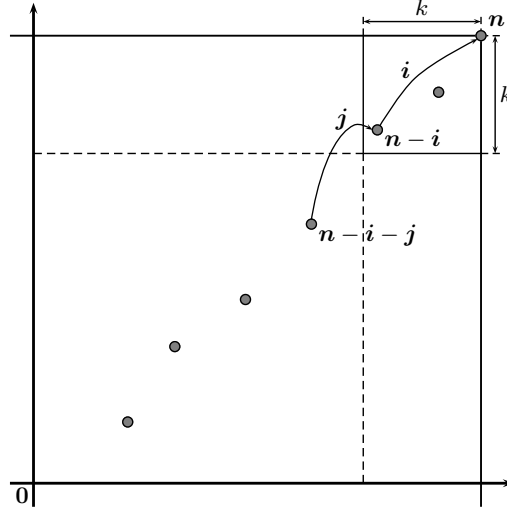


Figure 3.1: Decomposition of the partition function for the coarse-graining procedure.  $n - i$  denotes the first renewal epoch in the square  $\llbracket n - k + 1, n \rrbracket$ , and  $n - i - j$  is the renewal epoch before: it is either in the bottom left rectangle (e.g. the representation in the figure; those are the trajectories contributing to  $Z_n^{q,1}$ ), bottom right ( $Z_n^{q,2}$ ) or top left ( $Z_n^{q,3}$ ).

**Lemma 3.5.1.** For fixed  $\beta > 0$  and  $h \in \mathbb{R}$ , if there exist  $\eta \in (0, 1)$  and  $k \in \mathbb{N}$  such that  $\rho_1 + \rho_2 + \rho_3 \leq 1$ , where

$$\rho_s := C_1 \sum_{(i,j) \in D_{k,\infty}^s} K(\|j\|)^\eta A_i, \quad s \in \{1, 2, 3\}, \quad (3.5.7)$$

then  $F(\beta, h) = 0$ .

*Proof.* The proof is straightforward. Define  $\bar{A} := \sup\{A_i, i^{(1)} < k \text{ or } i^{(2)} < k\}$ . By Jensen's inequality, one obviously has  $A_i \leq \mathbb{E}[Z_i^q]^\eta \leq e^{\eta h \min(i^{(1)}, i^{(2)})}$ , because there are at most  $\min(i^{(1)}, i^{(2)})$  renewal points in  $\llbracket 1, i \rrbracket$ . Thus we have  $\bar{A} \leq e^{\eta h k}$ . Using the decomposition (3.5.5), (3.5.6) of  $A_n$  and  $\rho_1 + \rho_2 + \rho_3 \leq 1$ , we deduce (by induction)  $A_n \leq \bar{A}$  for any  $n \in \mathbb{N}^2$ . By applying Jensen's inequality, we conclude

$$F(\beta, h) = \lim_{\substack{n \rightarrow \infty \\ n=(n,n)}} \frac{1}{\eta n} \mathbb{E} \log [(Z_n^q)^\eta] \leq \lim_{\substack{n \rightarrow \infty \\ n=(n,n)}} \frac{1}{\eta n} \log A_n = 0. \quad (3.5.8)$$

□

*Proof of Theorem 3.2.3.* We now assume  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ . We fix  $h$  as in Theorem 3.2.3:

$$h = h(\beta) := \begin{cases} \beta^{\max(\frac{2\alpha}{2\alpha-1}, 2) + \varepsilon} & \text{if } m_1 \neq 0, \\ \beta^{\max(\frac{4\alpha}{2\alpha-1}, 4) + \varepsilon} & \text{if } m_1 = 0, m_4 > m_2^2, \end{cases} \quad (3.5.9)$$

where  $\varepsilon > 0$  is arbitrarily small, but fixed. Our goal is to choose  $\eta \in (0, 1)$  and  $k \in \mathbb{N}$  such that  $\rho_1, \rho_2$  and  $\rho_3$  (which is symmetric to  $\rho_2$ ) are small, so that Lemma 3.5.1 implies  $F(\beta, h) = 0$  and  $h_c(\beta) \geq h$ . First we pick

$$k = k(\beta) := \frac{1}{F(0, h)}, \quad (3.5.10)$$

which is the correlation length of the annealed system (actually we take the integer part of it, but we omit to write it for clarity purpose). Notice that  $k \rightarrow \infty$  as  $h \rightarrow 0$ , and recall that Theorem 3.1.2 allows us to write  $k = L_\alpha(1/h)h^{-1/\alpha}$  with  $L_\alpha$  a slowly varying function when  $\alpha \leq 1$ ; and  $k \sim c_\alpha h^{-1}$  for some  $c_\alpha > 0$  as  $h \searrow 0$  when  $\alpha > 1$  (here we slightly changed the notations from the theorem, to lighten upcoming computations).

Note that, if  $\eta$  is picked such that  $(2 + \alpha)\eta > 2$ , then we have

$$\sum_{j \succ k-i} K(\|j\|) \leq \frac{L_1(\|k-i\|)}{\|k-i\|^{(2+\alpha)\eta-2}}. \quad (3.5.11)$$

Therefore,

$$\rho_1 \leq \sum_{i \prec k} \frac{L_1(\|k-i\|)}{\|k-i\|^{(2+\alpha)\eta-2}} A_i \leq \sum_{i \prec k} \frac{L_2(k-i^{(2)})}{(k-i^{(2)})^{(2+\alpha)\eta-2}} A_i, \quad (3.5.12)$$

(because  $\|k-i\| \geq k-i^{(2)}$ ). Similarly,

$$\sum_{\substack{j^{(1)} < k-i^{(1)} \\ j^{(2)} \geq k-i^{(2)}}} K(\|j\|) \leq \sum_{j^{(2)} \geq k-i^{(2)}} \frac{L_3(j^{(2)})}{(j^{(2)})^{(2+\alpha)\eta-1}} \leq \frac{L_4(k-i^{(2)})}{(k-i^{(2)})^{(2+\alpha)\eta-2}}. \quad (3.5.13)$$

Thus,

$$\rho_2 \leq \sum_{i \prec k} \frac{L_4(k-i^{(2)})}{(k-i^{(2)})^{(2+\alpha)\eta-2}} A_i =: S. \quad (3.5.14)$$

Let us denote this sum  $S$ . Because  $\rho_3$  is symmetric to  $\rho_2$ , and because  $\rho_1$  and  $\rho_2$  are bounded by the same sum  $S$  (up to the slowly varying function, which doesn't change the behavior of the sum), Theorem 3.2.3 will be proven as soon as we show that  $S$  can be made small, by applying Lemma 3.5.1. For that we need some precise estimates on  $A_i$  for  $i \prec k$ . We will handle  $A_i$  differently depending on whether  $i$  is small or not, using Lemmas 3.5.2-3.5.3 below.

**Lemma 3.5.2.** *There exist constants  $C_2 > 0, h_1 > 0$  and a slowly varying function  $L_5$  such that for any  $h \in (0, h_1)$  and  $\mathbf{n}$  with  $1 \leq \|\mathbf{n}\| \leq 1/F(0, h)$ , one has*

$$Z_{\mathbf{n}, h} \leq \begin{cases} C_2 L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} & \text{if } \alpha \in (0, 1), \\ L_5(\|\mathbf{n}\|) \|\mathbf{n}\|^{-1/\min(\alpha, 2)} & \text{if } \alpha > 1; \end{cases} \quad (3.5.15)$$

where  $Z_{\mathbf{n}, h}$  is the partition function of the homogeneous model with parameter  $h$ .

When  $\alpha = 1$ , the upper bound  $Z_{\mathbf{n}, h} \leq L_5(1/\|\mathbf{n}\|) \|\mathbf{n}\|^{-1}$  holds for  $\|\mathbf{n}\| \leq F(0, h)^{-(1-\varepsilon^2)}$ .

This lemma implies that  $Z_{n,h}$  remains of the same order as  $Z_{n,0}$  as long as  $n$  does not exceed the correlation length  $1/F(0, h)$  (in the case  $\alpha = 1$  we added the assumption  $\|\mathbf{n}\| \leq F(0, h)^{-(1-\varepsilon^2)}$  only to avoid technicalities). Then we have by Jensen's inequality that  $A_i \leq (\mathbb{E}Z_i^q)^\eta = (Z_{i,h})^\eta$ , therefore Lemma 3.5.2 gives us a first bound on  $A_i$  valid for all  $i$  not too large.

When  $i$  is of greater order, we can obtain an improved bound on  $A_i$  as soon as  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ .

**Lemma 3.5.3.** *Assume that  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ . If  $\varepsilon$  has been fixed small enough, then there exist some constants  $C_3, C_4, C_5 > 0$  such that for any  $k^{(1-\varepsilon^2)} \leq \|\mathbf{n}\| \leq 2k$ , and for  $h$  defined as in (3.5.9), one has*

$$A_n \leq \begin{cases} C_3 \|\mathbf{n}\|^{-(2-\alpha+\varepsilon C_4)\eta} & \text{if } \alpha \in (1/2, 1], \\ C_3 \|\mathbf{n}\|^{-(\alpha+\varepsilon C_5)\eta} & \text{if } \alpha > 1. \end{cases} \quad (3.5.16)$$

We postpone the proof of Lemma 3.5.2 to Appendix 3.A (because it only concerns the homogeneous gPS model), and we prove Lemma 3.5.3 later in this section. We now have all the tools we need to finish the proof. Fix  $\varepsilon > 0$  sufficiently small for Lemma 3.5.3 to apply, then define  $\eta$  depending on  $\alpha > 1/2$ .

Case  $\alpha \in (1/2, 1)$ . Choose  $\eta$  such that

$$\max\left(\frac{4}{4+\varepsilon C_4}, \frac{4-2\varepsilon^2}{4-2\varepsilon^2+\varepsilon^2\alpha}, \frac{2}{2+\alpha}, \frac{1}{2-\alpha}\right) < \eta < 1. \quad (3.5.17)$$

In particular  $(2+\alpha)\eta > 2$  so (3.5.12) and (3.5.14) hold and we only have to control the sum  $S$  defined in (3.5.14). We introduce the notation  $u_k := k^{(1-\varepsilon^2)}$ , and we part the sum in two:

$$S := S_1 + S_2 := \sum_{\substack{i \prec k, \\ \|\mathbf{i}\| \geq u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} A_i + \sum_{\substack{i \prec k, \\ \|\mathbf{i}\| < u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} A_i. \quad (3.5.18)$$

To bound  $S_1$ , we use Lemma 3.5.3 to estimate  $A_i$ .

$$\begin{aligned} S_1 &\leq \sum_{\substack{i \prec k, \\ \|\mathbf{i}\| \geq u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} \frac{C_3}{\|\mathbf{i}\|^{(2-\alpha+\varepsilon C_4)\eta}} \\ &\leq \sum_{\mathbf{i}^{(2)}=0}^{k-1} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} \frac{C_6}{(\max(\mathbf{i}^{(2)}, u_k))^{(2-\alpha+\varepsilon C_4)\eta-1}}, \end{aligned} \quad (3.5.19)$$

where we used  $(2-\alpha+\varepsilon C_4)\eta \geq (2-\alpha)\eta > 1$  (recall (3.5.17) and  $\eta > \eta_0$ ), and summed over  $\mathbf{i}^{(1)}$ . We split this sum once more according to whether  $\mathbf{i}^{(2)} \leq k/2$  or  $\mathbf{i}^{(2)} > k/2$ , which gives

$$\begin{aligned} S_1 &\leq \left( \sum_{\mathbf{i}^{(2)}=0}^{k/2} + \sum_{\mathbf{i}^{(2)}=k/2+1}^{k-1} \right) \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} \frac{C_7}{(\max(\mathbf{i}^{(2)}, u_k))^{(2-\alpha+\varepsilon C_4)\eta-1}} \\ &\leq \frac{L_6(k)}{k^{(2+\alpha)\eta-2}} \frac{C_8}{k^{(2-\alpha+\varepsilon C_4)\eta-2}} + \frac{L_7(k)}{k^{(2+\alpha)\eta-3}} \frac{C_9}{k^{(2-\alpha+\varepsilon C_4)\eta-1}} \leq \frac{L_8(k)}{k^{(4+\varepsilon C_4)\eta-4}}, \end{aligned} \quad (3.5.20)$$

where, in the first term we used  $(2 - \alpha + \varepsilon C_4)\eta < 2$  and bounded uniformly in the first factor, and in the second we used  $(2 + \alpha)\eta < 3$  and bounded uniformly in the second factor. Finally we have  $(4 + \varepsilon C_4)\eta - 4 > 0$  because of (3.5.17), so  $S_1$  can be made small if  $k$  is large enough (*i.e.* if  $\beta$  is small enough). Note that it is crucial here to have  $C_4 > 0$  thanks to Lemma 3.5.3, which relies on the assumption that  $\mathbb{P} \neq \mathbb{P}_{\pm x}$  for all  $x > 0$ .

For  $S_2$ , we bound the first factor uniformly in  $\mathbf{i}^{(2)}$  (note that  $(k - \mathbf{i}^{(2)}) \geq k/2$  for  $\|\mathbf{i}\| < u_k$ ), and we estimate  $A_i$  with Lemma 3.5.2.

$$\begin{aligned} S_2 &\leq \frac{L_9(k)}{k^{(2+\alpha)\eta-2}} \sum_{\substack{\mathbf{i} \prec k, \\ 1 \leq \|\mathbf{i}\| < u_k}} \frac{C_1}{\|\mathbf{i}\|^{(2-\alpha)\eta} L(\|\mathbf{i}\|)^\eta} + \frac{L_4(k)}{k^{(2+\alpha)\eta-2}} A_0 \\ &\leq \frac{L_{10}(k)}{k^{(2+\alpha)\eta-2} u_k^{(2-\alpha)\eta-2}} + \frac{L_4(k)}{k^{(2+\alpha)\eta-2}} \\ &\leq \frac{L_{10}(k)}{k^{(2+\alpha)\eta-2+(1-\varepsilon^2)((2-\alpha)\eta-2)}} + \frac{L_{11}(k)}{k^{(2+\alpha)\eta-2}}, \end{aligned} \quad (3.5.21)$$

where we used  $(2 - \alpha)\eta \in (1, 2)$ , and we recall  $u_k = k^{(1-\varepsilon^2)}$  (note that we had to write separately the term  $\mathbf{i} = \mathbf{0}$ ). The exponent in the denominator of the first term can be written  $(4 - 2\varepsilon^2 + \varepsilon^2\alpha)\eta - 4 + 2\varepsilon^2$  which is positive because of (3.5.17), and the other exponent is also positive. Thus  $S_2$  is also small for  $k$  large (*i.e.*  $\beta$  small).

Therefore, there is some  $\beta_\varepsilon > 0$  such that for any  $\beta \in (0, \beta_\varepsilon)$ ,  $S \leq 1/3$ . Thereby  $\rho_1 + \rho_2 + \rho_3 \leq 1$  for our choice of  $k$  and  $\eta$ . Applying Lemma 3.5.1, this concludes the proof of Theorem 3.2.3 in the case  $\alpha \in (1/2, 1)$ .

*Case  $\alpha > 1$ .* This is very similar to the previous case. Choose  $\eta$  such that

$$\max \left( \frac{3}{2 + \alpha}, \frac{1}{\alpha + \varepsilon C_5}, \frac{4 - 2\varepsilon^2}{2 + \alpha + \frac{1-\varepsilon^2}{\min(\alpha, 2)}} \right) < \eta < 1. \quad (3.5.22)$$

The first two terms in the maximum are obviously strictly smaller than 1; regarding the third one, one should first notice that  $2 - \frac{1}{\min(\alpha, 2)}$  is strictly smaller than  $\alpha$  (it is obvious if  $\alpha \geq 2$ , and follows from  $(\alpha - 1)^2 > 0$  if  $1 < \alpha < 2$ ). This implies  $(1 - \varepsilon^2)(2 - \frac{1}{\min(\alpha, 1)}) < \alpha$ , which is equivalent to  $4 - 2\varepsilon^2 < 2 + \alpha + \frac{(1-\varepsilon^2)}{\min(\alpha, 2)}$ , thus  $\eta$  is well defined. Moreover  $(2 + \alpha)\eta > 2$  so (3.5.12) and (3.5.14) hold again and we only have to control  $S$ .

As in the case  $\alpha \in (1/2, 1)$ , we split it according to  $\|\mathbf{i}\| \geq u_k := k^{(1-\varepsilon^2)}$  and  $\|\mathbf{i}\| < u_k$  to obtain (3.5.18). We bound  $S_1$  with Lemma 3.5.3.

$$\begin{aligned} S_1 &\leq \sum_{\substack{\mathbf{i} \prec k, \\ \|\mathbf{i}\| \geq u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} \frac{C_3}{\|\mathbf{i}\|^{(\alpha+\varepsilon C_5)\eta}} \\ &\leq \left( \sum_{\mathbf{i}^{(2)}=0}^{k/2} + \sum_{\mathbf{i}^{(2)}=k/2+1}^{k-1} \right) \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{(2+\alpha)\eta-2}} \frac{C_{10}}{(\max(\mathbf{i}^{(2)}, u_k))^{(\alpha+\varepsilon C_5)\eta-1}} \\ &\leq \frac{L_{12}(k)}{k^{(2+\alpha)\eta-2}} \frac{C_{11}}{k^{(\alpha+\varepsilon C_5)\eta-2}} + \frac{L_{13}(k)}{k^{(2+\alpha)\eta-3}} \frac{C_{12}}{k^{(\alpha+\varepsilon C_5)\eta-1}} \leq \frac{L_{14}(k)}{k^{(2+2\alpha+\varepsilon C_5)\eta-4}}, \end{aligned} \quad (3.5.23)$$

where we used  $(2 + \alpha)\eta > 3$  and  $(\alpha + \varepsilon C_5)\eta > 1$  to estimate the sums, and the last bound goes to 0 as  $k \rightarrow \infty$  because  $(2 + 2\alpha + \varepsilon C_5)\eta > 4$ .

For  $S_2$ , we bound the first factor uniformly in  $\mathbf{i}^{(2)}$  and estimate  $A_i$  with Lemma 3.5.2.

$$\begin{aligned}
 S_2 &\leq \frac{L_{15}(k)}{k^{(2+\alpha)\eta-2}} \sum_{\substack{\mathbf{i} \prec k, \\ 1 \leq \|\mathbf{i}\| < u_k}} \frac{L_{16}(\|\mathbf{i}\|)}{\|\mathbf{i}\|^{\eta/\min(\alpha,2)}} + \frac{L_4(k)}{k^{(2+\alpha)\eta-2}} A_0 \\
 &\leq \frac{L_{17}(k)}{k^{(2+\alpha)\eta-2} u_k^{\eta/\min(\alpha,2)-2}} + \frac{L_4(k)}{k^{(2+\alpha)\eta-2}} \\
 &\leq \frac{L_{17}(k)}{k^{(2+\alpha)\eta-2+(1-\varepsilon^2)(\eta/\min(\alpha,2)-2)}} + \frac{L_{18}(k)}{k^{(2+\alpha)\eta-2}}.
 \end{aligned} \tag{3.5.24}$$

The exponent in the denominator of the first term can be written  $(2 + \alpha + \frac{1-\varepsilon^2}{\min(\alpha,2)})\eta - (4 - 2\varepsilon^2)$ , which is positive. Finally  $S_2$  vanishes too as  $k \rightarrow \infty$ , and this concludes the proof of Theorem 3.2.3 in the case  $\alpha > 1$  with Lemma 3.5.3.

Case  $\alpha = 1$ . Choose  $\eta$  such that

$$\max\left(\frac{2}{3}, \frac{4}{4 + \varepsilon C_4}\right) < \eta < 1. \tag{3.5.25}$$

We have  $(2 + \alpha)\eta = 3\eta > 2$  again, so we only have to control  $S$ , which we split it again in  $S_1 + S_2$  as in (3.5.18). First we control  $S_1$  with Lemma 3.5.3.

$$\begin{aligned}
 S_1 &\leq \sum_{\substack{\mathbf{i} \prec k, \\ \|\mathbf{i}\| \geq u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{3\eta-2}} \frac{C_3}{\|\mathbf{i}\|^{(1+\varepsilon C_4)\eta}} \\
 &\leq \left( \sum_{\mathbf{i}^{(2)=0} }^{k/2} + \sum_{\mathbf{i}^{(2)=k/2+1} }^{k-1} \right) \frac{L_4(k - \mathbf{i}^{(2)})}{(k - \mathbf{i}^{(2)})^{3\eta-2}} \frac{C_{13}}{(\max(\mathbf{i}^{(2)}, u_k))^{(1+\varepsilon C_4)\eta-1}} \\
 &\leq \frac{L_{19}(k)}{k^{3\eta-2}} \frac{C_{14}}{k^{(1+\varepsilon C_4)\eta-2}} + \frac{L_{20}(k)}{k^{3\eta-3}} \frac{C_{15}}{k^{(1+\varepsilon C_4)\eta-1}} \leq \frac{L_{21}(k)}{k^{(4+\varepsilon C_4)\eta-4}},
 \end{aligned} \tag{3.5.26}$$

and this sum decays because of (3.5.25). Then we control  $S_2$  with Lemma 3.5.2.

$$\begin{aligned}
 S_2 &\leq \sum_{\substack{\mathbf{i} \prec k, \\ 1 \leq \|\mathbf{i}\| < u_k}} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - i_2)^{3\eta-2}} \frac{L_5(\|\mathbf{i}\|)}{\|\mathbf{i}\|} + \frac{L_4(k)}{k^{3\eta-2}} A_0 \\
 &\leq \sum_{i_2=0}^{u_k} \frac{L_4(k - \mathbf{i}^{(2)})}{(k - i_2)^{3\eta-2}} C_{16} \log(u_k) + \frac{L_4(k)}{k^{3\eta-2}} \leq \frac{L_{22}(k)}{k^{3\eta-2}},
 \end{aligned} \tag{3.5.27}$$

so  $S_2$  vanishes too as  $k \rightarrow \infty$ . This concludes the proof of Theorem 3.2.3 for  $\alpha = 1$ .  $\square$

### 3.5.2 Fractional moment estimate: proof of Lemma 3.5.3

We prove Lemma 3.5.3 via a change of measure procedure. Even though we obtain the same estimate on the fractional moment in both cases  $m_1 \neq 0$  and  $m_1 = 0, m_4 > m_2^2$ , we have to handle them separately and with different values of  $h$ .

**Case  $m_1 \neq 0$ .** For any  $\mathbf{n} \in \mathbb{N}^2$ , let us define a new probability measure by tilting both sequences  $(\widehat{\omega}_i)_{i \leq \mathbf{n}^{(1)}}, (\bar{\omega}_i)_{i \leq \mathbf{n}^{(2)}}$  by some  $\delta \in (-\delta_0, \delta_0)$ :

$$\frac{d\mathbb{P}_{\mathbf{n},\delta}}{d\mathbb{P}}(\omega) = \frac{e^{\delta \sum_{i=1}^{\mathbf{n}^{(1)}} \widehat{\omega}_{i_1} + \delta \sum_{i_2=1}^{\mathbf{n}^{(2)}} \bar{\omega}_{i_2}}}{Q(\delta, 0)^{\|\mathbf{n}\|/2}}, \quad (3.5.28)$$

where

$$Q(\delta, \beta) := \mathbb{E}\left[e^{\beta \omega_1 + \delta \widehat{\omega}_1 + \delta \bar{\omega}_1}\right]. \quad (3.5.29)$$

It is well-defined for  $\beta \in [0, \beta_1)$  and  $\delta \in (-\delta_0, \delta_0)$ , for some  $\beta_1, \delta_0 > 0$ , because

$$\begin{aligned} Q(\delta, \beta) &\leq \mathbb{E}\left[e^{(\beta+2|\delta|)\max(|\omega_1|, \widehat{\omega}_1, \bar{\omega}_1)}\right] \leq \mathbb{E}\left[e^{(\beta+2|\delta|)|\omega_1|}\right] + 2\mathbb{E}\left[e^{(\beta+2|\delta|)\widehat{\omega}_1}\right] \\ &\leq \mathbb{E}\left[e^{(\beta+2|\delta|)(\widehat{\omega}_1^2 + \bar{\omega}_1^2)/2}\right] + 2\mathbb{E}\left[e^{(\beta+2|\delta|)(\widehat{\omega}_1^2+1)}\right] < \infty, \end{aligned}$$

where we used  $|\omega_1| \leq \frac{1}{2}(\widehat{\omega}_1^2 + \bar{\omega}_1^2)$  and  $\widehat{\omega}_1 \leq \widehat{\omega}_1^2 + 1$  for all  $\widehat{\omega}_1, \bar{\omega}_1 \in \mathbb{R}$ . Recall  $A_{\mathbf{n}} = \mathbb{E}\left[(Z_{\mathbf{n}}^q)^\eta\right]$  with  $\eta \in (0, 1)$ . Using Hölder's inequality, we write for any  $\mathbf{n} \in \mathbb{N}^2, \delta \in (-\delta_0, \delta_0)$ ,

$$\begin{aligned} A_{\mathbf{n}} &= \mathbb{E}\left[(Z_{\mathbf{n}}^q)^\eta\right] = \mathbb{E}_{\mathbf{n},\delta}\left[(Z_{\mathbf{n}}^q)^\eta \left(\frac{d\mathbb{P}_{\mathbf{n},\delta}}{d\mathbb{P}}(\omega)\right)^{-1}\right] \\ &\leq \left(\mathbb{E}_{\mathbf{n},\delta}[Z_{\mathbf{n}}^q]\right)^\eta \left(\mathbb{E}_{\mathbf{n},\delta}\left[\left(\frac{d\mathbb{P}_{\mathbf{n},\delta}}{d\mathbb{P}}(\omega)\right)^{-1/(1-\eta)}\right]\right)^{1-\eta} \\ &= \left(\mathbb{E}_{\mathbf{n},\delta}[Z_{\mathbf{n}}^q]\right)^\eta \left(Q(\delta, 0)^\eta Q(-\delta\eta/(1-\eta), 0)^{1-\eta}\right)^{\|\mathbf{n}\|/2}. \end{aligned} \quad (3.5.30)$$

Using a Taylor expansion of  $Q(\delta, 0) = \mathbb{E}[e^{\delta \widehat{\omega}_1}]^2$  as  $\delta \rightarrow 0$  (which we will detail below in (3.5.34)), we have

$$Q(\delta, 0)^\eta Q(-\delta\eta/(1-\eta), 0)^{1-\eta} = 1 + \delta^2(m_2 - m_1^2)\frac{\eta}{1-\eta} + O(|\delta|^3), \quad (3.5.31)$$

where  $m_2 - m_1^2 = \text{Var}(\widehat{\omega}_1) > 0$  (otherwise disorder is constant a.s.), so this expression is bounded by  $\exp(C_{17} \delta^2)$  for some  $C_{17} > 0$ , uniformly in  $\delta \in (-\delta_1, \delta_1)$  for some  $\delta_1 > 0$ . Therefore,

$$A_{\mathbf{n}} \leq \left(\mathbb{E}_{\mathbf{n},\delta}[Z_{\mathbf{n}}^q]\right)^\eta \exp\left(C_{17} \delta^2 \|\mathbf{n}\|/2\right). \quad (3.5.32)$$

We fix right away  $|\delta| := \|\mathbf{n}\|^{-1/2} > 0$ , so that the exponential is bounded by a constant—the sign of  $\delta$  will be determined below.

Let us now compute  $\mathbb{E}_{\mathbf{n}, \delta}[Z_{\mathbf{n}}^q]$ , and estimate it with our choice of  $\delta$ . Using Fubini's theorem, one has

$$\begin{aligned} \mathbb{E}_{\mathbf{n}, \delta}[Z_{\mathbf{n}}^q] &= \mathbf{E} \left[ \mathbb{E}_{\mathbf{n}, \delta} \left[ e^{\sum_{i \in \llbracket 1, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}} \right] \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right] \\ &= \mathbf{E} \mathbb{E} \left[ e^{(-\lambda(\beta) + h) |\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} e^{\sum_{i \in \tau \cap \llbracket 1, \mathbf{n} \rrbracket} \beta \omega_i} \prod_{i_1=1}^{\mathbf{n}^{(1)}} \frac{e^{\delta \widehat{\omega}_{i_1}}}{Q(\delta, 0)^{1/2}} \prod_{i_2=1}^{\mathbf{n}^{(2)}} \frac{e^{\delta \bar{\omega}_{i_2}}}{Q(\delta, 0)^{1/2}} \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right] \\ &= \mathbf{E} \left[ e^{(-\lambda(\beta) + h) |\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} \prod_{i \in \tau \cap \llbracket 1, \mathbf{n} \rrbracket} \frac{\mathbb{E} \left[ e^{\beta \omega_i + \delta \widehat{\omega}_{i(1)} + \delta \bar{\omega}_{i(2)}} \right]}{Q(\delta, 0)} \right. \\ &\quad \left. \times \prod_{\substack{i_1 \leq \mathbf{n}^{(1)} \\ i_1 \notin \tau^{(1)}}} \frac{\mathbb{E} \left[ e^{\delta \widehat{\omega}_{i_1}} \right]}{Q(\delta, 0)^{1/2}} \prod_{\substack{i_2 \leq \mathbf{n}^{(2)} \\ i_2 \notin \tau^{(2)}}} \frac{\mathbb{E} \left[ e^{\delta \bar{\omega}_{i_2}} \right]}{Q(\delta, 0)^{1/2}} \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right] \end{aligned}$$

where we used that  $\widehat{\omega}_i, \bar{\omega}_j, i, j \in \mathbb{N}$  are independent variables. Noticing that the factors with  $i_1 \notin \tau^{(1)}, i_2 \notin \tau^{(2)}$  simplify to 1, and recalling  $e^{\lambda(\beta)} = Q(0, \beta)$ , this finally gives

$$\mathbb{E}_{\mathbf{n}, \delta}[Z_{\mathbf{n}}^q] = \mathbf{E} \left[ \left( e^h \frac{Q(\delta, \beta)}{Q(\delta, 0)Q(0, \beta)} \right)^{|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right]. \quad (3.5.33)$$

Now we have to estimate  $\text{Frac}(\delta, \beta) := \frac{Q(\delta, \beta)}{Q(\delta, 0)Q(0, \beta)}$  as  $\beta, \delta \rightarrow 0$  (recall that they are related to  $h$  by (3.5.9), (3.5.10) and  $|\delta| = \|\mathbf{n}\|^{-1/2}$ , with  $k^{1-\varepsilon^2} \leq \|\mathbf{n}\| \leq 2k$ ). A Taylor expansion of  $Q(\delta, \beta)$  to the second order in  $(\delta, \beta)$ , combined with the finite exponential moment, gives

$$Q(\delta, \beta) = 1 + \beta m_1^2 + 2\delta m_1 + \frac{1}{2} \beta^2 m_2^2 + 2\delta \beta m_2 m_1 + \delta^2 (m_2 + m_1^2) + O(\cdot^3), \quad (3.5.34)$$

with  $O(\cdot^3) := O(\beta^3) + O(\beta^2 \delta) + O(\beta \delta^2) + O(\delta^3)$ . This leads to the following expansion:

$$\text{Frac}(\delta, \beta) = 1 + \delta \beta m_1 (m_2 - m_1^2) + O(\beta^p) + O(\delta^2) + o(\delta \beta), \quad (3.5.35)$$

where we can push the expansion to any order  $p \in \mathbb{N}$  in  $\beta$  because  $\text{Frac}(0, \beta) = 1$  for all  $\beta \geq 0$  and  $\text{Frac}(\delta, \beta)$  is  $\mathcal{C}^\infty$  on  $(-\delta_0, \delta_0) \times [0, \beta_1)$ . Recall that  $m_2 - m_1^2 > 0$  and that we are in the case  $m_1 \neq 0$ . Let us now prove that, with our choice of  $\delta$ , the order  $\delta \beta$  is dominating in (3.5.35), and deduce a bound on  $\mathbb{E}_{\mathbf{n}, \delta}[Z_{\mathbf{n}}^q]$ .

*Case  $\alpha \in (1/2, 1]$ :* Because of our assumptions  $k^{(1-\varepsilon^2)} \leq \|\mathbf{n}\| \leq 2k$ , (3.5.9), (3.5.10) and Theorem 3.1.2, and because we can bound the slowly varying function from Theorem 3.1.2 by  $h^{\varepsilon^{3/2}} \leq L_\alpha(1/h) \leq h^{-\varepsilon^{3/2}}$  for  $h$  sufficiently small, there exists some  $\beta_\varepsilon > 0$  such that for any  $\beta < \beta_\varepsilon$ ,

$$\begin{aligned} h^{\frac{1}{2}(\frac{1}{\alpha} + \varepsilon^{3/2})} &\leq |\delta| \leq h^{\frac{1}{2}(\frac{1}{\alpha} - \varepsilon^{3/2})(1-\varepsilon^2)}, \\ \beta^{\frac{1}{2\alpha-1} + \frac{1}{2\alpha}\varepsilon + c_1 \varepsilon^{3/2}} &\leq |\delta| \leq \beta^{\frac{1}{2\alpha-1} + \frac{1}{2\alpha}\varepsilon - c_1 \varepsilon^{3/2}}, \end{aligned} \quad (3.5.36)$$



for some constant  $c_1 > 0$  provided that  $\varepsilon$  was taken sufficiently small. Notice that  $\delta$  decays faster than  $\beta$  (especially if  $\alpha$  is close to  $1/2$ ), so  $\delta^2 \ll |\delta|\beta$ . Moreover we can fix  $p$  sufficiently large (depending on  $\alpha$ ) such that  $\beta^p \ll |\delta|\beta$ . Finally, letting  $\delta := -\text{sign}(m_1)\|\mathbf{n}\|^{-1/2}$ , there exist a constant  $c > 0$  such that for any  $\beta < \beta_\varepsilon$ , we have

$$\text{Frac}(\delta, \beta) \leq 1 - c|\delta|\beta \leq e^{-c|\delta|\beta}. \quad (3.5.37)$$

Eventually, noticing that  $h = \beta^{\frac{2\alpha}{2\alpha-1} + \varepsilon} \ll \beta^{\frac{2\alpha}{2\alpha-1} + \frac{1}{2\alpha}\varepsilon + c_1\varepsilon^{3/2}} \leq |\delta|\beta$  if  $\varepsilon$  is sufficiently small, and recalling  $|\delta| = \|\mathbf{n}\|^{-1/2}$ , we can bound (3.5.33) from above by

$$\mathbb{E}_{\mathbf{n}, \delta} [Z_{\mathbf{n}}^q] \leq \mathbf{E} \left[ \left( e^{-c'\|\mathbf{n}\|^{-1/2}\beta} \right)^{|\tau \cap [\mathbf{1}, \mathbf{n}]|} \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right], \quad (3.5.38)$$

for some  $c' > 0$ .

*Case  $\alpha > 1$ :* This is very similar to the previous case, with simpler bounds because the slowly varying function from Theorem 3.1.2 is replaced by a constant. So putting together (3.5.10), Theorem 3.1.2 and our assumption  $k^{(1-\varepsilon^2)} \leq \|\mathbf{n}\| \leq 2k$ , we deduce that there exist  $c_1, c_2 > 0$  and  $\beta_\varepsilon > 0$  such that for any  $\beta < \beta_\varepsilon$  we have  $c_1 h^{\frac{1}{2}} \leq |\delta| \leq c_2 h^{\frac{1-\varepsilon^2}{2}}$ . Recalling (3.5.9), this yields

$$c'_1 \beta^{1+\frac{\varepsilon}{2}} \leq |\delta| \leq c'_2 \beta^{\frac{1}{2}(2+\varepsilon)(1-\varepsilon^2)} \leq c'_2 \beta^{1+\frac{\varepsilon}{4}}. \quad (3.5.39)$$

Here again  $|\delta| \ll \beta$ , and we can fix  $p \in \mathbb{N}$  sufficiently large so that  $\beta^p \ll |\delta|\beta$ . Fixing  $\delta := -\text{sign}(m_1)\|\mathbf{n}\|^{-1/2} > 0$ , there is a constant  $c'' > 0$  such that for any  $\beta < \beta_\varepsilon$ , we have  $\text{Frac}(\delta, \beta) \leq e^{-c''|\delta|\beta}$ . Finally we have  $|\tau \cap [\mathbf{1}, \mathbf{n}]| \leq \|\mathbf{n}\| \leq C_{18}h^{-1}$  (recall Theorem 3.1.2), so  $e^{h|\tau \cap [\mathbf{1}, \mathbf{n}]|}$  is bounded uniformly by some constant. Thus:

$$\mathbb{E}_{\mathbf{n}, \delta} [Z_{\mathbf{n}}^q] \leq C_{19} \mathbf{E} \left[ \left( e^{-c''\|\mathbf{n}\|^{-1/2}\beta} \right)^{|\tau \cap [\mathbf{1}, \mathbf{n}]|} \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right]. \quad (3.5.40)$$

In both cases (3.5.38)-(3.5.40), we recognize the partition function of a homogeneous gPS model with parameter  $-c\|\mathbf{n}\|^{-1/2}\beta$ , for some  $c > 0$ . The following result (proven in Appendix 3.A) gives an estimate of this partition function.

**Lemma 3.5.4.** *For any  $\alpha > 0$ , fix some  $0 < \alpha_- < \alpha$ . Then there exist  $C_{20}, C_{21} > 0$  and  $u_0 > 0$  such that for any  $u \in (0, u_0)$  and  $\mathbf{n} \in \mathbb{N}^2$ , one has*

$$Z_{\mathbf{n}, -u} \leq C_{20} \frac{K(\|\mathbf{n}\|)}{u^2} + \mathbf{P}(\mathbf{n} \in \tau) e^{-C_{21}u\|\mathbf{n}\|^{\min(\alpha_-, 1)}}. \quad (3.5.41)$$

We conclude the proof of Lemma 3.5.3 by applying this result with  $u := c\|\mathbf{n}\|^{-1/2}\beta$  (which decays to 0 as  $\beta \searrow 0$ ), separating again the cases  $\alpha \in (1/2, 1]$  and  $\alpha > 1$ .

*Case  $\alpha \in (1/2, 1]$ .* We apply Lemma 3.5.4 to (3.5.38). Notice that, thanks to (3.5.36), we have

$$u \|\mathbf{n}\|^{\alpha_-} \geq \beta^2 \frac{\alpha - \alpha_-}{2\alpha - 1} - \varepsilon \frac{2\alpha - 1}{2\alpha} + c_1 \varepsilon^{3/2}. \quad (3.5.42)$$

Provided that  $\varepsilon$  is sufficiently small (so that  $\varepsilon \frac{2\alpha-1}{2\alpha} > 2c_1 \varepsilon^{3/2}$ ), we can choose  $\alpha_-$  sufficiently close to  $\alpha$  (depending on  $\varepsilon$ ) so that the exponent is negative; thus  $u\|\mathbf{n}\|^{\alpha_-}$  goes to infinity as a power of  $\beta$ . In particular the second term in (3.5.41) decays much faster than the first one (which decays at most polynomially in  $\beta$ , recall (3.5.36)). Hence, Lemma 3.5.4 and (3.5.36) give for any  $\beta$  sufficiently small,

$$Z_{\mathbf{n}, -c'\|\mathbf{n}\|^{-1/2}\beta} \leq C_{22} \frac{K(\|\mathbf{n}\|)}{\|\mathbf{n}\|^{-1} \beta^2} \leq \frac{C_{23}}{\|\mathbf{n}\|^{2-\alpha+\varepsilon C_{24}}}. \quad (3.5.43)$$

Recollecting (3.5.32) and (3.5.38), this finally proves the lemma in the case  $\alpha \in (1/2, 1]$ .

*Case  $\alpha > 1$ .* We apply Lemma 3.5.4 to (3.5.40). Recalling (3.5.39), we notice that  $u\|\mathbf{n}\| \geq C_{25}\beta^{-\varepsilon/4}$  diverges as a power of  $\beta$ : the second term in (3.5.41) decays again much faster than the first one as  $\beta \searrow 0$ , so Lemma 3.5.4 and (3.5.39) give for any  $\beta$  sufficiently small,

$$Z_{\mathbf{n}, -c''\|\mathbf{n}\|^{-1/2}\beta} \leq C_{26} \frac{K(\|\mathbf{n}\|)}{\|\mathbf{n}\|^{-1} \beta^2} \leq \frac{C_{27}}{\|\mathbf{n}\|^{\alpha+\varepsilon C_{28}}}. \quad (3.5.44)$$

With (3.5.32) and (3.5.40), this concludes the proof of the lemma in the case  $\alpha > 1$ .

**Case  $m_1 = 0, m_4 > m_2^2$ .** When  $m_1 = 0$ , one could push the expansion in (3.5.35) to the third order to find a dominating term  $\delta\beta^2 m_3^2$ . Assuming  $m_3 \neq 0$  and adjusting the definition of  $h$  (thus of  $\delta$ ), this would prove the fractional moment estimates with  $h = \beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)+\varepsilon}$ , hence a shift of the critical point of order  $\beta^{\max(\frac{4\alpha}{2\alpha-1}, 4)+\varepsilon}$ . However when  $m_1 = 0$ , we can define another change of measure which yields the same estimates on  $A_n$  without the additional assumption  $m_3 \neq 0$ . Indeed, let us define for all  $\mathbf{n} \in \mathbb{N}^2$  and  $\delta \in [0, \delta_0)$ ,

$$\frac{d\mathbb{P}'_{\mathbf{n}, \delta}(\omega)}{d\mathbb{P}} = \frac{e^{-\delta \sum_{i_1=1}^n \widehat{\omega}_{i_1}^2 - \delta \sum_{i_2=1}^n \bar{\omega}_{i_2}^2}}{R(\delta, 0)\|\mathbf{n}\|^{|\mathbf{n}|/2}}, \quad (3.5.45)$$

where we define

$$R(\delta, \beta) := \mathbb{E}\left[e^{\beta\omega_1 - \delta\widehat{\omega}_1^2 - \delta\bar{\omega}_1^2}\right], \quad (3.5.46)$$

This change of measure is a *quadratic* i.i.d tilt of the sequences  $\widehat{\omega}$  and  $\bar{\omega}$ . Similarly to (3.5.29), it is well-defined for  $\beta \in [0, \beta_1)$  and  $\delta \in [0, \delta_0)$ , for some  $\beta_1, \delta_0 > 0$  (notice that here we directly choose a non-positive tilt). As in the previous case, Hölder's inequality yields for any  $\mathbf{n} \in \mathbb{N}^2$ ,

$$A_n \leq \left(\mathbb{E}'_{\mathbf{n}, \delta}\left[Z_n^q\right]\right)^\eta \left(R(\delta, 0)^\eta R\left(-\delta\eta/(1-\eta), 0\right)^{1-\eta}\right)^{\|\mathbf{n}\|/2}. \quad (3.5.47)$$

Moreover a Taylor expansion of  $R(\delta, \beta)$  to the third order in  $(\delta, \beta)$  yields

$$\begin{aligned} R(\delta, \beta) &= 1 - 2\delta m_2 + \frac{\beta^2 m_2^2}{2} + \delta^2(m_4 + m_2^2) + \frac{\beta^3 m_3^2}{6} \\ &\quad - \beta^2 \delta m_2 m_4 + \beta \delta^2 m_3^2 - \delta^3 \left(\frac{m_6}{3} + m_2 m_4\right) + O(\cdot^4), \end{aligned} \quad (3.5.48)$$

with  $O(\cdot^4) := O(\beta^4) + O(\beta^3\delta) + \dots + O(\delta^4)$ , where we used  $m_1 = 0$ . Thus,

$$R(\delta, 0)^\eta R(-\delta\eta/(1-\eta), 0)^{1-\eta} = 1 + \delta^2(m_4 - m_2^2)\frac{\eta}{1-\eta} + O(\delta^3), \quad (3.5.49)$$

which is bounded by  $\exp(C_{29}\delta^2)$  for some  $C_{29} > 0$ , uniformly in  $\delta \in (-\delta_1, \delta_1)$  (recall that we assumed  $m_4 > m_2^2$ ). Therefore, letting  $\delta := \|\mathbf{n}\|^{-1/2} > 0$ , we have

$$A_{\mathbf{n}} \leq C_{30} \left( \mathbb{E}'_{\mathbf{n}, \delta} [Z_{\mathbf{n}}^q] \right)^\eta, \quad (3.5.50)$$

and similarly to (3.5.33), we write with a direct computation

$$\mathbb{E}'_{\mathbf{n}, \delta} [Z_{\mathbf{n}}^q] = \mathbf{E} \left[ \left( e^h \frac{R(\delta, \beta)}{R(\delta, 0)R(0, \beta)} \right)^{|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|} \mathbf{1}_{\{n \in \tau\}} \right]. \quad (3.5.51)$$

A Taylor expansion of  $\text{Frac}_2(\delta, \beta) := \frac{R(\delta, \beta)}{R(\delta, 0)R(0, \beta)}$  to the third order in  $(\delta, \beta)$  yields

$$\text{Frac}_2(\delta, \beta) = 1 - \delta\beta^2 m_2(m_4 - m_2^2) + O(\beta^p) + O(\delta\beta^3) + O(\delta^2\beta) + O(\delta^3), \quad (3.5.52)$$

where we can push the expansion to any order  $p \in \mathbb{N}$  in  $\beta$ . Notice that this is very similar to (3.5.35) above, with a leading term of order  $\beta\delta^2$  instead (and  $m_2(m_4 - m_2^2) > 0$  under our assumptions). By duplicating all arguments above, and recalling  $k^{(1-\varepsilon^2)} \leq \|\mathbf{n}\| \leq 2k$ , (3.5.10), Theorem 3.1.2 and our specific choice of  $h$  in (3.5.9), we finally obtain for some  $c > 0$  and any  $\alpha > 1/2$ ,

$$\mathbb{E}'_{\mathbf{n}, -\|\mathbf{n}\|^{-1/2}} [Z_{\mathbf{n}}^q] \leq \mathbf{E} \left[ \left( e^{-c\|\mathbf{n}\|^{-1/2}\beta^2} \right)^{|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|} \mathbf{1}_{\{n \in \tau\}} \right] = Z_{\mathbf{n}, -c\|\mathbf{n}\|^{-1/2}\beta^2}. \quad (3.5.53)$$

Eventually, we apply Lemma 3.5.4 (with  $u = c\|\mathbf{n}\|^{-1/2}\beta^2$ ) to estimate this homogeneous partition function, and we conclude the proof by recollecting (3.5.50) as in the case  $m_1 \neq 0$  (we do not write the details again).

### 3.6 Disorder relevance: shift of the critical point when

$$\mathbb{P} = \mathbb{P}_{\pm x}$$

In this section we prove the lower bound in Theorem 3.2.5 —recall that we assume  $m_1 = 0$  and  $m_4 = m_2^2$ , so  $\mathbb{P} = \mathbb{P}_{\pm x}$  for some  $x > 0$ — and in particular we discuss how the estimate of the fractional moment (*i.e.* Lemma 3.5.3) have to be adapted. Notice that we can assume  $\mathbb{P} = \mathbb{P}_{\pm 1}$ , that is  $x = 1$ , without loss of generality —one only has to replace  $\beta$  with  $\beta x^2$  in the partition function.

We can reproduce exactly the first part of the proof above, namely the coarse-graining procedure whose core is Lemma 3.5.1, and we can use Lemma 3.5.2 as it is. On the other hand, the change of measure argument needs important adaptation. In the case of an i.i.d. disorder, what plays the role of Lemma 3.5.3 is [18, Prop. 4.2]: there, the estimates of

fractional moments  $A_i$  use an i.i.d. tilt of the disorder, but only along an *extended diagonal*, i.e.

$$J_n := \{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket; |\mathbf{i}^{(1)} - \mathbf{i}^{(2)}| \leq 2\ell_n\}. \quad (3.6.1)$$

The width  $\ell_n$  is chosen depending on  $\alpha > 1$ , so that the renewal process  $\tau$  is very unlikely to deviate from the diagonal by more than  $\ell_n$  (see [18, Thm. A.5], or [20, Thm. 4.2] for a more general statement):

$$\ell_n := \begin{cases} (\mathbf{n}^{(1)})^{(1+\varepsilon^2)/\alpha} & \text{if } \alpha \in (1, 2], \\ C_1 \sqrt{\mathbf{n}^{(1)} \log \mathbf{n}^{(1)}} & \text{if } \alpha > 2, \end{cases} \quad (3.6.2)$$

(notice that  $|J_n| \leq 2\mathbf{n}^{(1)}\ell_n \ll \mathbf{n}^{(1)}\mathbf{n}^{(2)}$  when  $\mathbf{n}^{(1)} \approx \mathbf{n}^{(2)}$ ). The fact that disorder is not independent in our setting adds technicalities to this method, but they can be handled if we restrain ourselves to  $\mathbb{P}_{\pm 1}$ . We prove the following result, which plays the role of Lemma 3.5.3 in the case  $\mathbb{P} = \mathbb{P}_{\pm 1}$ .

**Proposition 3.6.1.** *Assume  $\alpha > 1$ ,  $\mathbb{P} = \mathbb{P}_{\pm 1}$ , recall  $k = 1/F(0, h)$ , and define  $\ell_n$  as in (3.6.2). Then there exist  $h_0 > 0$  and  $L_1$  such that for any  $h \in (0, h_0)$  and  $\sqrt{k} \leq \mathbf{n}^{(1)} \leq k$ ,  $\mathbf{n}^{(1)} \leq \mathbf{n}^{(2)} \leq \mathbf{n}^{(1)} + \ell_n$ , one has*

$$A_n \leq \begin{cases} L_1(k) k^{-(1+\varepsilon^2)\eta/\alpha} & \text{if } \alpha \in (1, 2], \\ L_1(k) k^{-(\alpha-1)\eta/2} & \text{if } \alpha > 2. \end{cases} \quad (3.6.3)$$

Note that these are exactly the same estimates on  $A_n$  as in [18, Prop. 4.2]. Once plugged in the computations of  $\rho_1, \rho_2$  and  $\rho_3$  from Lemma 3.5.1, they give the same lower bound for the shift of the critical point as in [18, Thm 1.4], which is

$$h_c(\beta) \geq \begin{cases} \beta^{\frac{2\alpha}{\alpha-1} + \varepsilon} & \text{if } \alpha \in (1, 2], \\ \beta^4 |\log \beta|^{-6} & \text{if } \alpha > 2. \end{cases} \quad (3.6.4)$$

This proves the left inequality in Theorem 3.2.5. We do not write the details here, because once we have the estimates on  $A_n$  from Proposition 3.6.1, the computations of  $\rho_1, \rho_2$  and  $\rho_3$  are the same as in [18] and do not depend on the setting of disorder.

*Proof of Proposition 3.6.1.* This follows the same scheme as [18, Prop. 4.2]. We define the same change of measure on  $J_n$  as [18, (4.18)], that is

$$\frac{d\bar{\mathbb{P}}_{n,\delta}}{d\mathbb{P}}(\omega) := \frac{\prod_{i \in J_n} e^{-\delta\omega_i}}{\bar{Q}_{J_n}(\delta)}, \quad \text{where } \bar{Q}_{J_n}(\delta) := \mathbb{E} \left[ \prod_{i \in J_n} e^{-\delta\omega_i} \right]. \quad (3.6.5)$$

Applying Hölder's inequality to  $A_n$  similarly to (3.5.47), we have

$$\begin{aligned} A_n &\leq \left( \bar{\mathbb{E}}_{n,\delta} [Z_n^q] \right)^\eta \left( \bar{\mathbb{E}}_{n,\delta} \left[ \left( \frac{d\bar{\mathbb{P}}_{n,\delta}}{d\mathbb{P}}(\omega) \right)^{-1/(1-\eta)} \right] \right)^{1-\eta} \\ &= \left( \bar{\mathbb{E}}_{n,\delta} [Z_n^q] \right)^\eta \bar{Q}_{J_n}(\delta)^\eta \bar{Q}_{J_n}(-\delta\eta/(1-\eta))^{1-\eta}. \end{aligned} \quad (3.6.6)$$

Let us fix  $\delta := (\mathbf{n}^{(1)}\ell_n)^{-\frac{1}{2}(1+\varepsilon^3)} > 0$ —this is the same as in [18] where we added the power  $(1 + \varepsilon^3)$  to avoid technicalities. We claim the following:

**Lemma 3.6.2.** *There exists  $C_2 > 0$  and  $\delta_1 > 0$  such that if  $\delta^2 \mathbf{n}^{(1)} \ell_n \leq \delta_1$ , then*

$$1 \leq \overline{Q}_{J_n}(\delta) \leq C_2. \quad (3.6.7)$$

*Proof of Lemma 3.6.2.* The lower bound follows directly from Jensen's inequality, so let us focus on the upper bound. For all  $1 \leq i \leq \mathbf{n}^{(1)}$ , we define

$$J_n(i) := \{j \in \llbracket 1, \mathbf{n}^{(2)} \rrbracket; (i, j) \in J_n\}, \quad \text{and} \quad \sigma_i = \sigma_i(\bar{\omega}) := \sum_{j \in J_n(i)} \bar{\omega}_j, \quad (3.6.8)$$

in particular  $|\sigma_i| \leq |J_n(i)| \leq 2\ell_n$ . Then, computing first the expectation conditionally on  $\bar{\omega}$ , we have

$$\overline{Q}_{J_n}(\delta) = \mathbb{E} \left[ \exp \left( -\delta \sum_{i=1}^{\mathbf{n}^{(1)}} \hat{\omega}_i \sigma_i(\bar{\omega}) \right) \right] = \mathbb{E} \left[ \prod_{i=1}^{\mathbf{n}^{(1)}} \cosh \left( \delta \sigma_i(\bar{\omega}) \right) \right], \quad (3.6.9)$$

where we used that  $\hat{\omega}$  is a sequence of independent variables, is independent from  $\bar{\omega}$ , and that  $\mathbb{E}[e^{-x\hat{\omega}_i}] = \cosh(-x\hat{\omega}_i) = \cosh(x)$  for all  $x \in \mathbb{R}$ . Notice that for all  $x \in \mathbb{R}$ , one has  $\cosh(x) \leq e^{\frac{x^2}{2}}$ : we therefore have

$$\overline{Q}_{J_n}(\delta) \leq \mathbb{E} \left[ \exp \left( \sum_{i=1}^{\mathbf{n}^{(1)}} \frac{\delta^2 \sigma_i^2}{2} \right) \right] \leq \prod_{i=1}^{\mathbf{n}^{(1)}} \mathbb{E} \left[ \exp \left( \frac{1}{2} \mathbf{n}^{(1)} \delta^2 \sigma_i^2 \right) \right]^{1/\mathbf{n}^{(1)}} \quad (3.6.10)$$

by Cauchy-Schwarz inequality. Here  $\sigma_i$  is a sum of  $|J_n(i)|$  i.i.d. bounded variables, so  $\frac{\sigma_i}{\sqrt{|J_n(i)|}}$  converges in distribution to some standard gaussian  $Z \sim \mathcal{N}(0, 1)$  as  $|J_n(i)| \rightarrow \infty$ . This implies

$$\overline{Q}_{J_n}(\delta) \leq C_3 + C'_3 \mathbb{E} \left[ \exp \left( \mathbf{n}^{(1)} \delta^2 |J_n(i)| Z^2 \right) \right] \quad (3.6.11)$$

uniformly in  $\mathbf{n}$  (where  $C_3$  handles the indexes  $i$  for which  $|J_n(i)|$  is small). Noticing that  $\mathbf{n}^{(1)} \delta^2 |J_n(i)| \leq 2 \delta^2 \mathbf{n}^{(1)} \ell_n \leq 2\delta_1$ , and that  $Z^2$  has some finite exponential moments, this concludes the proof of the lemma.  $\square$

Note that  $\delta^2 \mathbf{n}^{(1)} \ell_n \rightarrow 0$  as  $\mathbf{n}^{(1)} \rightarrow \infty$ , so this lemma implies that the last two factors in the right-hand side of (3.6.6) are uniformly bounded. It remains to deal with the expectation in (3.6.6). One has

$$\begin{aligned} \overline{\mathbb{E}}_{\mathbf{n}, \delta} [Z_n^q] &= \mathbb{E} \left[ \overline{\mathbb{E}}_{\mathbf{n}, \delta} \left[ e^{\sum_{i \in \llbracket 1, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}} \right] \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right] \\ &= \mathbb{E} \left[ \frac{e^{h|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|}}{\overline{Q}_{J_n}(\delta)} \mathbb{E} \left[ e^{-\lambda(\beta)|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} \prod_{i \in \llbracket 1, \mathbf{n} \rrbracket} e^{\beta \omega_i \mathbf{1}_{\{i \in \tau\}} - \delta \omega_i \mathbf{1}_{\{i \in J_n\}}} \right] \mathbf{1}_{\{\mathbf{n} \in \tau\}} \right], \end{aligned} \quad (3.6.12)$$

Under the assumptions of Proposition 3.6.1 we have  $|\tau \cap \llbracket 1, \mathbf{n} \rrbracket| \leq \|\mathbf{n}\| \leq C_4 h^{-1}$  (recall Theorem 3.1.2), so this implies that  $e^{h|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|}$  is bounded by some uniform constant (and so is  $\overline{Q}_{J_n}(\delta)^{-1}$ ).

Let us fix a realization of  $\tau$  with  $\mathbf{n} \in \tau$ , and let  $B_{\mathbf{n},\tau}^{\beta,\delta}$  be the expectation over  $\mathbb{P}$  in (3.6.12). Recall the definition of  $J_{\mathbf{n}}(i)$  and  $\sigma_i$  in Lemma 3.6.2, and define

$$\bar{\sigma}_i = \bar{\sigma}_i(\bar{\omega}) := \sum_{j=1}^{\mathbf{n}^{(2)}} \bar{\omega}_j \mathbf{1}_{\{(i,j) \in \tau\}}, \quad (3.6.13)$$

and notice that  $|\bar{\sigma}_i| = 1$  if  $i \in \llbracket 1, \mathbf{n}^{(1)} \rrbracket \cap \tau^{(1)}$ , and  $\bar{\sigma}_i = 0$  otherwise. Again, as for  $\bar{Q}_{J_{\mathbf{n}}}$ , computing the expectation first conditionally on  $\bar{\omega}$  gives

$$\begin{aligned} B_{\mathbf{n},\tau}^{\beta,\delta} &= \mathbb{E} \left[ e^{-\lambda(\beta)|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} \exp \left( \sum_{i=1}^{\mathbf{n}^{(1)}} \hat{\omega}_i (\beta \bar{\sigma}_i(\bar{\omega}) - \delta \sigma_i(\bar{\omega})) \right) \right] \\ &= \mathbb{E} \left[ e^{-\lambda(\beta)|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} \prod_{i=1}^{\mathbf{n}^{(1)}} \cosh(\beta \bar{\sigma}_i(\bar{\omega}) - \delta \sigma_i(\bar{\omega})) \right]. \end{aligned} \quad (3.6.14)$$

Notice that  $\lambda(\beta) = \log \cosh(\beta)$  when  $\mathbb{P} = \mathbb{P}_{\pm 1}$ . Using that  $\cosh$  is even and that  $\cosh 0 = 1$ , we get that

$$e^{\lambda(\beta)|\tau \cap \llbracket 1, \mathbf{n} \rrbracket|} = \prod_{i=1}^{\mathbf{n}^{(1)}} \cosh(\beta \mathbf{1}_{\{i \in \tau^{(1)}\}}) = \prod_{i=1}^{\mathbf{n}^{(1)}} \cosh(\beta \bar{\sigma}_i), \quad (3.6.15)$$

where we used that  $\bar{\sigma}_i = \pm 1$  when it is not equal to 0. Then (3.6.14) becomes

$$\begin{aligned} B_{\mathbf{n},\tau}^{\beta,\delta} &= \mathbb{E} \left[ \prod_{i=1}^{\mathbf{n}^{(1)}} \frac{\cosh(\beta \bar{\sigma}_i - \delta \sigma_i)}{\cosh(\beta \bar{\sigma}_i)} \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^{\mathbf{n}^{(1)}} \left( \cosh(\delta \sigma_i) - \sinh(\delta \sigma_i) \tanh(\beta \bar{\sigma}_i) \right) \right], \end{aligned} \quad (3.6.16)$$

where we simply used trigonometric identities to expand  $\cosh(\beta \bar{\sigma}_i - \delta \sigma_i)$ . Notice that  $\sinh(\delta \sigma_i) \tanh(\beta \bar{\sigma}_i) = \sinh(\delta \sigma_i \bar{\sigma}_i) \tanh(\beta)$  (because  $\bar{\sigma}_i \in \{-1, 0, 1\}$  and  $\sinh, \tanh$  are antisymmetric), and recall that  $|\delta \sigma_i| \ll (\ell_{\mathbf{n}}/\mathbf{n}^{(1)})^{1/2} \rightarrow 0$  as  $\mathbf{n}^{(1)} \rightarrow \infty$ . Therefore, some standard Taylor expansions give for  $\mathbf{n}^{(1)}$  sufficiently large,

$$\begin{aligned} \cosh(\delta \sigma_i) &\leq 1 + (\delta \sigma_i)^2, \\ -\sinh(\delta \sigma_i \bar{\sigma}_i) &\leq -\delta \sigma_i \bar{\sigma}_i + (\delta \sigma_i \bar{\sigma}_i)^2 \leq -\delta \sigma_i \bar{\sigma}_i + (\delta \sigma_i)^2. \end{aligned} \quad (3.6.17)$$

Finally, for  $\mathbf{n}^{(1)}$  sufficiently large, one has for all  $\bar{\omega} \in \{-1, 1\}^{\mathbb{N}}$ ,

$$\cosh(\delta \sigma_i) - \sinh(\delta \sigma_i) \tanh(\beta \bar{\sigma}_i) \leq 1 + 2(\delta \sigma_i)^2 - \delta \tanh(\beta) \sigma_i \bar{\sigma}_i, \quad (3.6.18)$$

(where we also used  $\tanh(\beta) \leq 1$ ). Since all the factors in (3.6.16) are positive for  $\mathbf{n}^{(1)}$  sufficiently large, we get

$$\begin{aligned} B_{\mathbf{n},\tau}^{\beta,\delta} &\leq \mathbb{E} \left[ \prod_{i=1}^{\mathbf{n}^{(1)}} \left( 1 + 2(\delta \sigma_i)^2 - \delta \tanh(\beta) \sigma_i \bar{\sigma}_i \right) \right] \\ &\leq \mathbb{E} \left[ \exp \left( \sum_{i=1}^{\mathbf{n}^{(1)}} 2(\delta \sigma_i)^2 - \delta \tanh(\beta) \sigma_i \bar{\sigma}_i \right) \right]. \end{aligned} \quad (3.6.19)$$

Let us define  $\sigma'_i := \sum_{j \in J_n(i)} \bar{\omega}_j \mathbf{1}_{\{(i,j) \notin \tau\}}$ , which is a small modification of  $\sigma_i$  (we removed  $\bar{\sigma}_i$  if it appears in the sum in  $\sigma_i$ ), so that we have

$$\begin{aligned} \sigma_i \bar{\sigma}_i &= \sigma'_i \bar{\sigma}_i + \left( \sum_{j \in J_n(i)} \bar{\omega}_j \mathbf{1}_{\{(i,j) \in \tau\}} \right) \left( \sum_{j=1}^{n^{(2)}} \bar{\omega}_j \mathbf{1}_{\{(i,j) \in \tau\}} \right) \\ &= \sigma'_i \bar{\sigma}_i + \mathbf{1}_{\{\exists j \in J_n(i); (i,j) \in \tau\}}. \end{aligned} \quad (3.6.20)$$

Hence, we get that

$$B_{\mathbf{n}, \tau}^{\beta, \delta} \leq e^{-\delta \tanh(\beta) |\tau \cap J_n|} \mathbb{E} \left[ \exp \left( 4 \sum_{i=1}^{n^{(1)}} \delta^2 \sigma_i^2 \right) \right]^{1/2} \mathbb{E} \left[ \exp \left( -2\delta \tanh(\beta) \sum_{i=1}^{n^{(1)}} \sigma'_i \bar{\sigma}_i \right) \right]^{1/2}. \quad (3.6.21)$$

where we applied Cauchy-Schwarz inequality. In the proof of Lemma 3.6.2 we already showed that the second factor is bounded. We claim that the third factor is also bounded.

**Lemma 3.6.3.** *There exists  $\delta_2 > 0$  such that if  $|\delta| \sqrt{\mathbf{n}^{(1)} \ell_n \log(\mathbf{n}^{(1)})} \leq \delta_2$ , and if  $\beta \leq 1$ , then*

$$1 \leq \mathbb{E} \left[ \exp \left( 2\delta \tanh(\beta) \sum_{i=1}^{n^{(1)}} \sigma'_i \bar{\sigma}_i \right) \right] \leq 2. \quad (3.6.22)$$

We prove this Lemma at the end of this section. Since  $\delta \sqrt{\mathbf{n}^{(1)} \ell_n \log(\mathbf{n}^{(1)})} \rightarrow 0$ , the assumption of the lemma is clearly satisfied by  $-\delta$  if  $\mathbf{n}^{(1)}$  is large enough: going back to (3.6.12) (and writing  $\tanh(\beta) > \beta/2$  for small  $\beta > 0$ ), we finally obtain

$$\bar{\mathbb{E}}_{\mathbf{n}, \delta} [Z_{\mathbf{n}}^q] \leq C_5 \mathbf{E} \left[ e^{-\delta \beta |\tau \cap J_n| / 2} \mathbf{1}_{\{n \in \tau\}} \right], \quad (3.6.23)$$

which is the same as [18, (4.25)]. Then we finish the proof of Proposition 3.6.1 by applying [18, Lemma 4.3] and Lemma 3.5.4 (we do not write the details here because it is a replica of [18]).  $\square$

*Proof of Lemma 3.6.3.* The left inequality is a straight consequence to Jensen's inequality, so we focus on the upper bound. We introduce some notations specific to this lemma. Let  $r := |\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|$  and denote  $(a_l, b_l) = \tau_l$ ,  $1 \leq l \leq r$  (in particular  $(a_r, b_r) = \tau_r = \mathbf{n}$ ). Moreover for all  $1 \leq l \leq r$ , denote  $J_n(a_l) = \llbracket c_l, d_l \rrbracket$ . Then we can write

$$\begin{aligned} \sum_{i=1}^{n^{(1)}} \sigma'_i \bar{\sigma}_i &= \sum_{l=1}^r \bar{\omega}_{b_l} \left( \sum_{j=c_l}^{d_l} \bar{\omega}_j \mathbf{1}_{\{(a_l, j) \notin \tau\}} \right) \\ &= \sum_{l=1}^r \bar{\omega}_{b_l} \left( \sum_{j=c_l}^{\min(d_l, b_l-1)} \bar{\omega}_j + \sum_{j=\max(b_l+1, c_l)}^{d_l} \bar{\omega}_j \right), \end{aligned} \quad (3.6.24)$$

where we parted the sum according to indices of  $\bar{\omega}$  before and after  $b_l$ . Let us define  $X_0 = Y_0 = 0$  and for all  $1 \leq t \leq r$ ,

$$X_t := \sum_{l=1}^t \bar{\omega}_{b_l} \left( \sum_{j=c_l}^{\min(d_l, b_l-1)} \bar{\omega}_j \right) \quad \text{and} \quad Y_t := \sum_{l=1}^t \bar{\omega}_{b_l} \left( \sum_{j=\max(b_l+1, c_l)}^{d_l} \bar{\omega}_j \right), \quad (3.6.25)$$

so that  $\sum_{i=1}^{\mathbf{n}^{(1)}} \sigma'_i \bar{\sigma}_i = X_r + Y_r$ , and by Cauchy-Schwarz inequality,

$$\mathbb{E} \left[ \exp \left( 2\delta \tanh(\beta) \sum_{i=-\ell_n}^{\mathbf{n}^{(1)} + \ell_n} \sigma'_i \bar{\sigma}_i \right) \right] \leq \mathbb{E} \left[ \exp(4\delta \tanh(\beta) X_r) \right]^{1/2} \mathbb{E} \left[ \exp(4\delta \tanh(\beta) Y_r) \right]^{1/2}. \quad (3.6.26)$$

We only treat the first factor, the second one being symmetric. We define  $\mathcal{F}_0$  the trivial  $\sigma$ -algebra, and for all  $1 \leq t \leq r$ ,

$$\mathcal{F}_t := \sigma(\{\bar{\omega}_j, 1 \leq j \leq b_{t+1} - 1\}), \quad (3.6.27)$$

with  $b_{r+1} := \mathbf{n}^{(2)} + 1$ : it is then easily checked that  $(X_t)_{0 \leq t \leq r}$  is a  $(\mathcal{F}_t)_{0 \leq t \leq r}$ -martingale. We also define a ‘‘truncated’’ version of  $X$ : for all  $w > 0$ ,  $X_0^{(w)} := 0$  and for all  $1 \leq t \leq r$ , we set

$$X_t^{(w)} := \sum_{l=1}^t \bar{\omega}_{b_l} \left( \sum_{j=c_l}^{\min(d_l, b_l - 1)} \bar{\omega}_j \right) \mathbf{1}_{\left\{ \left| \sum_{j=c_l}^{\min(d_l, b_l - 1)} \bar{\omega}_j \right| \leq w \sqrt{\ell_n \log(\mathbf{n}^{(1)})} \right\}}. \quad (3.6.28)$$

Notice that  $(X_t^{(w)})_{0 \leq t \leq r}$  is also a  $(\mathcal{F}_t)_{0 \leq t \leq r}$ -martingale, and it has bounded increments:  $|X_t^{(w)} - X_{t-1}^{(w)}| \leq w \sqrt{\ell_n \log(\mathbf{n}^{(1)})}$  for all  $1 \leq t \leq r$ . Therefore we deduce from the Azuma-Hoeffding inequality that for all  $u, w > 0$ ,

$$\mathbb{P} \left( |X_r^{(w)}| > u \sqrt{r \ell_n \log(\mathbf{n}^{(1)})} \right) \leq 2 \exp \left( - \frac{u^2}{2w^2} \right). \quad (3.6.29)$$

Moreover we have

$$\begin{aligned} \mathbb{P}(X_r \neq X_r^{(w)}) &= \mathbb{P} \left( \exists 1 \leq l \leq r; \left| \sum_{j=c_l}^{\min(d_l, b_l - 1)} \bar{\omega}_j \right| > w \sqrt{\ell_n \log(\mathbf{n}^{(1)})} \right) \\ &\leq \sum_{l=1}^r \mathbb{P} \left( \left| \sum_{j=c_l}^{\min(d_l, b_l - 1)} \bar{\omega}_j \right| > w \sqrt{\ell_n \log(\mathbf{n}^{(1)})} \right), \end{aligned} \quad (3.6.30)$$

so by Hoeffding’s inequality, and because  $|J_n(i)| \leq 2\ell_n$  and  $r \leq \mathbf{n}^{(1)}$ ,

$$\mathbb{P}(X_r \neq X_r^{(w)}) \leq \mathbf{n}^{(1)} \exp \left( - \frac{w^2}{4} \log(\mathbf{n}^{(1)}) \right). \quad (3.6.31)$$

We deduce from (3.6.31) and (3.6.29) that for all  $u, w > 0$ ,

$$\mathbb{P} \left( |X_r| > u \sqrt{r \ell_n \log(\mathbf{n}^{(1)})} \right) \leq 2 \exp \left( - \frac{u^2}{2w^2} \right) + \mathbf{n}^{(1)} \exp \left( - \frac{w^2}{4} \log(\mathbf{n}^{(1)}) \right), \quad (3.6.32)$$

in particular with  $w^2 = \frac{u\sqrt{2}}{\sqrt{\log(\mathbf{n}^{(1)})}}$ ,

$$\begin{aligned} \mathbb{P} \left( |X_r| > u \sqrt{r \ell_n \log(\mathbf{n}^{(1)})} \right) &\leq (2 + \mathbf{n}^{(1)}) \exp \left( - \frac{u \sqrt{\log(\mathbf{n}^{(1)})}}{2\sqrt{2}} \right) \\ &\leq \exp(-Cu), \end{aligned} \quad (3.6.33)$$



where  $C > 0$  is a constant sufficiently small such that this inequality holds for all  $u \geq 1$  and  $\mathbf{n}^{(1)} \in \mathbb{N}$ . Hence, (3.6.33) implies that  $\frac{|X_r|}{\sqrt{r\ell_{\mathbf{n}} \log(\mathbf{n}^{(1)})}}$  is dominated under some coupling by  $C' + Z$ , with  $C'$  a constant and  $Z$  an exponential random variable with parameter  $C$ . In particular,

$$\mathbb{E} \left[ \exp \left( 4\delta \tanh(\beta) X_r \right) \right] \leq \mathbb{E} \left[ \exp \left( 4|\delta|\beta \left( C' + \sqrt{r\ell_{\mathbf{n}} \log(\mathbf{n}^{(1)})} Z \right) \right) \right], \quad (3.6.34)$$

and this can be made arbitrarily small if  $|\delta|\beta\sqrt{r\ell_{\mathbf{n}} \log(\mathbf{n}^{(1)})} \leq |\delta|\sqrt{\mathbf{n}^{(1)}\ell_{\mathbf{n}} \log(\mathbf{n}^{(1)})}$  is small, which concludes the proof of the lemma.  $\square$

## Appendices to Chapter 3

### 3.A Estimates on the homogeneous gPS model

Here we prove Lemma 3.5.4 and Lemma 3.5.2.

*Proof of Lemma 3.5.4.* This Lemma is already proven for  $\alpha > 1$  in [18, Lemma 4.4], so we replicate the proof here for  $\alpha \in (0, 1]$  (notice that this proof follows the lines of [64, Prop. A.2.] in the context of the homogeneous PS model).

We decompose the partition function according to the number of renewal points until  $\mathbf{n}$ :

$$Z_{\mathbf{n},-u} = \sum_{j=1}^{\|\mathbf{n}\|} e^{-ju} \mathbf{P}(\tau_j = \mathbf{n}) = \sum_{j=1}^{\|\mathbf{n}\|^{\alpha_-}} e^{-ju} \mathbf{P}(\tau_j = \mathbf{n}) + \sum_{j=\|\mathbf{n}\|^{\alpha_-}+1}^{\|\mathbf{n}\|} e^{-ju} \mathbf{P}(\tau_j = \mathbf{n}). \quad (3.A.1)$$

For the first sum, (recall  $\alpha_- < \alpha \leq 1$ ) we use Proposition 3.B.1 to bound  $\mathbf{P}(\tau_j = \mathbf{n})$  by a constant times  $jK(\|\mathbf{n}\|)$ , both for  $\alpha < 1$  and  $\alpha = 1$  (for the latter, notice that  $b_j \ll \|\mathbf{n}\|$ ). We obtain

$$\sum_{j=1}^{\|\mathbf{n}\|^{\alpha_-}} e^{-ju} \mathbf{P}(\tau_j = \mathbf{n}) \leq C_1 \frac{K(\|\mathbf{n}\|)}{u^2} \sum_{j=1}^{\|\mathbf{n}\|^{\alpha_-}} ju^2 e^{-ju} \leq C_2 \frac{K(\|\mathbf{n}\|)}{u^2}, \quad (3.A.2)$$

where we used a Riemann-sum approximation of the last sum to get that it is bounded by a constant times  $\int_{\mathbb{R}_+} xe^{-x} dx = 1$  for any  $u \leq u_0$ . For the terms  $j > \|\mathbf{n}\|^{\alpha_-}$ , we simply bound  $j$  from below:

$$\sum_{j=\|\mathbf{n}\|^{\alpha_-}+1}^{+\infty} e^{-ju} \mathbf{P}(\tau_j = \mathbf{n}) \leq e^{-u\|\mathbf{n}\|^{\alpha_-}} \mathbf{P}(\mathbf{n} \in \tau), \quad (3.A.3)$$

and the proof is complete.  $\square$

*Proof of Lemma 3.5.2.* We first take care of the case  $\alpha \geq 1$  with very rough estimates: if  $\alpha > 1$ , recalling Theorem 3.1.2, there is some  $c_\alpha > 0$  such that  $F(0, h) \sim c_\alpha h$  as  $h \searrow 0$ . In particular  $h|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket$  is bounded uniformly in  $h \in (0, h_1)$  and  $1 \leq \|\mathbf{n}\| \leq 1/F(0, h)$ . Thus,

$$Z_{\mathbf{n},h} = \mathbf{E}[e^{h|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket}|_{\mathbf{n} \in \tau}] \leq C_3 \mathbf{P}(\mathbf{n} \in \tau). \quad (3.A.4)$$

Then we conclude with Proposition 3.B.3.

If  $\alpha = 1$ , then  $F(0, h) \sim L_\alpha(1/h)h$  as  $h \searrow 0$  where  $L_\alpha$  is slowly varying. Using the additional assumption  $\|\mathbf{n}\| \leq F(0, h)^{-(1-\varepsilon^2)}$ ,  $h|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|$  is also bounded (it decays to 0) and the former upper bound still holds. Then we conclude with Proposition 3.B.3.

Those rough estimates do not apply when  $\alpha \in (0, 1)$ , since we do not have uniform bounds on  $h|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|$ . We follow the line of proof of [64, Lem. 4.1] for the PS model. A first step in the proof consists in dealing with the indicator function in the partition function, to compare it with its free counterpart, see (3.A.9) below. Let us define

$$T_n := \left\{ \mathbf{i} \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket; \|\mathbf{i}\| \leq \frac{1}{2}\|\mathbf{n}\| \right\}, \quad (3.A.5)$$

the lower left half of the rectangle  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket$ , where we recall that  $\|\cdot\|$  is the  $L^1$  norm on  $\mathbb{N}^2$ . Because, conditionally on  $\mathbf{n} \in \tau$ , the time-reversed process  $\tilde{\tau}$  in  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket \setminus T_n$  starting from  $\mathbf{n}$  has same law as  $\tau$  in  $T_n$ , the partition function is bounded with a Cauchy-Schwarz inequality by

$$Z_{n,h} = \mathbf{E} \left[ e^{h|\tau \cap \llbracket \mathbf{1}, \mathbf{n} \rrbracket|} \mid \mathbf{n} \in \tau \right] \mathbf{P}(\mathbf{n} \in \tau) \leq \mathbf{E} \left[ e^{2h|\tau \cap T_n|} \mid \mathbf{n} \in \tau \right] \mathbf{P}(\mathbf{n} \in \tau), \quad (3.A.6)$$

(if  $\|\mathbf{n}\|$  is even the anti-diagonal  $\{\|\mathbf{i}\| = \|\mathbf{n}\|/2\}$  is counted twice, but the upper bound still holds). Let us define  $X_n := \sup\{\mathbf{i} \in T_n \cap \tau\}$  the last renewal point in  $T_n$  (we take the supremum for the natural order on  $\tau \subset \mathbb{N}^2$ : recall that  $\tau$  is strictly increasing on both coordinates). Because  $|\tau \cap T_n|$  and  $1_{\{\mathbf{n} \in \tau\}}$  are independent conditionally to  $\{X_n = \mathbf{i}\}$ , we can write:

$$\mathbf{E} \left[ e^{2h|\tau \cap T_n|} \mid \mathbf{n} \in \tau \right] = \sum_{\mathbf{i} \in T_n} \mathbf{E} \left[ e^{2h|\tau \cap T_n|} \mid X_n = \mathbf{i} \right] \mathbf{P}(X_n = \mathbf{i} \mid \mathbf{n} \in \tau). \quad (3.A.7)$$

then we can use the following Lemma, which is proven afterwards.

**Lemma 3.A.1.** *Assume  $\alpha \in (0, 1)$ . There exists  $C_4 > 0$  such that for any  $\mathbf{n} \in \mathbb{N}$  and  $\mathbf{i} \in T_n$ ,*

$$\mathbf{P}(X_n = \mathbf{i}, \mathbf{n} \in \tau) \leq C_4 L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} \mathbf{P}(X_n = \mathbf{i}). \quad (3.A.8)$$

Thanks to this Lemma and (3.A.7), the inequality (3.A.6) becomes

$$Z_{n,h} \leq \frac{C_4 L(\|\mathbf{n}\|)^{-1}}{\|\mathbf{n}\|^{2-\alpha}} \mathbf{E} \left[ e^{2h|\tau \cap T_n|} \right], \quad (3.A.9)$$

so Lemma 3.5.2 will be proven once we show that the above expectation is bounded by a constant, uniformly for  $\|\mathbf{n}\| \leq 1/F(0, h)$ . Notice that  $\|\tau\| = (\|\tau_i\|)_{i \geq 1}$  is a renewal process on  $\mathbb{N}$ , so we may write the upper bound

$$\mathbf{E} \left[ e^{2h|\tau \cap T_n|} \right] \leq \mathbf{E} \left[ \exp \left( 2h \sum_{i=1}^{\|\mathbf{n}\|/2} \mathbf{1}_{\{i \in \|\tau\|\}} \right) \right], \quad (3.A.10)$$

which is the partition function of a homogeneous PS model with underlying univariate renewal  $\|\tau\|$ , which easily verifies  $\mathbf{P}(\|\tau_1\| = k) \sim L(k)k^{-(1+\alpha)}$  as  $k \rightarrow +\infty$ . The right side of (3.A.10) has already been studied in [64, Lemma 4.1] when  $\alpha \in (0, 1)$ , and we therefore get that the expectation is bounded by a constant uniformly in  $\|\mathbf{n}\| \leq 1/F(0, h)$ .  $\square$

*Proof of Lemma 3.A.1.* Fix  $\mathbf{n}, \mathbf{i} \in \mathbb{N}^2$ . Recall and the definitions of  $T_n$  in (3.A.5) and  $X_n$ , and that we assumed  $\alpha \in (0, 1)$ . We write:

$$\begin{aligned} \mathbf{P}(X_n = \mathbf{i}, \mathbf{n} \in \tau) &= \mathbf{P}(\mathbf{i} \in \tau) \mathbf{P}\left(\|\tau_1\| > \frac{1}{2}\|\mathbf{n}\| - \|\mathbf{i}\|, \mathbf{n} - \mathbf{i} \in \tau\right) \\ &= \mathbf{P}(\mathbf{i} \in \tau) \sum_{\mathbf{j} \in Q_i^n} \mathbf{P}(\tau_1 = \mathbf{j}) \mathbf{P}(\mathbf{n} - \mathbf{i} - \mathbf{j} \in \tau), \end{aligned} \quad (3.A.11)$$

with

$$Q_i^n := \left\{ \mathbf{j} \in \mathbb{N}^2; \mathbf{j} \preccurlyeq \mathbf{n} - \mathbf{i}, \|\mathbf{j}\| > \frac{1}{2}\|\mathbf{n}\| - \|\mathbf{i}\| \right\}. \quad (3.A.12)$$

Then, we can use Proposition 3.B.2 to get that  $\mathbf{P}(\mathbf{n} - \mathbf{i} - \mathbf{j} \in \tau)$  is bounded by a constant times  $L(\|\mathbf{n} - \mathbf{i} - \mathbf{j}\|)^{-1} \|\mathbf{n} - \mathbf{i} - \mathbf{j}\|^{-(2-\alpha)}$ . We get that the sum in (3.A.11) is bounded by a constant times

$$\begin{aligned} &\sum_{\mathbf{j} \in Q_i^n} \mathbf{P}(\tau_1 = \mathbf{j}) L(\|\mathbf{n} - \mathbf{i} - \mathbf{j}\|)^{-1} \|\mathbf{n} - \mathbf{i} - \mathbf{j}\|^{-(2-\alpha)} \\ &\leq C_5 L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} \sum_{\mathbf{j} \in Q_i^n, \|\mathbf{j}\| \leq \|\mathbf{n}\|/4} \mathbf{P}(\tau_1 = \mathbf{j}) \\ &\quad + C_5 L(\|\mathbf{n}\|) \|\mathbf{n}\|^{-(2+\alpha)} \sum_{\ell \leq \|\mathbf{n}\|} L(\|\ell\|)^{-1} \|\ell\|^{-(2-\alpha)}, \end{aligned}$$

where we decomposed the sum according to whether  $\|\mathbf{j}\| \leq \|\mathbf{n}\|/4$  or not. We used that if  $\|\mathbf{j}\| \leq \|\mathbf{n}\|/4$  then  $\|\mathbf{n} - \mathbf{i} - \mathbf{j}\| \geq \|\mathbf{n}\|/4$ , and if  $\|\mathbf{j}\| > \|\mathbf{n}\|/4$  then we can bound  $\mathbf{P}(\mathbf{j} \in \tau)$  uniformly thanks to (3.1.1), and we wrote  $\ell := \mathbf{n} - \mathbf{i} - \mathbf{j}$ .

Using that the last sum above is bounded by a constant times  $L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^\alpha$ , we therefore get that (3.A.11) is bounded by a constant times

$$\begin{aligned} &\mathbf{P}(\mathbf{i} \in \tau) L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} \left( \mathbf{P}\left(\|\tau_1\| > \frac{1}{2}\|\mathbf{n}\| - \|\mathbf{i}\|\right) + L(\|\mathbf{n}\|) \|\mathbf{n}\|^{-\alpha} \right) \\ &\leq C_6 L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} \mathbf{P}(\mathbf{i} \in \tau) \mathbf{P}\left(\|\tau_1\| > \frac{1}{2}\|\mathbf{n}\| - \|\mathbf{i}\|\right) \\ &\leq C_6 L(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)} \mathbf{P}(X_n = \mathbf{i}), \end{aligned}$$

which concludes the proof.  $\square$

### 3.B Some properties of bivariate renewals

We provide here some estimates on the bivariate renewal  $\tau$  that we use (recall its inter-arrival distribution (3.1.1)). They can be found in [18, Appendix A] or in [17] in a more general setting, and rely on the fact that  $\tau$  is in the domain of attraction of a  $\min(\alpha, 2)$ -stable distribution. We define the *scaling sequence*  $a_n$ ,

$$a_n := \psi(n)n^{1/\min(\alpha,2)}, \quad (3.B.1)$$

where  $\psi$  is some slowly varying function (we do not detail it here, see [17, 18]; notice that if  $\alpha > 2$ ,  $\psi$  is a constant). We also define the *recentering sequence*  $b_n$ ,

$$b_n := 0 \text{ if } \alpha \in (0, 1), \quad b_n := n\mu(a_n) \text{ if } \alpha = 1, \quad b_n := n\mu \text{ if } \alpha > 1. \quad (3.B.2)$$

Here  $\mu(x) = \mathbf{E}[\tau_1^{(1)}1_{\{\tau_1^{(1)} \leq x\}}]$  is the truncated moment, and  $\mu = \lim_{x \rightarrow +\infty} \mu(x)$ . When  $\alpha = 1$ , notice that either  $\mu < +\infty$  or  $\mu(x) \rightarrow +\infty$  as a slowly varying function.

**Proposition 3.B.1** ([17], Thm. 2.4). *There exists some  $C_1 > 0$  such that*

$$\mathbf{P}(\tau_j = \mathbf{n} + b_j \mathbf{1}) \leq C_1 j \mathbf{P}(\tau_1 = \mathbf{n}), \quad (3.B.3)$$

for any  $j \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^2$  such that  $\|\mathbf{n}\| \geq a_j$ .

**Proposition 3.B.2** ([17], Thms. 3.1-4.1). *Assume  $\alpha \in (0, 1)$ . There is some constant  $C > 0$  such that for all  $\mathbf{n} \in \mathbb{N}^2$ ,*

$$\mathbf{P}(\mathbf{n} \in \tau) \leq CL(\|\mathbf{n}\|)^{-1} \|\mathbf{n}\|^{-(2-\alpha)}. \quad (3.B.4)$$

Note that this upper bound is sharp when  $\mathbf{n}^{(1)}$  and  $\mathbf{n}^{(2)}$  are of the same order (*i.e.* when  $\mathbf{n}$  is close to the diagonal: we refer to [154, Cor. 3-B] for a precise statement).

**Proposition 3.B.3** ([17], Thms. 3.3-3.4 and 4.2-4.3). *Assume  $\alpha \geq 1$ , then there is a constant  $C > 0$  such that for any  $\mathbf{n} \in \mathbb{N}^2$ ,*

$$\mathbf{P}(\mathbf{n} \in \tau) \leq \frac{C}{\mu(n)a_{n/\mu(n)}}, \quad (3.B.5)$$

with  $n = \min(\mathbf{n}^{(1)}, \mathbf{n}^{(2)})$ . When  $\alpha > 1$ ,  $\mu(n)$  can be replaced by 1.

We stress that this upper bound is sharp when  $\mathbf{n}$  is close to the diagonal, we refer to [17, Thms. 3.3-3.4] for the precise statements. Also, notice that  $n \mapsto \mu(n)a_{n/\mu(n)}$  is regularly varying with index  $1/\min(\alpha, 2)$ .

Let us also recall some results on intersections of renewals, either univariate or bivariate.

**Proposition 3.B.4.**

1. Let  $\tau$  be a renewal in  $\mathbb{N}$ , with  $\mathbf{P}(\tau_1 = n) = L(n)/n^{1+\alpha}$ , for  $\alpha > 0$  and  $L$  slowly varying. Let  $\tau'$  be an independent copy of  $\tau$ . Then  $\tau \cap \tau'$  is also a renewal process, and it is terminating a.s. if  $\sum_{n \geq 1} \frac{n^{2\alpha-2}}{L(n)^2} < \infty$  (in particular  $\alpha \leq 1/2$  is necessary, while  $\alpha < 1/2$  is sufficient). Otherwise it is persistent a.s..
2. Let  $\tau$  be a renewal in  $\mathbb{N}^2$ , with inter-arrival distribution as in (3.1.1). Let  $\tau'$  be an independent copy of  $\tau$ . Then  $\tau \cap \tau'$  is also a renewal process, and it is terminating a.s. if  $\alpha < 1$ , or if  $\alpha = 1$  and  $\sum_{n \geq 1} \frac{1}{nL(n)\mu(n)} < \infty$ . Otherwise it is persistent a.s..

This comes from the following observation:  $\tau \cap \tau'$  (resp.  $\tau \cap \tau'$ ) is a univariate (resp. bivariate) renewal process, so  $|\tau \cap \tau'|$  (resp.  $|\tau \cap \tau'|$ ) is either infinite a.s., or a finite geometric variable. So one only has to estimate  $\mathbf{E}[|\tau \cap \tau'|]$  (resp.  $\mathbf{E}[|\tau \cap \tau'|]$ ) to determine if the intersection is persistent or terminating. For univariate processes, those estimate can be obtained thanks to the asymptotic behavior of the renewal mass function  $\mathbf{P}(i \in \tau)$  (see [67] if  $\alpha \in (0, 1)$ , [70] if  $\alpha = 1$ , the case  $\alpha > 1$  being simply the renewal theorem). Regarding the bivariate case, it has recently been treated in [18, Prop. A.3 and Rem. A.7], thanks to the estimates on the renewal mass function  $\mathbf{P}(i \in \tau)$  collected in [17].

### 3.C Computation of the second moment in the Gaussian case

In this appendix we compute the upper bound of the second moment of the partition function under the assumption that  $\widehat{\omega}, \bar{\omega}$  have a centered Gaussian distribution, for  $\beta \in (0, \beta_1)$  sufficiently small (recall Remark 3.3.5). We can assume without loss of generality that  $\widehat{\omega}, \bar{\omega} \sim \mathcal{N}(0, 1)$  (by adjusting  $\beta$ ). Recalling our computations from Section 3.3.2 and the decomposition from Proposition 3.3.4, we have to derive the correlations along a chain of points.

**Proposition 3.C.1.** *Let  $(X_k)_{k \geq 1}$  be a sequence of i.i.d. variables with distribution  $\mathcal{N}(0, 1)$ . Then we have for all  $\beta \leq 1/2, \ell \in \mathbb{N}$ ,*

$$\mathbb{E} \left[ e^{\beta(X_1 X_2 + X_2 X_3 + \dots + X_\ell X_{\ell+1})} \right] = \xi_1 \xi_2 \dots \xi_\ell, \tag{3.C.1}$$

where  $\xi_0 = 1$  and for all  $k \geq 0, \xi_{k+1} = (1 - \beta^2 \xi_k^2)^{-1/2}$ .

Note that the assumption  $\beta \leq 1/2$  is required for  $\xi_k$  to be well-defined for all  $k \geq 1$ .

*Proof.* The proof relies on the following identity: if  $X \sim \mathcal{N}(0, 1)$ , then for all  $t \in \mathbb{R}, \beta < 1$  and  $k \geq 0$ ,

$$\mathbb{E} \left[ e^{\beta t X} e^{\beta^2 \xi_k^2 X^2 / 2} \right] = \xi_{k+1} e^{\beta^2 \xi_{k+1}^2 t^2 / 2}, \tag{3.C.2}$$

(this follows from a direct computation). Hence for any  $\ell \geq 1$ , we condition the expectation over  $(X_1, \dots, X_\ell)$  to write

$$\begin{aligned} \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell} X_k X_{k+1}}\right] &= \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell-1} X_k X_{k+1}} \mathbb{E}\left[e^{\beta X_\ell X_{\ell+1}} \mid (X_1, \dots, X_\ell)\right]\right] \\ &= \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell-1} X_k X_{k+1}} e^{\beta^2 X_\ell^2 / 2}\right]. \end{aligned}$$

Then, by conditioning over  $(X_1, \dots, X_{\ell-1})$  and applying (3.C.2), we obtain by induction,

$$\begin{aligned} \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell} X_k X_{k+1}}\right] &= \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell-2} X_k X_{k+1}} \mathbb{E}\left[e^{\beta X_{\ell-1} X_\ell} e^{\beta^2 X_\ell^2 / 2} \mid (X_1, \dots, X_{\ell-1})\right]\right] \\ &= \xi_1 \mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell-2} X_k X_{k+1}} e^{\beta^2 \xi_1^2 X_{\ell-1}^2 / 2}\right] \\ &= \xi_1 \dots \xi_{\ell-1} \mathbb{E}\left[e^{\beta^2 \xi_{\ell-1}^2 X_1^2 / 2}\right] = \prod_{k=1}^{\ell} \xi_k, \end{aligned}$$

which concludes the proof.  $\square$

When  $\beta \leq 1/2$ , one can prove that the sequence  $(\xi_k)_{k \geq 1}$  is non-decreasing and that it converges to  $\xi_\infty := \frac{1}{\beta\sqrt{2}} \sqrt{1 - \sqrt{1 - 4\beta^2}}$ . Moreover the application  $f : x \mapsto (1 - \beta^2 x^2)^{-1/2}$  is convex, hence we have for all  $k \geq 1$ ,

$$\xi_{k+1} - \xi_k \leq f'(\xi_\infty)(\xi_k - \xi_{k-1}), \quad (3.C.3)$$

with  $f'(\xi_\infty) = \beta^2 \xi_\infty^4 \leq C_1 \beta^2$  for some  $C_1 > 0$ . Thus we can prove by induction,

$$\xi_k \leq \xi_1 + (\xi_1 - 1)(C_1 \beta^2 + (C_1 \beta^2)^2 + \dots + (C_1 \beta^2)^{k-1}). \quad (3.C.4)$$

Assuming  $\beta < \beta_1 < C_1^{-1/2}$ , there exist  $C_2 > 0$  such that for all  $k \geq 1$ ,

$$\sum_{i=1}^k (C_1 \beta^2)^i \leq \frac{C_1 \beta^2}{1 - C_1 \beta^2} \leq C_2 \beta^2. \quad (3.C.5)$$

Moreover a Taylor expansion yields  $\xi_1 - 1 = \frac{1}{2}\beta^2 + o(\beta^2)$ . Hence there exist  $\beta_1 > 0$  and  $C_3 > 0$  such that for all  $\beta \in (0, \beta_1)$  and  $k \geq 1$ , one has

$$\frac{\xi_k}{\xi_1} \leq 1 + C_3 \beta^4 \leq e^{C_3 \beta^4}. \quad (3.C.6)$$

Finally, noticing that  $e^{\lambda(\beta)} = \xi_1$ , we conclude

$$\mathbb{E}\left[e^{\beta \sum_{k=1}^{\ell} X_k X_{k+1} - \ell \lambda(\beta)}\right] = \prod_{k=1}^{\ell} \frac{\xi_k}{\xi_1} \leq e^{C_3 \ell \beta^4}. \quad (3.C.7)$$

The left-hand side in (3.C.7) corresponds exactly to the contribution of a chain of length  $\ell$  to the second moment of the partition function. Going back to (3.3.19) and recalling that the expectation with respect to intersection points gives a contribution  $e^{|\nu_n|(\lambda(2\beta)-2\lambda(\beta))}$  and that the isolated points do not contribute, we get that

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta,q,free} \right)^2 \right] \leq \mathbf{E}_{(\tau,\tau')} \left[ e^{|\nu_n|(\lambda(2\beta)-2\lambda(\beta))} \prod_{m \in \mathbb{N}} e^{C_3 \beta^4 |\sigma_{n,m}|} \right].$$

Applying Cauchy-Schwarz inequality, we obtain

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta,q,free} \right)^2 \right] \leq \mathbf{E}_{(\tau,\tau')} \left[ e^{2(\lambda(2\beta)-2\lambda(\beta))|\nu_n|} \right]^{1/2} \mathbf{E}_{(\tau,\tau')} \left[ e^{4C_3 \beta^4 (|\tau_{\leq n}^{(1)} \cap \tau'_{\leq n}{}^{(1)}| + |\tau_{\leq n}^{(2)} \cap \tau'_{\leq n}{}^{(2)}|)} \right]^{1/2}, \quad (3.C.8)$$

where we also used that  $\sum_m |\sigma_{n,m}| = |\mathfrak{S}_n| \leq 2 (|\tau_{\leq n}^{(1)} \cap \tau'_{\leq n}{}^{(1)}| + |\tau_{\leq n}^{(2)} \cap \tau'_{\leq n}{}^{(2)}|)$ , recall (3.3.22). Now, for  $\alpha < 1$ , Proposition 3.B.4-(2) gives that  $\nu = \tau \cap \tau'$  is a.s. finite; in fact,  $|\nu|$  is a geometric random variable. Hence, if  $\beta$  is small enough the first factor in the r.h.s. of (3.C.8) is bounded by 2. Applying Cauchy-Schwarz inequality for the other factor, we therefore end up with

$$\mathbb{E} \left[ \left( Z_{n,0}^{\beta,q,free} \right)^2 \right] \leq 2 \mathbf{E}_{(\tau,\tau')} \left[ e^{8C_3 \beta^4 |\tau_{\leq n}^{(1)} \cap \tau'_{\leq n}{}^{(1)}|} \right]^{1/2},$$

and one can remove the exponent 1/2 since the expectation on the r.h.s. is larger than 1. This proves the inequality (3.3.28) claimed in Section 3.3.3 when  $\alpha < 1$ , and thus gives the expected  $n_\beta$  and upper bound on the critical point shift.

If  $\alpha \geq 1$ , one concludes by proving that the first factor in the r.h.s. of (3.C.8) remains bounded as long as  $\beta^2 N^{1/\min(\alpha,2)} \sim 1$ , and the second factor does as long as  $\beta^4 N \sim 1$  (up to slowly varying factors). Hence we have  $n_\beta \sim \beta^{-4}$ , which gives the expected upper bound on the critical point shift. All the required estimates can be found in [18, 70], so we do not write all the details here.





# Scaling limit of the Generalized Poland-Scheraga model

This Chapter is a work in progress with Quentin Berger; hence we warn the reader that the claims, proofs and notation displayed here are neither perfect or final.

## 4.1 Introduction

### 4.1.1 The generalized Poland-Scheraga model

Let us quickly introduce some definitions for this chapter, some from previous Chapter 3 and a few new ones.

We will denote elements of  $\mathbb{N}^2$ ,  $\mathbb{R}^2$  with bold characters, and elements of  $\mathbb{N}$ ,  $\mathbb{R}$  with plain characters (in particular we note  $\mathbf{0} := (0, 0)$  and  $\mathbf{1} := (1, 1)$ ); moreover for  $\mathbf{t} \in \mathbb{R}^2$ ,  $t^{(a)}$  will denote its projection on its  $a$ -th coordinate,  $a \in \{1, 2\}$ , and when there is no risk of confusion, we may also write more simply  $\mathbf{t} = (t_1, t_2)$  and so on. Let  $\boldsymbol{\tau} = (\tau_k)_{k \geq 0}$  be a bivariate renewal process, with  $\tau_0 = \mathbf{0}$  and inter-arrival distribution

$$\mathbf{P}(\boldsymbol{\tau}_1 = (\ell_1, \ell_2)) := K(\ell_1 + \ell_2) = \frac{L(\ell_1 + \ell_2)}{(\ell_1 + \ell_2)^{2+\alpha}}, \quad \forall \boldsymbol{\ell} = (\ell_1, \ell_2) \in \mathbb{N}^2, \quad (4.1.1)$$

with  $\mathbf{P}(|\boldsymbol{\tau}_1| < +\infty) = 1$ . With a slight abuse of notation, we also interpret  $\boldsymbol{\tau}$  as a set  $\{\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots\}$  (we will always omit  $\boldsymbol{\tau}_0$ ).

Let  $\hat{\omega} = (\hat{\omega}_i)_{i \geq 1}$  and  $\bar{\omega} = (\bar{\omega}_i)_{i \geq 1}$  be two independent sequences of i.i.d. random variables, whose laws are denoted  $\hat{\mathbb{P}}$  and  $\bar{\mathbb{P}}$  respectively —let us also denote  $\mathbb{P} := \hat{\mathbb{P}} \otimes \bar{\mathbb{P}}$ . For simplicity's sake we will assume  $\hat{\mathbb{P}} = \bar{\mathbb{P}}$ , and for  $\mathbf{i} \in \mathbb{N}^2$ , we denote  $\omega_i = \omega_{i^{(1)}, i^{(2)}} :=$

$f(\widehat{\omega}_{i(1)}, \widehat{\omega}_{i(2)})$ , where  $f(\cdot, \cdot)$  is a function describing the interactions between the monomers; in the remainder of the Chapter we will assume  $f(x, y) = xy$  for  $x, y \in \mathbb{R}$ . We stress again that  $\omega := (\omega_i)_{i \in \mathbb{N}^2}$  is an ergodic, strongly correlated field. We assume  $\lambda(\beta) := \log \mathbb{E}[e^{\beta \omega_1}] < +\infty$  for  $\beta \in [0, \beta_0)$ , which can be ensured by assuming  $\mathbb{E}[e^{\frac{1}{2}\beta \widehat{\omega}_1^2}] < +\infty$  for  $\beta \in [0, \beta_0)$  (in particular it is satisfied by Gaussian or bounded laws). For a fixed realization of  $\omega$  (*quenched disorder*), we define, for  $\beta \geq 0$  (the disorder strength) and  $h \in \mathbb{R}$  (the pinning potential), the following *polymer measures*: for any  $\mathbf{n} \in \mathbb{N}^2$ , representing the respective lengths of the strands,

$$\frac{d\mathbf{P}_{\mathbf{n},h}^{\beta,\omega,q}}{d\mathbf{P}}(\tau) := \frac{1}{Z_{\mathbf{n},h}^{\beta,\omega,q}} \exp\left(\sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}\right) \mathbf{1}_{\{\mathbf{n} \in \tau\}}, \quad (4.1.2)$$

where

$$Z_{\mathbf{n},h}^{\beta,\omega,q} := \mathbf{E}\left[\exp\left(\sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}\right) \mathbf{1}_{\{\mathbf{n} \in \tau\}}\right] \quad (4.1.3)$$

is the partition function of the system. We used the notation  $\llbracket \mathbf{1}, \mathbf{n} \rrbracket := \llbracket 1, n_1 \rrbracket \times \llbracket 1, n_2 \rrbracket$ , and  $\llbracket 1, n \rrbracket := [1, n] \cap \mathbb{N}$ . This corresponds to giving a reward (or penalty if it is negative)  $\beta \omega_i + h$  if  $\mathbf{i} = (i_1, i_2) \in \tau$ , that is monomer  $i_1$  of the first strand is paired with monomer  $i_2$  of the second strand. The term  $-\lambda(\beta)$  is only present for renormalization purposes, and even though  $Z_{\mathbf{n},h}^{\beta,\omega,q}$  depends on the realization of  $\omega$ , we will drop it in the notation for conciseness. We also define the *conditioned* and *free* partition functions for  $\mathbf{n} \in \tau$ , respectively by

$$Z_{\mathbf{n},h}^{\beta,\text{cond}} := \mathbf{E}\left[\exp\left(\sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}\right) \middle| \mathbf{n} \in \tau\right], \quad (4.1.4)$$

and

$$Z_{\mathbf{n},h}^{\beta,\text{free}} := \mathbf{E}\left[\exp\left(\sum_{i \in \llbracket \mathbf{1}, \mathbf{n} \rrbracket} (\beta \omega_i - \lambda(\beta) + h) \mathbf{1}_{\{i \in \tau\}}\right)\right]. \quad (4.1.5)$$

We finally introduce some notation for the remainder of the Chapter. We define some orders  $\prec, \preceq$  on  $\mathbb{R}^2$ , by

$$\begin{aligned} \mathbf{x} \prec \mathbf{y} & \text{ if } \mathbf{x}^{(1)} < \mathbf{y}^{(1)}, \mathbf{x}^{(2)} < \mathbf{y}^{(2)}, \\ \text{and } \mathbf{x} \preceq \mathbf{y} & \text{ if } \mathbf{x}^{(1)} \leq \mathbf{y}^{(1)}, \mathbf{x}^{(2)} \leq \mathbf{y}^{(2)}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^2. \end{aligned}$$

For  $\mathbf{t} \succcurlyeq \mathbf{s} \succcurlyeq \mathbf{0}$ , we use the notation  $[\mathbf{s}, \mathbf{t}]$  for the rectangle  $[\mathbf{s}^{(1)}, \mathbf{t}^{(1)}] \times [\mathbf{s}^{(2)}, \mathbf{t}^{(2)}]$  (and similarly  $[\mathbf{s}, \mathbf{t}] := [\mathbf{s}^{(1)}, \mathbf{t}^{(1)}] \times [\mathbf{s}^{(2)}, \mathbf{t}^{(2)}]$  and so on), and  $\llbracket \mathbf{s}, \mathbf{t} \rrbracket := [\mathbf{s}, \mathbf{t}] \cap \mathbb{N}^2$ . For  $s, t \in \mathbb{R}$ , we define  $s \wedge t = \min(s, t)$ , resp.  $s \vee t = \max(s, t)$ , and for  $\mathbf{s}, \mathbf{t} \in \mathbb{R}^2$ ,

$$\begin{aligned} \mathbf{s} \wedge \mathbf{t} & := (\mathbf{s}^{(1)} \wedge \mathbf{t}^{(1)}, \mathbf{s}^{(2)} \wedge \mathbf{t}^{(2)}), \\ \text{resp. } \mathbf{s} \vee \mathbf{t} & := (\mathbf{s}^{(1)} \vee \mathbf{t}^{(1)}, \mathbf{s}^{(2)} \vee \mathbf{t}^{(2)}). \end{aligned}$$

For  $\mathbf{s} \in \mathbb{R}^2$ , we denote  $\lfloor \mathbf{s} \rfloor = (\lfloor s_1 \rfloor, \lfloor s_2 \rfloor)$ . Finally, we will say that  $\mathbf{s}, \mathbf{t} \in \mathbb{R}^2$  are *aligned* if they are on the same line or column (that is  $\mathbf{s}^{(1)} = \mathbf{t}^{(1)}$  or  $\mathbf{s}^{(2)} = \mathbf{t}^{(2)}$ ), and we will write  $\mathbf{s} \leftrightarrow \mathbf{t}$ ; otherwise we will write  $\mathbf{s} \nleftrightarrow \mathbf{t}$ .

### 4.1.2 Disorder relevance via chaos expansion and intermediate disorder scaling limit

As we advertised in Chapter 3, our aim is to complete our results on disorder (ir)-relevance for the gPS model, by taking inspiration from [3, 44]. In those papers the authors proved for some disordered systems (notably the disordered pinning model in [44]) that, by choosing a disorder intensity  $\beta_n$  decaying to 0 as  $n \rightarrow \infty$ , it was possible to exhibit an *intermediate disorder* regime, laying in-between the homogeneous ( $\beta = 0$ ) and disordered ( $\beta > 0$  constant) ones. A specificity of this approach is that it only seems to yield a non-trivial, random limit when disorder is relevant in the meaning of Harris, i.e. when a shift of the critical point (and smoothing) occurs.

Thereby, we will try this approach in our setting of a gPS model with correlated disorder, more specifically for  $1/2 < \alpha < 1$ : in that case we proved in Chapter 3 that disorder (ir)-relevance depends on the choice of disorder distribution  $\mathbb{P}$ , therefore we expect to find a random limit only when  $\mathbb{P} \neq \mathbb{P}_{\pm x}$ , where we define  $\mathbb{P}_{\pm x}(\widehat{\omega}_1 = x) = \mathbb{P}_{\pm x}(\widehat{\omega}_1 = -x) = 1/2$  (and same law for  $\bar{\omega}_1$ ).

We define  $\mathfrak{P}$  the set of disorder distributions  $\mathbb{P}$  satisfying all our previous assumptions. For  $\mathbb{P} \in \mathfrak{P}$ , we denote  $m_k := \mathbb{E}[\widehat{\omega}_1^k]$ ,  $k \in \mathbb{N}$  (in particular  $\mathbb{E}[\omega_1^k] = m_k^2$ ), and we note that distributions  $\mathbb{P}_{\pm x}$  are characterized by the identities  $m_1 = 0$  and  $m_4 - m_2^2 = 0$ —that is  $\widehat{\omega}^2$  is constant  $\mathbb{P}$ -a.s.. To match our expectations, we partition  $\mathfrak{P}$  into three sets,

$$\begin{aligned} \mathfrak{P}_0 &:= \left\{ \mathbb{P}_{\pm x}; x > 0 \right\}, \\ \mathfrak{P}_1 &:= \left\{ \mathbb{P} \in \mathfrak{P}; m_1 \neq 0 \right\}, \\ \mathfrak{P}_2 &:= \left\{ \mathbb{P} \in \mathfrak{P}; m_1 = 0, m_4 > m_2^2 \right\}, \end{aligned} \quad (4.1.6)$$

and we suspect to find different results depending on which set  $\mathbb{P}$  lies in.

#### Heuristics of the chaos expansion

As in [3, 44], we look for scaling limits of the partition functions by computing polynomial expansions, starting with the free variant (4.1.5) for convenience. Let us define  $\zeta_i = \zeta_i(\beta) := e^{\beta\omega_i - \lambda(\beta)} - 1$ , so that for  $\mathbf{t} \in (\mathbb{R}_+^*)^2$  and  $n \in \mathbb{N}$  such that  $n\mathbf{t} \in \mathbb{N}^2$ ,

$$\begin{aligned} Z_{n\mathbf{t}, h}^{\beta, \text{free}} &= \mathbf{E} \left[ \prod_{i \in \llbracket 1, n\mathbf{t} \rrbracket} \left( 1 + (e^h \zeta_i + e^h - 1) 1_{\{i \in \tau\}} \right) \right] \\ &= 1 + \sum_{k=1}^{(n\mathbf{t}^{(1)}) \wedge (n\mathbf{t}^{(2)})} \sum_{\mathbf{0} = i_0 < i_1 < \dots < i_k \leq n\mathbf{t}} \prod_{l=1}^k \left( (e^h \zeta_{i_l} + e^h - 1) u(\mathbf{i}_l - \mathbf{i}_{l-1}) \right) \end{aligned} \quad (4.1.7)$$

where we denoted  $u(\mathbf{i}) := \mathbf{P}(\mathbf{i} \in \tau)$  the *renewal function*. Let us focus on the convergence of the term  $k = 1$ , for simplicity. As  $h \rightarrow 0$ , it is

$$(1 + o(1)) \sum_{i \in \llbracket 1, n\mathbf{t} \rrbracket} \zeta_i u(\mathbf{i}) + (1 + o(1)) \sum_{i \in \llbracket 1, n\mathbf{t} \rrbracket} h u(\mathbf{i}). \quad (4.1.8)$$

Regarding the second term, we need to estimate the renewal mass function  $u(\mathbf{i})$  as  $\mathbf{i} \rightarrow \infty$ . When  $\alpha \in (0, 1)$ , it is given in [154].

**Theorem 4.1.1** ([154]). *Assume  $0 < \alpha < 1$  in (4.1.1). Then for  $\mathbf{s} \in (\mathbb{R}_+^*)^2$ , we have*

$$\lim_{n \rightarrow +\infty} n^{2-\alpha} L(n) \mathbf{P}(\lfloor n\mathbf{s} \rfloor \in \tau) = \varphi(\mathbf{s}),$$

for some continuous function  $\varphi : (\mathbb{R}_+^*)^2 \rightarrow \mathbb{R}_+$ . Writing  $\mathbf{s}$  in the polar form  $\mathbf{s} = re^{i\theta}$ , we get that  $\varphi(\mathbf{s}) = r^{\alpha-2} a(\theta)$ , for some continuous function  $a : [0, \pi/2] \rightarrow \mathbb{R}_+$ , which is equal to 0 at  $\theta = 0$  and  $\theta = \pi/2$ .

This theorem and a Riemann sum approximation imply that

$$\frac{L(n)}{n^\alpha} \sum_{\mathbf{i} \in \llbracket \mathbf{1}, n\mathbf{t} \rrbracket} u(\mathbf{i}) \xrightarrow{n \rightarrow +\infty} \int_{[0, \mathbf{t}]} \varphi(\mathbf{s}) d\mathbf{s},$$

and the integral is finite because  $\varphi(\mathbf{s}) \leq c \|\mathbf{s}\|^{\alpha-2}$  with  $0 < \alpha < 1$ . Hence, in order to make the second term converge in (4.1.8), we have to take  $h = h_n := \hat{h} L(n) n^{-\alpha}$  for some  $\hat{h} \in \mathbb{R}$ .

Regarding the first term in (4.1.8), the main idea in [3, 44] is to prove that the field  $(\zeta_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^2}$ , with an appropriate scaling, converges to a white noise; thus the sum converges to an integral of  $\varphi$  against the white noise. However this relies on the assumption that the disorder field is i.i.d., which is not the case in our setting since  $\zeta$ , as  $\omega$ , is strongly correlated on each line and column of  $\mathbb{N}^2$ . We overcome this hurdle by proving a convergence to another Gaussian random field  $\mathcal{M}$  which encapsulates the correlation structure of  $\zeta$ , and by defining a 2-dimensional integral against that field (similarly to [41]).

To that end, define for  $\mathbf{s} \in (\mathbb{R}_+)^2$  and  $n \in \mathbb{N}$ ,

$$M_n(\mathbf{s}) := \sum_{\mathbf{i} \in \llbracket \mathbf{1}, \lfloor n\mathbf{s} \rfloor \rrbracket} \zeta_{\mathbf{i}}, \quad (4.1.9)$$

with the convention  $M_n(\mathbf{s}) = 0$  if  $\mathbf{s}^{(1)} < 1/n$  or  $\mathbf{s}^{(2)} < 1/n$ ; and let  $\mathcal{M}$  be a Gaussian field on  $(\mathbb{R}_+)^2$  with zero-mean and covariance matrix

$$K(\mathbf{u}, \mathbf{v}) := \left( \mathbf{u}^{(1)} \wedge \mathbf{v}^{(1)} \right) \left( \mathbf{u}^{(2)} \wedge \mathbf{v}^{(2)} \right) \left( \mathbf{u}^{(1)} \vee \mathbf{v}^{(1)} + \mathbf{u}^{(2)} \vee \mathbf{v}^{(2)} \right), \quad \mathbf{u}, \mathbf{v} \in (\mathbb{R}_+)^2. \quad (4.1.10)$$

Let us recall the partition (4.1.6) of the set  $\mathfrak{P}$  of disorder distributions. Then we claim the following.

**Theorem 4.1.2.** *Let  $\mathbf{t} \in (\mathbb{R}_+)^2$  and recall that  $\zeta_{\mathbf{i}} = e^{\beta_n \omega_{\mathbf{i}} - \lambda(\beta_n)} - 1$ ,  $\mathbf{i} \in \mathbb{N}^2$ .*

(i) *If  $\mathbb{P} \in \mathfrak{P}_1$  and if  $\lim_{n \rightarrow \infty} \beta_n = 0$  and  $\lim_{n \rightarrow \infty} n\beta_n^2 = +\infty$  then*

$$\left( \frac{1}{n^{3/2} \beta_n} M_n(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} \xrightarrow{(d)} \left( m_1 \sqrt{m_2 - m_1^2} \mathcal{M}(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} . \quad (4.1.11)$$

(ii) *If  $\mathbb{P} \in \mathfrak{P}_2$  and if  $\lim_{n \rightarrow \infty} \beta_n = 0$  and  $\lim_{n \rightarrow \infty} n\beta_n^4 = +\infty$  then*

$$\left( \frac{1}{n^{3/2} \beta_n^2} M_n(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} \xrightarrow{(d)} \left( \frac{m_2}{2} \sqrt{m_4 - m_2^2} \mathcal{M}(\mathbf{s}) \right)_{\mathbf{s} \in [0, \mathbf{t}]} . \quad (4.1.12)$$

This can be summarized as follows: for  $r \in \{1, 2\}$ , if  $\mathbb{P} \in \mathfrak{P}_r$  and if  $\lim_{n \rightarrow \infty} \beta_n = 0$  and  $\lim_{n \rightarrow \infty} n\beta_n^{2r} = +\infty$  then

$$\left( \frac{1}{c_r n^{3/2} \beta_n^r} M_n(\mathbf{s}) \right)_{\mathbf{s} \in [\mathbf{0}, \mathbf{t}]} \xrightarrow{(d)} \left( \mathcal{M}(\mathbf{s}) \right)_{\mathbf{s} \in [\mathbf{0}, \mathbf{t}]}, \quad (4.1.13)$$

with the constant  $c_r$  defined by

$$c_1 := m_1 \sqrt{m_2 - m_1^2}, \quad \text{and} \quad c_2 := \frac{m_2}{2} \sqrt{m_4 - m_2^2}. \quad (4.1.14)$$

Those claims hold for the topology of uniform convergence on  $[\mathbf{0}, \mathbf{t}]$ : more precisely, define

$$L^\infty([\mathbf{0}, \mathbf{t}]) := \left\{ f : [\mathbf{0}, \mathbf{t}] \rightarrow \mathbb{R}; f \text{ is } (\text{Bor}([\mathbf{0}, \mathbf{t}]), \text{Bor}(\mathbb{R}))\text{-measurable and bounded} \right\}, \quad (4.1.15)$$

and give it the norm  $\|\cdot\|_\infty$  (and ensuing Borel sigma-algebra). Then, for any random variables  $W_n, n \geq 1$  and  $\mathcal{W}$  in  $L^\infty([\mathbf{0}, \mathbf{t}])$ , we have  $W_n \xrightarrow{(d)} \mathcal{W}$  if for any bounded function  $h : L^\infty([\mathbf{0}, \mathbf{t}]) \rightarrow \mathbb{R}$  that is continuous (for the aforementioned topology),  $\mathbb{E}[h(W_n)] \rightarrow \mathbb{E}[h(\mathcal{W})]$  as  $n \rightarrow \infty$ . Notice that the fields  $M_n$  and  $\mathcal{M}$  defined above are a.s. bounded so this convergence is well-posed.

To discern notation, the  $L^p$  norms on spaces of functions from  $[\mathbf{0}, \mathbf{t}]$  to  $\mathbb{R}$  (i.e.  $L^p([\mathbf{0}, \mathbf{t}])$ ) will be noted  $\|\cdot\|_p$ , whereas on spaces of real random variables (i.e.  $L^p(\mathbb{P})$ ) they will be noted  $\|\cdot\|_{L^p(\mathbb{P})}$  or  $\|\cdot\|_{L^p}$ ,  $p \in [1, \infty]$ .

A (non-trivial) consequence of Theorem 4.1.2, that we show below, is the convergence in distribution and in  $L^2(\mathbb{P})$

$$\frac{L(n)}{n^{\alpha-2}} \frac{1}{c_r n^{3/2} \beta_n^r} \sum_{\mathbf{i} \in \llbracket 1, n\mathbf{t} \rrbracket} \zeta_{\mathbf{i}} u(\mathbf{i}) \xrightarrow{n \rightarrow +\infty} \int_{[\mathbf{0}, \mathbf{t}]} \varphi(\mathbf{s}) d\mathcal{M}(\mathbf{s}), \quad (4.1.16)$$

for  $\mathbb{P} \in \mathfrak{P}_r$  and  $r \in \{1, 2\}$ . The definition of the integral with respect to  $\mathcal{M}(d\mathbf{s})$ , together with the fact that the integral on the right-hand side of (4.1.16) is well-defined, is discussed below. This suggests that one should take  $\beta_n := \widehat{\beta} (n^{\frac{1}{2}-\alpha} L(n))^{\frac{1}{r}}$  with  $r \in \{1, 2\}$  and  $\widehat{\beta} \in [0, +\infty)$  in order for the first term in (4.1.8) to converge, and that one needs to have  $1/2 < \alpha < 1$  to satisfy the requirements of Theorem 4.1.2.

### 4.1.3 Integrating with respect to the field $\mathcal{M}$

Let  $\mathbf{t} \in (\mathbb{R}_+)^2$ , and let

$$\mathcal{S}_{\mathbf{t}} := \left\{ [\mathbf{u}, \mathbf{v}]; \mathbf{0} \preceq \mathbf{u} \preceq \mathbf{v} \preceq \mathbf{t} \right\} \cup \{\emptyset\}, \quad (4.1.17)$$

denote the set of sub-rectangles of  $[\mathbf{0}, \mathbf{t}]$ , closed at the bottom and left and open at the top and right. Notice that it is a semi-ring (i.e. it is non-empty, stable by finite intersection, and

for  $A, B \in \mathcal{S}_t$ ,  $A \setminus B$  is a finite union of disjoint elements of  $\mathcal{S}_t$ ), and that it generates the sigma-algebra  $\sigma(\mathcal{S}_t) = \text{Bor}([0, t])$ .

We call a  $L^2(\mathbb{P})$ -random field on  $\mathcal{S}_t$  any family  $M$  of random variables  $M(A) \in L^2(\mathbb{P})$ ,  $A \in \mathcal{S}_t$ . We will sometimes call it a “random field on  $[0, t]$ ”. For any application  $M : [0, t] \rightarrow L^2(\mathbb{P})$  (which we may also call “ $(L^2(\mathbb{P}))$ -random function from  $[0, t]$  to  $\mathbb{R}$ ”), we can define a random field  $(M(A))_{A \in \mathcal{S}_t}$ , by letting for any rectangle  $[u, v] \in \mathcal{S}_t$ ,  $u \preceq v$ ,

$$M([u, v]) := M(v_1, v_2) - M(u_1, v_2) - M(v_1, u_2) + M(u_1, u_2), \quad (4.1.18)$$

and  $M(\emptyset) := 0$ , which we call the *increment* of  $M$  on  $A$ . Notice that any random function generates a field via its increments, but some fields are not constructed by pointwise-defined functions (e.g. the white noise). Most of our upcoming results hold for generic random fields, thus we do not distinguish notation between increments of functions and fields; and we will explicitly mention whenever we assume that a field is generated via the increments of some pointwise-defined random function.

With these definitions we notice that  $\zeta_i(\beta_n) = M_n(\frac{1}{n}(\mathbf{i} - \mathbf{1}) + \Delta_n)$  where we have set  $\Delta_n := [0, \frac{1}{n}\mathbf{1}]$ ,  $\mathbf{i} \in \mathbb{N}^2$ ,  $n \in \mathbb{N}$ ; hence the first order term from the polynomial expansion of  $Z_{nt, h}^{\beta, \text{free}}$ , more precisely the first term from (4.1.8), can be rewritten as

$$\sum_{i \in [1, nt]} \zeta_i u(i) = \frac{c_r \beta_n^r n^{\alpha-1/2}}{L(n)} \sum_{i \in [1, nt]} \varphi_n(\frac{1}{n}\mathbf{i}) \overline{M}_n(\frac{1}{n}(\mathbf{i} - \mathbf{1}) + \Delta_n), \quad (4.1.19)$$

where we have defined  $\overline{M}_n := \frac{1}{c_r n^{3/2} \beta_n^r} M_n$  and the piecewise constant function

$$\varphi_n(\mathbf{s}) = \varphi_n(\mathbf{s}^{[n]}) := n^{2-\alpha} L(n) u(n\mathbf{s}^{[n]}), \quad (4.1.20)$$

with  $u$  the renewal mass function,  $r \in \{1, 2\}$  and  $\mathbf{s}^{[n]} := \frac{1}{n} \lfloor n\mathbf{s} \rfloor \in \frac{1}{n} \mathbb{Z}^2$ . If  $\beta_n \sim \widehat{\beta} (n^{\frac{1}{2}-\alpha} L(n))^{\frac{1}{r}}$ , then the first factor in the r.h.s. of (4.1.19) goes to  $c_r \widehat{\beta}^r$  as  $n \rightarrow \infty$ . Moreover Theorem 4.1.1 states that  $\varphi_n(\mathbf{s}) \rightarrow \varphi(\mathbf{s})$  for  $\mathbf{s} \in [0, t]$ , and Theorem 4.1.2 gives that  $\overline{M}_n \rightarrow \mathcal{M}$  in distribution as  $n \rightarrow \infty$ , assuming that  $\alpha \in (1/2, 1)$ . Moreover, the r.h.s. in (4.1.19) is bilinear in  $(\varphi_n, \overline{M}_n)$ , so it would be natural to write it as an integral “ $\int_{[0, t]} \varphi_n d\overline{M}_n$ ” and extend its definition, so to obtain some convergence for the real-valued random variables:

$$\int_{[0, t]} \varphi_n d\overline{M}_n \xrightarrow{n \rightarrow \infty} \int_{[0, t]} \varphi d\mathcal{M}, \quad (4.1.21)$$

at least in distribution.

However, this construction is challenging in our setting because the integrating field  $\mathcal{M}$  is strongly correlated. Let us give a one-dimensional example: taking inspiration from the construction of Itô’s integral and Itô’s isometry, a way to define an integral against a convenient process  $(X_s)_{0 \leq s \leq t}$  (e.g. the Brownian motion) is to define some measure  $\mu_X$  on  $[0, t)$  such that, for all  $f : [0, t) \rightarrow \mathbb{R}$  “integrable” against  $X$ , one has

$$\int_{[0, t)} f^2 d\mu_X = \mathbb{E} \left[ \left( \int_{[0, t)} f dX \right)^2 \right], \quad (4.1.22)$$

More precisely, if for  $[u, v] \subset [0, t]$  we define  $\int_{[0,t]} \mathbf{1}_{[u,v]} dX := X([u, v]) := X(v) - X(u)$ , the isometry from (4.1.22) ensures that this definition can be extended linearly on a set of integrable functions (i.e. satisfying  $\int_0^t f^2 d\mu_X < \infty$ ). Unfortunately, this construction fails if  $(X_s)_{0 \leq s \leq t}$  is a more generic process: indeed, for  $[u, v], [u', v'] \subset [0, t]$  disjoint,

$$\begin{aligned} \mathbb{E} \left[ \left( X([u, v]) + X([u', v']) \right)^2 \right] &= \mathbb{E} \left[ X([u, v])^2 \right] + \mathbb{E} \left[ X([u', v'])^2 \right] \\ &\quad + \mathbb{E} \left[ X([u, v]) X([u', v']) \right]. \end{aligned}$$

If  $X$  is a martingale, the last term is zero; but this cannot be guaranteed for a general process, therefore no isometry such as (4.1.22) may be constructed in general.

As far as multidimensional stochastic integrals are concerned, they have been developed in [41] among others, but still focuses on multidimensional martingales. However, in our setting,  $\mathcal{M}$  is not a martingale, and up to our knowledge attempts at expanding stochastic integrals to non-martingales have mostly focused on stationary processes so far (such as the fractional Brownian motion). In this thesis we propose to bypass those limitations with another approach: instead of defining a measure on  $[0, t] \subset \mathbb{R}^2$  (or  $[0, t]$  in one dimension), we define a measure  $\nu_{\mathcal{M}}$  on  $[0, t]^2 \subset \mathbb{R}^4$  (or  $[0, t]^2$ ) such that for “integrable” functions  $f, g$  on  $[0, t]$ , we have

$$\int_{[0,t]^2} f(\mathbf{u})g(\mathbf{v}) d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) = \mathbb{E} \left[ \left( \int_{[0,t]} f d\mathcal{M} \right) \left( \int_{[0,t]} g d\mathcal{M} \right) \right]. \quad (4.1.23)$$

The set of “integrable” functions will be constructed by setting  $\int_{[0,t]} \mathbf{1}_{[u,v]} d\mathcal{M} := \mathcal{M}([u, v])$  as previously, then extending linearly this definition to the set of functions  $f$  satisfying  $\int_{[0,t]^2} f(\mathbf{u})f(\mathbf{v})d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) < \infty$  (using that both terms in (4.1.23) are bilinear in  $(f, g)$ ). We will call  $\nu_{\mathcal{M}}$  the *covariance measure* of the field  $\mathcal{M}$  on  $[0, t]^2$ .

We prove below the following result.

**Proposition 4.1.3.** *Let  $\mathcal{M}$  be a Gaussian field on  $(\mathbb{R}_+)^2$  with zero-mean and covariance matrix  $K(\mathbf{u}, \mathbf{v})$  given in (4.1.10). There is a unique (finite) measure  $\nu_{\mathcal{M}}$  on  $\text{Bor}([0, t]^2)$  such that for any  $A, B \in \mathcal{S}_t$ ,  $\nu_{\mathcal{M}}(A, B) = \mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)]$ . Moreover, for any measurable functions  $g, h : [0, t] \rightarrow \mathbb{R}$  non-negative or bounded, we have*

$$\int_{[0,t]^2} g(\mathbf{u})h(\mathbf{v})d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) = \int_{[0,t]} g(\mathbf{u}) \left( \int_0^{t_1} h(x, u_2)dx + \int_0^{t_2} h(u_1, y)dy \right) d\mathbf{u}. \quad (4.1.24)$$

**Remark 4.1.4.** Note that the covariance measure  $\nu_{\mathcal{M}}$  explicitly displays the correlation structure of the field  $\mathcal{M}$  on lines and columns, as can be seen in (4.1.24). For the sake of comparison, let us consider the white noise  $W$  on  $[0, t] \subset \mathbb{R}^2$ , which is a Gaussian field with covariance  $K(x, y) = \mathbf{1}_{x=y}$ . One can define the covariance measure  $\nu_W$  of  $W$ , by extending  $\nu_W(A, B) := \mathbb{E}[W(A)W(B)] = \lambda_2(A \cap B)$  for  $A, B \in \mathcal{S}_t$  to all  $E \in \text{Bor}([0, t]^2)$ :

$$\nu_W(E) = \lambda_2(\mathbf{u} \in [0, t]; (\mathbf{u}, \mathbf{u}) \in E), \quad (4.1.25)$$



where  $\lambda_2$  denotes the 2-dimensional Lebesgue measure. For non-negative and measurable  $g, h : [0, \mathbf{t}] \rightarrow \mathbb{R}_+$ , we have

$$\int_{[0, \mathbf{t}]^2} g(\mathbf{u})h(\mathbf{v}) d\nu_W(\mathbf{u}, \mathbf{v}) = \int_{[0, \mathbf{t}]^2} g(\mathbf{u})h(\mathbf{u}) d\lambda_2(\mathbf{u}). \quad (4.1.26)$$

Whereas  $\nu_{\mathcal{M}}$  can be rewritten as a Lebesgue measure supported on a 3-dimensional set and displays the correlation structure of  $\mathcal{M}$ ,  $\nu_W$  is supported on the diagonal of  $[0, \mathbf{t}]^2$  which is 2-dimensional, and thus displays the absence of correlation in the white noise.

Let  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  be a random field on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , and assume that it admits a (finite, non-negative) covariance measure  $\nu$  on  $\text{Bor}([0, \mathbf{t}]^2)$ , i.e. such that for  $A, B \in \mathcal{S}_t$ ,

$$\mathbb{E}[X(A)X(B)] = \nu(A \times B). \quad (4.1.27)$$

We then give some properties on that measure  $\nu$ , which are useful to construct the stochastic integral against  $X$ . First, let us define

$$L_X^2 := \left\{ g : [0, \mathbf{t}] \rightarrow \mathbb{R} \text{ measurable ; } \int_{[0, \mathbf{t}]^2} |g(\mathbf{u})g(\mathbf{v})| d\nu(\mathbf{u}, \mathbf{v}) < +\infty \right\}, \quad (4.1.28)$$

and for  $g, h \in L_X^2$ , write with an abuse of notation,

$$\langle g, h \rangle_\nu := \int_{[0, \mathbf{t}]^2} g(\mathbf{u})h(\mathbf{v}) d\nu(\mathbf{u}, \mathbf{v}), \quad \text{and} \quad \|g\|_\nu := \sqrt{\langle g, g \rangle_\nu}. \quad (4.1.29)$$

In the next proposition, we notably prove that those definitions are well-posed.

**Proposition 4.1.5.** *Let  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  be a random field which admits a finite, non-negative covariance measure  $\nu$  on  $\text{Bor}([0, \mathbf{t}]^2)$ .*

- (i) *The set  $L_X^2$  is a vector space and contains  $L^\infty([0, \mathbf{t}])$ . Moreover  $\langle g, h \rangle_\nu$  is well-posed for  $g, h \in L_X^2$ .*
- (ii) *The application  $\langle \cdot, \cdot \rangle_\nu$  is bilinear, symmetric and semi-definite positive. In particular  $\|g\|_\nu$  is well-posed for  $g \in L_X^2$ .*
- (iii) (Cauchy-Schwarz) *For  $g, h \in L_X^2$ , one has  $\langle g, h \rangle_\nu \leq \|g\|_\nu \|h\|_\nu$ .*
- (iv) (Triangle inequality) *For  $g, h \in L_X^2$ , one has  $\|g + h\|_\nu \leq \|g\|_\nu + \|h\|_\nu$ .*

**Remark 4.1.6.** Despite all those properties,  $\langle \cdot, \cdot \rangle_\nu$  is not actually a scalar product on  $L_X^2$  in general (in particular  $\| \cdot \|_\nu$  is not a norm). Indeed, let us assume  $\mathbf{t} = \mathbf{1}$  for simplicity, and recall the expression of  $\nu_{\mathcal{M}}$  from (4.1.24). Then, letting  $h : [0, \mathbf{1}] \rightarrow \mathbb{R}$  be defined by

$$h = \mathbf{1}_{[0, 1/2] \times [0, 1/2]} - \mathbf{1}_{[1/2, 1] \times [0, 1/2]} - \mathbf{1}_{[0, 1/2] \times [1/2, 1]} + \mathbf{1}_{[1/2, 1] \times [1/2, 1]},$$

we have  $\|h\|_{\nu_{\mathcal{M}}} = \|\mathbf{1}_{[0, 1]}\|_{\nu_{\mathcal{M}}} = 2 < \infty$  so  $h \in L_X^2$ ; however  $\|h\|_{\nu_{\mathcal{M}}} = 0$ , and more generally  $\langle g, h \rangle_{\nu_{\mathcal{M}}} = 0$  for all  $g \in L_{\mathcal{M}}^2$ .

On the contrary, when considering the white noise  $W$  on  $[0, 1)$ , one has for any function  $g : [0, 1) \rightarrow \mathbb{R}$ ,

$$\|g\|_{\nu_W}^2 = \int_{[0,t)} (g(\mathbf{u}))^2 d\lambda_2(\mathbf{u}) ,$$

(where we used (4.1.26)). This proves that  $\|g\|_{\nu_W} = 0$  if and only if  $g = 0$   $\lambda_2$ -almost everywhere.

Finally we state in the next theorem that, for  $X$  a random field on  $\mathcal{S}_t$  which admits a covariance measure  $\nu$ , a stochastic integral against  $X$  is well-defined on  $L_X^2$ . This is a general result, of its own interest.

**Theorem 4.1.7.** *Let  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  be a random field which admits a finite, non-negative covariance measure  $\nu$  on  $\text{Bor}([0, t]^2)$ . For  $A \in \mathcal{S}_t$ , define*

$$1_A \diamond X := X(A) \in L^2(\mathbb{P}) .$$

*Then the application  $f \mapsto f \diamond X$  can be extended to  $L_X^2$  such that (i) it is linear, (ii) for  $g \in L_X^2$ ,  $g \diamond X \in L^2(\mathbb{P})$  is defined almost everywhere on  $(\Omega, \mathcal{F}, \mathbb{P})$ , and (iii) for  $g, h \in L_X^2$ ,*

$$\mathbb{E}[(g \diamond X)(h \diamond X)] = \langle g, h \rangle_\nu . \tag{4.1.30}$$

The random variable  $g \diamond X$  will be called the *integral of  $g$  against  $X$* , and sometimes denoted

$$\int g dX := g \diamond X .$$

Notice that the observation  $L^\infty([0, t]) \subset L_X^2$  from Proposition 4.1.5 is actually obvious since  $\nu([0, t]^2) < \infty$  by assumption; however it is noteworthy, since it shows that the extension constructed in Theorem 4.1.7 is not dull.

**Remark 4.1.8.** (i) Let us point out that we restricted Proposition 4.1.3 and Theorem 4.1.7 to bounded rectangles  $[0, t) \subset (\mathbb{R}_+)^2$ ; however they should also apply to random fields  $X$  defined on the full quadrant  $(\mathbb{R}_+)^2$ , with  $\nu_X$  a sigma-finite measure on  $\text{Bor}((\mathbb{R}_+)^2)$ , and a set of integrable functions  $L_X^2$  containing all bounded, measurable functions with compact support. In this thesis we restricted ourselves to a bounded rectangle since it is sufficient in our setting, but the more general construction will be included in a latter version of this work.

(ii) In our setting we are only interested in 2-dimensional fields, i.e. defined on sub-rectangles of  $(\mathbb{R}_+)^2$ ; however this construction can straightforwardly be adapted to any  $d$ -dimensional field,  $d \in \mathbb{N}$ . We discuss this in Section 4.3.4, when constructing the limits of terms of rank  $k > 1$  in the polynomial expansion.

We conclude this section by a result ensuring that the function  $\varphi$  appearing in Theorem 4.1.1 is integrable with respect to the field  $\mathcal{M}$  with covariance  $K(\cdot, \cdot)$  given in (4.1.10), that is  $\varphi \in L_{\mathcal{M}}^2$ .

**Proposition 4.1.9.** *Let  $g : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$  be the function defined by  $g(\mathbf{s}) := (\|\mathbf{s}\|_1)^{\alpha-2}$  for  $\mathbf{s} \in [0, t) \setminus \{0\}$ . If  $\alpha \in (\frac{1}{2}, 1)$ , then  $g \in L_{\mathcal{M}}^2$ , and in particular  $\varphi \in L_{\mathcal{M}}^2$ .*

#### 4.1.4 Statement of the main result

We now have the tools to state our main result. At the current state of this document it is still a conjecture, and we are currently working on proving the last remaining parts.

**Conjecture 4.1.10.** *Let  $\tau$  satisfy (4.1.1) with  $\alpha \in (1/2, 1)$ . Let  $r \in \{1, 2\}$  and  $\mathbb{P} \in \mathfrak{P}_r$ , and let  $(\beta_n)_{n \geq 1}$ ,  $(h_n)_{n \geq 1}$  be such that*

$$\lim_{n \rightarrow +\infty} \beta_n \left( n^{\alpha - \frac{1}{2}} L(n)^{-1} \right)^{\frac{1}{r}} = \widehat{\beta} \in [0, +\infty), \quad \text{and} \quad \lim_{n \rightarrow +\infty} h_n n^\alpha L(n)^{-1} = \widehat{h} \in \mathbb{R}. \quad (4.1.31)$$

Then for  $\mathbf{t} \succ \mathbf{0}$ , we have the convergence in distribution

$$\mathbf{Z}_{[\mathbf{nt}], h_n}^{\beta_n, \omega, \mathbf{q}} \xrightarrow[n \rightarrow +\infty]{(d)} \mathbf{Z}_{\mathbf{t}, \widehat{h}}^{\widehat{\beta}, \mathcal{M}, \mathbf{q}}, \quad (4.1.32)$$

where the random variable  $\mathbf{Z}_{\mathbf{t}, \widehat{h}}^{\widehat{\beta}, \mathcal{M}, \mathbf{q}}$  is given by the chaos expansion, which converges in  $L^2(\mathbb{P})$ ,

$$\mathbf{Z}_{\mathbf{t}, \widehat{h}}^{\widehat{\beta}, \mathcal{M}, \mathbf{q}} := 1 + \sum_{k=1}^{+\infty} \int \cdots \int_{\mathbf{0} \prec \mathbf{s}_1 \prec \cdots \prec \mathbf{s}_k \prec \mathbf{t}} \psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k) \prod_{j=1}^k \left( c_r \widehat{\beta}^r d\mathcal{M}(\mathbf{s}_j) + \widehat{h} d\mathbf{s}_j \right). \quad (4.1.33)$$

In (4.1.33), the constant  $c_r > 0$  is defined in (4.1.14), and  $\psi_{\mathbf{t}}^{\mathbf{q}}$  is symmetric, defined by

$$\psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \varphi(\mathbf{s}_1) \varphi(\mathbf{s}_2 - \mathbf{s}_1) \cdots \varphi(\mathbf{t} - \mathbf{s}_k) \quad (4.1.34)$$

for  $\mathbf{0} =: \mathbf{s}_0 \prec \mathbf{s}_1 \prec \cdots \prec \mathbf{s}_k \prec \mathbf{t}$ , and  $\psi_{\mathbf{t}}^{\mathbf{q}}(\cdot) = 0$  if its arguments cannot be ordered with  $\prec$ .

As far as the free<sup>free</sup> and conditioned<sup>cond</sup> partition function are concerned, the same result holds by replacing  $\psi_{\mathbf{t}}^{\mathbf{q}}$  respectively by  $\psi_{\mathbf{t}}^{\text{free}}$  symmetric such that

$$\psi_{\mathbf{t}}^{\text{free}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \varphi(\mathbf{s}_1) \varphi(\mathbf{s}_2 - \mathbf{s}_1) \cdots \varphi(\mathbf{s}_k - \mathbf{s}_{k-1}), \quad (4.1.35)$$

for  $\mathbf{0} =: \mathbf{s}_0 \prec \mathbf{s}_1 \prec \cdots \prec \mathbf{s}_k \prec \mathbf{t}$  (and 0 otherwise); and by

$$\psi_{\mathbf{t}}^{\text{cond}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \frac{\psi_{\mathbf{t}}^{\mathbf{q}}(\mathbf{s}_1, \dots, \mathbf{s}_k)}{\varphi(\mathbf{t})}. \quad (4.1.36)$$

We have already proven that the series from (4.1.33) is well-defined in  $L^2(\mathbb{P})$  (in particular integrals of all order are well-posed), and that the first term of the polynomial expansion converges to the first term of the continuous chaos expansion, that is (4.1.16)—see Propositions 4.3.7 and 4.4.1 below. What remains to be shown is the convergence of terms of order  $k \in \mathbb{N}$  to their continuous counterpart (for which we only lack a few technical estimates at present); then the convergence (4.1.32) is a straightforward consequence. These will be added in a later version of this work, along with some further comments on our results.

## 4.2 Convergence of the field $M_n$ to $\mathcal{M}$ : proof of Theorem 4.1.2

In this section, we prove the convergence in Theorem 4.1.2, and we also provide useful estimates on the process  $(M_n(\mathbf{s}))_{\mathbf{s} \geq \mathbf{0}}$ .

### 4.2.1 Preliminary result: the covariance structure

We start with some preliminaries, controlling the covariances of the  $\zeta_i$ 's. Recall that  $\zeta_i = \zeta_i(\beta_n) := e^{\beta_n \omega_i - \lambda(\beta_n)} - 1$  has mean 0, and that  $m_k = \mathbb{E}[\widehat{\omega}_1^k]$ ,  $k \in \mathbb{N}$ .

**Lemma 4.2.1.** *Let  $\mathbf{i}, \mathbf{j} \in \mathbb{N}^2$ . Then, as  $\beta_n \rightarrow 0$*

$$\mathbb{E}[\zeta_i \zeta_j] = \begin{cases} 0 & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \\ 0 & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_0, \\ m_1^2(m_2 - m_1^2)\beta_n^2 + o(\beta_n^2) & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_1, \\ \frac{m_2^2}{4}(m_4 - m_2^2)\beta_n^4 + o(\beta_n^4) & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_2, \\ (m_2^2 - m_1^4)\beta_n^2 + o(\beta_n^2) & \text{if } \mathbf{i} = \mathbf{j}. \end{cases} \quad (4.2.1)$$

Notice that all three non-trivial coefficients displayed above are positive:  $m_2 > m_1^2$  because we assume  $\widehat{\omega}$  non-constant  $\mathbb{P}$ -a.s.,  $m_1 \neq 0$  if  $\mathbb{P} \in \mathfrak{P}_1$  and  $m_4 > m_2^2$  if  $\mathbb{P} \in \mathfrak{P}_2$ . Notice that for  $r \in \{1, 2\}$  and  $\mathbb{P} \in \mathfrak{P}_r$ , we can rewrite Lemma 4.2.1 as

$$\mathbb{E}[\zeta_i \zeta_j] = \begin{cases} 0 & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \\ c_r^2 \beta_n^{2r} + o(\beta_n^{2r}) & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j}, \\ c_3^2 \beta_n^2 + o(\beta_n^2) & \text{if } \mathbf{i} = \mathbf{j}, \end{cases} \quad (4.2.2)$$

where  $c_r$  are defined in (4.1.14) and  $(c_3)^2 := m_2^2 - m_1^4 = \text{Var}(\omega_1)$ .

*Proof.* First of all, if  $i_1 \neq j_1$  and  $i_2 \neq j_2$ , then  $\omega_i$  and  $\omega_j$  are independent, and we clearly have that  $\mathbb{E}[\zeta_i \zeta_j] = 0$ . Otherwise, we use a straightforward expansion of  $e^{\lambda(\beta_n)}$  as  $\beta_n \rightarrow 0$ , which gives

$$e^{\lambda(\beta_n)} = \mathbb{E}[e^{\beta_n \widehat{\omega}_1}] = 1 + m_1^2 \beta_n + \frac{m_2^2}{2} \beta_n^2 + \frac{m_3^2}{6} \beta_n^3 + \frac{m_4^2}{24} \beta_n^4 + o(\beta_n^4). \quad (4.2.3)$$

If  $\mathbf{i} = \mathbf{j}$ , then we find  $\mathbb{E}[\zeta_i \zeta_j] = e^{\lambda(2\beta_n)} - 1 = (m_2^2 - m_1^4)\beta_n^2 + o(\beta_n^2)$  as  $\beta_n \rightarrow 0$ .

If  $i_1 = j_1$  but  $i_2 \neq j_2$ , then we have  $\mathbb{E}[\zeta_i \zeta_j] = \mathbb{E}[e^{\beta_n \widehat{\omega}_{i_1}(\widehat{\omega}_{i_2} + \widehat{\omega}_{j_2})}] e^{-2\lambda(\beta_n)} - 1$ . With a similar expansion as in (4.2.3) we get that, as  $\beta_n \rightarrow 0$ ,

$$\begin{aligned} \mathbb{E}[e^{\beta_n \widehat{\omega}_{i_1}(\widehat{\omega}_{i_2} + \widehat{\omega}_{j_2})}] &= 1 + 2m_1^2 \beta_n + m_2(m_2 + m_1^2)\beta_n^2 + \frac{m_3(m_3 + 3m_1 m_2)}{3} \beta_n^3 \\ &+ \frac{m_4(m_4 + 4m_1 m_3 + 3m_2^2)}{12} \beta_n^4 + o(\beta_n^4). \end{aligned} \quad (4.2.4)$$

Together with (4.2.3), if  $\mathbf{i}$  and  $\mathbf{j}$  are aligned but not equal, we get after simplifications that  $\mathbb{E}[\zeta_i \zeta_j] = m_1^2(m_2 - m_1^2)\beta_n^2 + o(\beta_n^2)$ . If  $m_1 = 0$ , we push the expansion to the 4-th order and get  $\mathbb{E}[\zeta_i \zeta_j] = \frac{m_2^2(m_4 - m_2^2)}{4} \beta_n^4 + o(\beta_n^4)$ .

Finally, notice that if  $\mathbb{P} \in \mathfrak{P}_0$ , that is  $\mathbb{P} = \mathbb{P}_{\pm x}$ ,  $x > 0$ , then  $e^{\lambda(\beta_n)} = \cosh(\beta_n x^2)$  and

$$\mathbb{E}\left[e^{\beta_n \widehat{\omega}_{i_1}(\widehat{\omega}_{i_2} + \widehat{\omega}_{j_2})}\right] = \frac{1}{4} \left( e^{2\beta_n x^2} + e^{-2\beta_n x^2} + 2 \right) = e^{2\lambda(\beta_n)},$$

so  $\mathbb{E}[\zeta_i \zeta_j] = 0$  as soon as  $\mathbf{i} \neq \mathbf{j}$ .  $\square$

**Remark 4.2.2.** Lemma 4.2.1 can actually be adapted to the case of a more general interaction function  $f(\cdot, \cdot)$ . Let  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a symmetric function and define  $\omega_i := f(\widehat{\omega}_{i(1)}, \widehat{\omega}_{i(2)})$  where  $(\widehat{\omega}_i)_{i \geq 1}$  and  $(\bar{\omega}_i)_{i \geq 1}$  are two independent sequences of i.i.d. random variables with the same distribution  $\mathbb{P}$ . We assume that  $\lambda(\beta) := \log \mathbb{E}[e^{\beta \omega_i}] < +\infty$  for  $\beta$  small enough. Then, defining again  $\zeta_i := e^{\beta \omega_i - \lambda(\beta)} - 1$ , we can show an analogue to Lemma 4.2.1 for the covariance structure of  $(\zeta_i)_{i \in \mathbb{N}^2}$ :

$$\mathbb{E}[\zeta_i \zeta_j] = \begin{cases} 0 & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \\ 0 & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_0, \\ c_1^2 \beta_n^2 + o(\beta_n^2) & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_1, \\ c_2^2 \beta_n^4 + o(\beta_n^4) & \text{if } \mathbf{i} \leftrightarrow \mathbf{j}, \mathbf{i} \neq \mathbf{j} \text{ and } \mathbb{P} \in \mathfrak{P}_2, \\ c_3^2 \beta_n^2 + o(\beta_n^2) & \text{if } \mathbf{i} = \mathbf{j}. \end{cases} \quad (4.2.5)$$

Here the partition of  $\mathfrak{P}$  depends on the function  $f$ : we have  $\mathfrak{P}_0 = \{\mathbb{P}, \omega_i^2 \text{ is } \mathbb{P}\text{-a.s. constant}\}$ ,  $\mathfrak{P}_1 = \{\mathbb{P}, \mathbb{E}[\omega_i \mid \widehat{\omega}_{i(1)}] \neq \mathbb{E}[\omega_i] \text{ with positive } \mathbb{P}\text{-probability}\}$ ,  $\mathfrak{P}_2 = \{\mathbb{P}, \mathbb{E}[\omega_i \mid \widehat{\omega}_{i(1)}] = \mathbb{E}[\omega_i] \text{ } \mathbb{P}\text{-a.s.}\}$ . The constants  $c_1, c_2, c_3$  are given by

$$c_1^2 = \text{Cov}(\omega_{1,1}, \omega_{1,2}), \quad c_2^2 = \frac{1}{4} \text{Cov}(\omega_{1,1}^2, \omega_{1,2}^2), \quad c_3^2 = \text{Var}(\omega_1),$$

and they are all positive (because of how  $\mathfrak{P}_1, \mathfrak{P}_2$  are defined). We have chosen to stick to the presentation of the case  $f(x, y) = xy$  for coherence with Chapter 3 and since it already encapsulates all possible behaviors. We stress that the proof of Theorem 4.1.2 entirely relies on Lemma 4.2.1, so (4.2.5) (which is of the same form as (4.2.2)) is enough to conclude in the case of a general interaction function  $f(\cdot, \cdot)$ .

Let us define, for  $\mathbf{s} \in \mathbb{R}_+ \times \mathbb{R}_+$ ,

$$\overline{M}_n(\mathbf{s}) := \frac{1}{c_r n^{3/2} \beta_n^r} M_n(\mathbf{s}) = \frac{1}{c_r n^{3/2} \beta_n^r} \sum_{i \in \llbracket \mathbf{1}, n\mathbf{s} \rrbracket} \zeta_i, \quad (4.2.6)$$

where  $\frac{1}{c_r n^{3/2} \beta_n^r}$  is the scaling advertised in Theorem 4.1.2. Thanks to Lemma 4.2.1, we easily identify the covariance structure of  $(\overline{M}_n(\mathbf{s}))_{\mathbf{s} \geq \mathbf{0}}$ . Let  $\mathbf{s} = (s_1, s_2), \mathbf{t} = (t_1, t_2) \in \mathbb{R}_+^2$ , and let us compute

$$\mathbb{E}[\overline{M}_n(\mathbf{s}) \overline{M}_n(\mathbf{t})] = \frac{1}{c_r^2 n^3 \beta_n^{2r}} \sum_{i \in \llbracket \mathbf{1}, n\mathbf{s} \rrbracket} \sum_{j \in \llbracket \mathbf{1}, n\mathbf{t} \rrbracket} \mathbb{E}[\zeta_i \zeta_j]. \quad (4.2.7)$$

Then, in view of Lemma 4.2.1 (or (4.2.2)), we distinguish in the sum indices  $(i, j)$  that are equal (there are  $(1 + o(1))(s_1 \wedge t_1)(s_2 \wedge t_2)n^2$  of them), aligned indices that are not equal (there are  $(1 + o(1))(s_1 \wedge t_1)(s_2 \wedge t_2)(s_1 \vee t_1 + s_2 \vee t_2)n^3$  of them, see Figure 4.1), and other indices (which do not contribute to the sum).

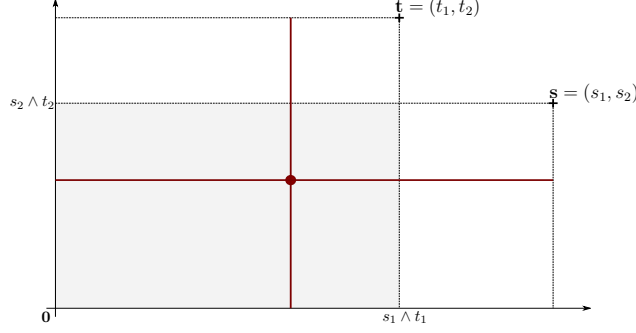


Figure 4.1: Graphical representation of indices  $i \in \llbracket \mathbf{1}, ns \rrbracket, j \in \llbracket \mathbf{1}, nt \rrbracket$  with  $i \leftrightarrow j$  and  $i \neq j$ . One of  $i, j$  must be in  $\llbracket \mathbf{1}, ns \rrbracket \cap \llbracket \mathbf{1}, nt \rrbracket$  (there are  $(s_1 \wedge t_1)(s_2 \wedge t_2)n^2$  possible locations, represented as the red dot), and the other one has to be aligned with it (there are  $(s_1 \vee t_1 + s_2 \vee t_2)n - 1$  possibilities, represented by the red lines).

Therefore, in view of (4.2.2), as  $n \rightarrow +\infty$  the correlation  $\mathbb{E}[\overline{M}_n(\mathbf{s})\overline{M}_n(\mathbf{t})]$  is asymptotic to

$$(s_1 \wedge t_1)(s_2 \wedge t_2) \frac{c_3^2}{c_r^2 n \beta_n^{2r}} + (s_1 \wedge t_1)(s_2 \wedge t_2)(s_1 \vee t_1 + s_2 \vee t_2).$$

If  $n\beta_n^{2r} \rightarrow +\infty$ , which is one assumption of Theorem 4.1.2, then we end up with

$$\lim_{n \rightarrow +\infty} \mathbb{E}[\overline{M}_n(\mathbf{s})\overline{M}_n(\mathbf{t})] = K(\mathbf{s}, \mathbf{t}) = (s_1 \wedge t_1)(s_2 \wedge t_2)(s_1 \vee t_1 + s_2 \vee t_2), \quad (4.2.8)$$

which is the correlation function in Theorem 4.1.2.

## 4.2.2 Finite-dimensional convergence

We now prove the convergence in distribution of vectors  $(\overline{M}_n(\mathbf{s}_1), \dots, \overline{M}_n(\mathbf{s}_m))$ , for  $m \in \mathbb{N}$  and  $\mathbf{s}_1, \dots, \mathbf{s}_m \in \mathbb{R}_+^2$ . Let  $\Sigma_{\mathbf{s}_1, \dots, \mathbf{s}_m}(i, j) := K(\mathbf{s}_i, \mathbf{s}_j)$  be the covariance matrix of  $(\mathcal{M}(\mathbf{s}_1), \dots, \mathcal{M}(\mathbf{s}_m))$ , where we recall that  $K(\cdot, \cdot)$  is defined in (4.1.10). In view of (4.2.8),  $\Sigma$  is the limit of the sequence of the covariance matrices of  $(\overline{M}_n(\mathbf{s}_1), \dots, \overline{M}_n(\mathbf{s}_m))$ ; in particular, it is positive semi-definite.

**Proposition 4.2.3.** *Let  $m \in \mathbb{N}$  and  $\mathbf{s}_1, \dots, \mathbf{s}_m \in \mathbb{R}_+^2$ . As  $n \rightarrow \infty$ , if  $\beta_n \rightarrow 0$  and  $n\beta_n^{2r} \rightarrow +\infty$ , then  $(\overline{M}_n(\mathbf{s}_1), \dots, \overline{M}_n(\mathbf{s}_m))$  converges in distribution toward a Gaussian vector, centered and with covariance matrix  $\Sigma_{\mathbf{s}_1, \dots, \mathbf{s}_m}$ .*

Before we prove this proposition, let us start with the case  $m = 1$ , which already encapsulates the combinatorial difficulty and will ease the understanding of the general case

$m \in \mathbb{N}$ . We show the convergence of the moments of  $\overline{M}_n(\mathbf{s})$  to the moments of a Gaussian variable, which implies the convergence in distribution.

**Lemma 4.2.4.** *Let  $\ell \in \mathbb{N}$  and  $\mathbf{s} \in \mathbb{R}_+^2$ . Then  $\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell]$  is well defined if  $n$  is large enough, and if  $\beta_n \rightarrow 0$ ,  $n\beta_n^{2r} \rightarrow +\infty$ , then we have*

$$\lim_{n \rightarrow +\infty} \mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell] = \begin{cases} 0 & \text{if } \ell \text{ is odd,} \\ (K(\mathbf{s}, \mathbf{s}))^{\ell/2} \frac{\ell!}{2^{\ell/2}(\ell/2)!} & \text{if } \ell \text{ is even.} \end{cases} \quad (4.2.9)$$

where  $K(\mathbf{s}, \mathbf{t})$  is defined in (4.1.10), so  $K(\mathbf{s}, \mathbf{s}) = s_1 s_2 (s_1 + s_2)$ .

*Proof.* We write

$$\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell] = \left( \frac{1}{c_r n^{3/2} \beta_n^r} \right)^\ell \sum_{\mathbf{i}_1 \in \llbracket \mathbf{1}, ns \rrbracket \cap \mathbb{N}^2} \cdots \sum_{\mathbf{i}_\ell \in \llbracket \mathbf{1}, ns \rrbracket} \mathbb{E}[\zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_\ell}]. \quad (4.2.10)$$

Now, notice that  $\mathbb{E}[\zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_\ell}]$  depends only on the relative positions of the indices  $(\mathbf{i}_k)_{k=1}^\ell$ . For instance, if one of the  $\mathbf{i}_k$  is isolated (i.e. not aligned with any other index  $\mathbf{i}$ ) then the expectation is equal to 0.

Let us define the classes of sets of “ $m$ -aligned” indices (with possible repetitions of the indices) as

$$\mathcal{A}_m := \left\{ \{\mathbf{i}_1, \dots, \mathbf{i}_m\} \subset \mathbb{N}^2; \begin{array}{l} \forall 1 \leq k, k' \leq m, \exists k_0 = k, k_1, \dots, k_p = k' \\ \text{such that } \forall 1 \leq a \leq p, \mathbf{i}_{k_a} \leftrightarrow \mathbf{i}_{k_{a-1}} \end{array} \right\}, \quad (4.2.11)$$

in other words,  $\{\mathbf{i}_1, \dots, \mathbf{i}_m\} \in \mathcal{A}_m$  if for all pair  $(\mathbf{i}_k, \mathbf{i}_{k'})$ ,  $1 \leq k, k' \leq m$ , there is a path from  $\mathbf{i}_k$  to  $\mathbf{i}_{k'}$  of subsequently aligned indices in  $\{\mathbf{i}_1, \dots, \mathbf{i}_m\}$ . Then, for any  $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_\ell) \in \llbracket \mathbf{1}, ns \rrbracket^\ell$ , there is a unique partition  $\mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_k\}$  of  $\{1, \dots, \ell\}$  into sets of “ $m$ -aligned” indices; in particular  $\{\mathbf{i}_j\}_{j \in J_a} \in \mathcal{A}_{|J_a|}$  for all  $1 \leq a \leq k$ , and  $\mathbf{i}_j \not\leftrightarrow \mathbf{i}_{j'}$  for any  $j \in J_a, j' \in J_b$  with  $a \neq b$ . One can view  $(J_a)_{1 \leq a \leq k}$  as equivalence classes, for the following equivalence relation (defined for  $\mathbf{I}$  fixed):  $j \rightsquigarrow j'$  if and only if there exists a path  $j_0 = j, j_1, \dots, j_b = j'$  in  $\{1, \dots, \ell\}$  satisfying  $\mathbf{i}_{j_a} \leftrightarrow \mathbf{i}_{j_{a+1}}$  for all  $0 \leq a < b$ .

Therefore, we may factorize

$$\mathbb{E}[\zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_\ell}] = \prod_{a=1}^k \mathbb{E} \left[ \prod_{j \in J_a} \zeta_{\mathbf{i}_j} \right]. \quad (4.2.12)$$

As a consequence, denoting  $\mathcal{P}_{k,\ell}$  the set of partitions of  $\{1, \dots, \ell\}$  with  $k$  sets, we write

$$\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell] = \left( \frac{1}{c_r n^{3/2} \beta_n^r} \right)^\ell \sum_{k=1}^{\ell} \sum_{\{J_1, \dots, J_k\} \in \mathcal{P}_{k,\ell}} \sum_{\substack{\mathbf{I}=(\mathbf{i}_1, \dots, \mathbf{i}_\ell) \in \llbracket \mathbf{1}, ns \rrbracket^\ell \\ \mathcal{J}(\mathbf{I})=\{J_1, \dots, J_k\}}} \prod_{a=1}^k \mathbb{E} \left[ \prod_{j \in J_a} \zeta_{\mathbf{i}_j} \right]. \quad (4.2.13)$$

First of all, in (4.2.13), we can restrict the sum to having only  $|J_a| \geq 2$ : if  $|J_a| = 1$  we obviously have  $\mathbb{E}[\prod_{j \in J_a} \zeta_{\mathbf{i}_j}] = \mathbb{E}[\zeta_{\mathbf{i}_1}] = 0$  (note that it also restricts the sum to  $k \leq \ell/2$ ).

Now, we claim that the main contribution in the sum comes from having  $|J_a| = 2$  for all  $a$ , or in other words from the term  $k = \ell/2$ .

Indeed let us show that, for any  $m \geq 3$ ,

$$\frac{1}{(c_r n^{3/2} \beta_n^r)^m} \sum_{I \in \mathcal{A}_m(ns)} \mathbb{E} \left[ \prod_{i \in I} \zeta_i \right] \xrightarrow{n \rightarrow +\infty} 0, \quad (4.2.14)$$

where we introduced the notation  $\mathcal{A}_m(ns) = \{I \in \mathcal{A}_m, I \subset \llbracket \mathbf{1}, ns \rrbracket\}$  (note that we still allow repetitions of the indices).

We use Hölder's inequality to get that  $\mathbb{E}[\prod_{i \in I} \zeta_i] \leq \mathbb{E}[|\zeta_1|^m]$  for any  $I \in \mathcal{A}_m$ . We stress that  $\mathbb{E}[|\zeta_1|^m]$  is finite provided that  $\beta_n$  is small enough, and that a simple Taylor expansion of  $\zeta_1$  gives that there is a constant  $c_m$  such that for  $\beta_n$  sufficiently small,  $\mathbb{E}[|\zeta_1|^m] \leq c_m (\beta_n)^m$ . Then, we may bound  $\text{Card}(\mathcal{A}_m(ns)) \leq m!(\|s\|_1 n)^{1+m}$  by induction, and we get that the left-hand side of (4.2.14) is bounded by a constant (depending on  $m$  and  $s$ ) times  $(n^{3/2} \beta_n^r)^{-m} (\beta_n)^m n^{m+1} = (n^{1/2-1/m} \beta_n^{r-1})^{-m}$ .

If  $r = 1$  (that is  $\mathbb{P} \in \mathfrak{P}_1$ ), then this goes to 0 as  $n \rightarrow +\infty$  for any  $m \geq 3$ . On the other hand, if  $r = 2$  (that is  $\mathbb{P} \in \mathfrak{P}_2$ ), then this goes to 0 provided that  $m \geq 4$ , using the assumption that  $n \beta_n^{2r} \rightarrow 0$ , which implies that  $n^{1/2-1/m} \beta_n^{r-1} \geq n^{1/4} \beta_n \rightarrow +\infty$ .

It therefore remains to treat the case  $m = 3$ , for  $r = 2$ , that is  $\mathbb{P} \in \mathfrak{P}_2$ : we now estimate  $\mathbb{E}[\zeta_{i_1} \zeta_{i_2} \zeta_{i_3}]$ . If two of the indices are equal, then we use the same idea as above: we bound  $\mathbb{E}[\zeta_{i_1} \zeta_{i_2} \zeta_{i_3}]$  by  $c_m (\beta_n)^3$ , and since there are at most  $2(\|s\|_1 n)^3$  such triplets of indices in  $\mathcal{A}_3$ , we get that the contribution of these terms in the left-hand side of (4.2.14) is bounded by a constant times  $n^{-3/2} \beta_n^{-3}$ , which goes to 0 as  $n \rightarrow +\infty$  (we have  $n^{1/2} \beta_n \rightarrow +\infty$ ). If all indices are different (but still form a 3-aligned set), a Taylor expansion of the type (4.2.3)-(4.2.4) gives that  $\mathbb{E}[\zeta_{i_1} \zeta_{i_2} \zeta_{i_3}] = O(\beta_n^4)$  (we skip the details here). Then, since  $\text{Card}(\mathcal{A}_3(ns)) \leq 6(\|s\|_1 n)^4$ , we get that the contribution of distinct indices  $i_1, i_2, i_3$  to the left-hand side of (4.2.14) is bounded by a constant times  $n^{-1/2} \beta_n^{-2}$ , which goes to 0 as  $n \rightarrow +\infty$  (recall  $n^{1/4} \beta_n \rightarrow +\infty$ ).

In addition to (4.2.14), recall that when  $m = 2$  then (4.2.8) shows that

$$(c_r n^{3/2} \beta_n^r)^{-m} \sum_{I \in \mathcal{A}_m(ns)} \mathbb{E} \left[ \prod_{i \in I} \zeta_i \right] \xrightarrow{n \rightarrow +\infty} K(s, s),$$

in particular these terms are bounded. All together, for any  $1 \leq k \leq \ell$  and any fixed partition  $\{J_1, \dots, J_k\} \in \mathcal{P}_{k,\ell}$  with at least one  $|J_a| \geq 3$ , we have

$$\begin{aligned} & \frac{1}{(c_r n^{3/2} \beta_n^r)^\ell} \sum_{\substack{I \in \llbracket \mathbf{1}, ns \rrbracket^\ell \\ \mathcal{J}(I) = \{J_1, \dots, J_k\}}} \prod_{a=1}^k \mathbb{E} \left[ \prod_{j \in J_a} \zeta_j \right] \\ & \leq \prod_{a=1}^k \left( \frac{1}{(c_r n^{3/2} \beta_n^r)^{|J_a|}} \sum_{I_r \in \mathcal{A}_{|J_a|}(ns)} \mathbb{E} \left[ \prod_{i \in I_a} \zeta_i \right] \right) \xrightarrow{n \rightarrow +\infty} 0, \end{aligned} \quad (4.2.15)$$

where we simply dropped the condition that  $i_j \not\leftrightarrow i_{j'}$  if  $j, j'$  are in different  $J_a$ 's, and wrote  $I_a = \{i_j\}_{j \in J_a}$ .



Recall (4.2.13). From the above, we can now restrict the sum to having all  $|J_a| \geq 2$ .

(i) If  $\ell$  is odd, then it imposes that one  $|J_a|$  is larger or equal than 3. Hence  $\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell]$  goes to 0, which proves the first part of (4.2.9).

(ii) If  $\ell$  is even, then the only part contributing to the sum in (4.2.13) comes from having all  $|J_a| = 2$  for all  $a$ : we denote  $\mathcal{P}_\ell$  the set of pairings of  $\{1, \dots, \ell\}$  (i.e. the sets of partitions  $\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell/2, \ell}$  with  $|J_a| = 2$  for all  $a$ ), so that

$$\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell] = o(1) + \sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_\ell} \sum_{\substack{\mathbf{I} \in \llbracket \mathbf{1}, ns \rrbracket^\ell \\ \mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_{\ell/2}\}}} \prod_{a=1}^{\ell/2} \frac{1}{c_r^2 n^3 \beta_n^{2r}} \mathbb{E} \left[ \prod_{j \in J_a} \zeta_{i_j} \right]. \quad (4.2.16)$$

Now, denote  $\Upsilon_\ell = \{(\mathbf{i}_1, \dots, \mathbf{i}_\ell) \in (\mathbb{N}^2)^\ell, \exists j, j' \text{ s.t. } \mathbf{i}_j = \mathbf{i}_{j'}\}$ . Analogously to (4.2.15), we get that for a fixed  $\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_\ell$ , the sum over  $\mathbf{I} \in \llbracket \mathbf{1}, ns \rrbracket^\ell \cap \Upsilon_\ell$  with  $\mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_{\ell/2}\}$  goes to 0: indeed, there must be some  $J_a = \{j_1, j_2\}$  with  $\mathbf{i}_{j_1} = \mathbf{i}_{j_2}$ , and Lemma 4.2.1 gives that  $\mathbb{E}[\zeta_{i_{j_1}}^2] = O(\beta_n^2)$ , so that  $(n^3 \beta_n^{2r})^{-1} \sum_{i_{j_1} \in \llbracket \mathbf{1}, ns \rrbracket} \mathbb{E}[\zeta_{i_{j_1}}^2] = O(n^{-2} \beta_n^{2(1-r)})$  goes to 0 (and all the other terms are bounded).

As a consequence, the restriction of the sum to  $\mathbf{I} \in \llbracket \mathbf{1}, ns \rrbracket^\ell \cap \Upsilon_\ell$  in (4.2.16) goes to 0, and we have

$$\mathbb{E}[(\overline{M}_n(\mathbf{s}))^\ell] = o(1) + \sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_\ell} \sum_{\substack{\mathbf{I} \in \llbracket \mathbf{1}, ns \rrbracket^\ell \cap \Upsilon_\ell^c \\ \mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_{\ell/2}\}}} \prod_{a=1}^{\ell/2} \frac{1}{c_r^2 n^3 \beta_n^{2r}} \mathbb{E} \left[ \prod_{j \in J_a} \zeta_{i_j} \right]. \quad (4.2.17)$$

Then, Lemma 4.2.1 (or (4.2.2)) gives that  $\mathbb{E}[\prod_{j \in J_a} \zeta_{i_j}] = (1 + o(1)) c_r^2 \beta_n^{2r}$  for all  $a$ . Moreover, there are  $((1 + o(1)) s_1 s_2 (s_1 + s_2) n^3)^{\ell/2}$  terms in the sum (recall Figure 4.1). All together, we get the second part of (4.2.9), using that the number of pairings  $\text{Card}(\mathcal{P}_\ell)$  is the correct combinatorial factor.  $\square$

*Proof of Proposition 4.2.3.* We now prove the finite-dimensional convergence. Let  $m \in \mathbb{N}$ ,  $\mathbf{s}_1, \dots, \mathbf{s}_m \in \mathbb{R}_+^2$  and  $(\mathcal{M}(\mathbf{s}_1), \dots, \mathcal{M}(\mathbf{s}_m))$  be a Gaussian vector of covariance matrix  $\Sigma_{\mathbf{s}_1, \dots, \mathbf{s}_m}$ . We show that for any  $u_1, \dots, u_m \in \mathbb{R}$ ,  $(\sum_{k=1}^m u_k \overline{M}_n(\mathbf{s}_k))_{n \geq 1}$  converges in distribution to  $\sum_{k=1}^m u_k \mathcal{M}(\mathbf{s}_k)$ , by showing the convergence of its moments. Let  $\ell \in \mathbb{N}$ , and let us compute

$$\mathbb{E} \left[ \left( \sum_{k=1}^m u_k \overline{M}_n(\mathbf{s}_k) \right)^\ell \right] = \sum_{k_1=1}^m \cdots \sum_{k_\ell=1}^m \mathbb{E} \left[ \prod_{j=1}^{\ell} u_{k_j} \overline{M}_n(\mathbf{s}_{k_j}) \right]. \quad (4.2.18)$$

We fix  $k_1, \dots, k_\ell \in \{1, \dots, m\}$ , and consider

$$\mathbb{E} \left[ \prod_{j=1}^{\ell} \overline{M}_n(\mathbf{s}_{k_j}) \right] = \left( \frac{1}{c_r n^{3/2} \beta_n^r} \right)^\ell \sum_{\mathbf{i}_1 \in \llbracket \mathbf{1}, s_{k_1} \rrbracket} \cdots \sum_{\mathbf{i}_\ell \in \llbracket \mathbf{1}, s_{k_\ell} \rrbracket} \mathbb{E} \left[ \prod_{j=1}^{\ell} \zeta_{i_j} \right]. \quad (4.2.19)$$

Then, we proceed as for the proof of Lemma 4.2.4: to each  $\ell$ -uple  $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_\ell)$ , we associate a partition  $\mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_k\}$  by decomposing  $\mathbf{I}$  into disjoint “ $|J_a|$ -aligned”

indices. As we showed above, cf. (4.2.15), the contribution of the terms with some  $|J_a| \neq 2$  goes to 0. First of all, this implies that if  $\ell$  is odd, then  $\mathbb{E}[\prod_{j=1}^{\ell} \overline{M}_n(\mathbf{s}_{k_j})]$  goes to 0 as  $n \rightarrow +\infty$ . If  $\ell$  is even, then analogously to (4.2.16)-(4.2.17), the main contribution to (4.2.19) comes from pairings  $\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell}$ , and from  $\mathbf{I} \in \llbracket \mathbf{1}, n\mathbf{s}_{k_1} \rrbracket \times \dots \times \llbracket \mathbf{1}, n\mathbf{s}_{k_{\ell}} \rrbracket$  with distinct entries. We therefore have that

$$\begin{aligned} \mathbb{E} \left[ \prod_{j=1}^{\ell} \overline{M}_n(\mathbf{s}_{k_j}) \right] &= o(1) + \sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell}} \sum_{\substack{\mathbf{I} \in \llbracket \mathbf{1}, n\mathbf{s}_{k_1} \rrbracket \times \dots \times \llbracket \mathbf{1}, n\mathbf{s}_{k_{\ell}} \rrbracket, \\ \mathbf{I} \notin \Upsilon_{\ell}, \mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_{\ell/2}\}}} \prod_{a=1}^{\ell/2} \frac{1}{c_r^2 n^3 \beta_n^4} \mathbb{E} \left[ \prod_{j \in J_a} \zeta_{i_j} \right] \\ &= (1 + o(1)) \sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell}} \prod_{a=1}^{\ell/2} K(\mathbf{s}_{k_{j_a}}, \mathbf{s}_{k_{j'_a}}), \end{aligned}$$

where we denoted  $J_a = \{j_a, j'_a\}$ . Here, we used again Lemma 4.2.1, and that for any fixed  $\{J_1, \dots, J_{\ell/2}\}$ , the number of terms in the sum over  $\mathbf{I}$  with  $\mathcal{J}(\mathbf{I}) = \{J_1, \dots, J_{\ell/2}\}$  has  $(1 + o(1)) \prod_{a=1}^{\ell/2} (K(\mathbf{s}_{k_{j_a}}, \mathbf{s}_{k_{j'_a}}) n^3)$  terms, see Figure 4.1.

Going back to (4.2.18), we get that if  $\ell$  is odd, then the  $\ell$ -th moment goes to 0. If  $\ell$  is even, we have that it is  $(1 + o(1))$  times

$$\sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell}} \sum_{k_1, \dots, k_{\ell/2}=1}^m \prod_{a=1}^{\ell/2} u_{k_{j_a}} u_{k_{j'_a}} K(\mathbf{s}_{k_{j_a}}, \mathbf{s}_{k_{j'_a}}) = \sum_{\{J_1, \dots, J_{\ell/2}\} \in \mathcal{P}_{\ell}} \left( \sum_{k, k'=1}^m u_k u_{k'} K(\mathbf{s}_k, \mathbf{s}_{k'}) \right)^{\ell/2}.$$

All together, we have shown that for any  $u_1, \dots, u_m \in \mathbb{R}$  and any  $\ell \in \mathbb{N}$ ,

$$\mathbb{E} \left[ \left( \sum_{k=1}^m u_k \overline{M}_n(\mathbf{s}_k) \right)^{\ell} \right] \xrightarrow{n \rightarrow +\infty} \frac{\ell!}{2^{\ell/2} (\ell/2)!} \left( \sum_{k, k'=1}^m u_k u_{k'} K(\mathbf{s}_k, \mathbf{s}_{k'}) \right)^{\ell/2} \quad (4.2.20)$$

if  $\ell$  is even (it converges to 0 if  $\ell$  is odd). Note that the term raised to the power  $\ell/2$  is the variance of  $\sum_{k=1}^m u_k \mathcal{M}(\mathbf{s}_k)$ . This shows that for any  $\ell \in \mathbb{N}$ , the  $\ell$ -th moment of  $\sum_{k=1}^m u_k \overline{M}_n(\mathbf{s}_k)$  converges as  $n \rightarrow +\infty$  to the  $\ell$ -th moment of  $\sum_{k=1}^m u_k \mathcal{M}(\mathbf{s}_k)$ . Since  $(\mathcal{M}(\mathbf{s}_k))_{1 \leq k \leq m}$  is a Gaussian vector, this implies the convergence in distribution of  $(\overline{M}_n(\mathbf{s}_k))_{1 \leq k \leq m}$  to  $(\mathcal{M}(\mathbf{s}_k))_{1 \leq k \leq m}$ .  $\square$

### 4.2.3 Convergence of $(\overline{M}_n(\mathbf{s}))_{\mathbf{s} \in [0, \mathbf{t}]}$

In this section we prove that the sequence  $(\overline{M}_n(\mathbf{s}))_{\mathbf{s} \in [0, \mathbf{t}]}$ , converges in distribution to  $(\mathcal{M}(\mathbf{s}))_{\mathbf{s} \in [0, \mathbf{t}]}$ . Let us first introduce a continuous interpolation of  $\overline{M}_n$ . For any  $n \in \mathbb{N}$  and  $\mathbf{s} \in \mathbb{R}_+^2$ , let  $\mathbf{s}^{[n]} := \frac{1}{n} \lfloor n\mathbf{s} \rfloor$  (not to be confused with the projection  $\mathbf{s}^{(a)} \in \mathbb{R}$ ,  $a \in \{1, 2\}$ ); notice that  $\mathbf{s}^{[n]} \in \frac{1}{n} \mathbb{Z}^2$ , and  $\|\mathbf{s}^{[n]} - \mathbf{s}\|_1 \leq \frac{2}{n}$ . Define for any  $\mathbf{s} \in [0, \mathbf{t}]$ ,

$$\begin{aligned} \widetilde{M}_n(\mathbf{s}) &= (1 - \gamma_1)(1 - \gamma_2) \overline{M}_n(\mathbf{s}^{[n]}) + \gamma_1(1 - \gamma_2) \overline{M}_n(\mathbf{s}^{[n]} + \frac{1}{n}(1, 0)) \\ &\quad + \gamma_2(1 - \gamma_1) \overline{M}_n(\mathbf{s}^{[n]} + \frac{1}{n}(0, 1)) + \gamma_2 \gamma_1 \overline{M}_n(\mathbf{s}^{[n]} + \frac{1}{n}(1, 1)), \end{aligned} \quad (4.2.21)$$

where  $(\gamma_1, \gamma_2) := n(\mathbf{s} - \mathbf{s}^{[n]}) \in [0, 1]^2$ , and where we use the convention  $\overline{M}_n(\mathbf{s}) := \overline{M}_n(\mathbf{s} \wedge \mathbf{t})$  if  $\mathbf{s} \in \mathbb{R}_+^2 \setminus [0, \mathbf{t}]$ . Note that  $\widetilde{M}_n$  is a continuous random field, which satisfies  $\widetilde{M}_n(\mathbf{s}^{[n]}) = \overline{M}_n(\mathbf{s}^{[n]})$  for all  $\mathbf{s} \in [0, \mathbf{t}]$ ,  $n \in \mathbb{N}$ .

We prove the following Lemma.

**Lemma 4.2.5.** *Under the assumptions of Theorem 4.1.2, one has for any  $p < +\infty$ ,*

$$\|\widetilde{M}_n - \overline{M}_n\|_\infty \xrightarrow[n \rightarrow \infty]{L^p} 0. \quad (4.2.22)$$

Recall that  $\|\widetilde{M}_n - \overline{M}_n\|_\infty := \sup_{\mathbf{0} \preceq \mathbf{s} \preceq \mathbf{t}} |\widetilde{M}_n(\mathbf{s}) - \overline{M}_n(\mathbf{s})|$ , which is a real-valued random variable. Recall also that for  $p \in \mathbb{N}$ ,  $\zeta_1 \in L^p(\mathbb{P})$  as soon as  $n$  is sufficiently large.

*Proof.* We prove the result for  $p \in 2\mathbb{N}$  (we take  $p > 8$ ) and  $n$  sufficiently large, so that in particular  $\zeta_1 \in L^p(\mathbb{P})$ . First, notice that for  $n \in \mathbb{N}$  and  $\mathbf{s} \in [0, \mathbf{t}]$ , one has

$$|\widetilde{M}_n(\mathbf{s}) - \overline{M}_n(\mathbf{s})| \leq \max \left\{ \begin{array}{l} \left| \overline{M}_n(\mathbf{s}^{[n]}) - \overline{M}_n(\mathbf{s}^{[n]} + (\frac{1}{n}, 0)) \right|, \\ \left| \overline{M}_n(\mathbf{s}^{[n]}) - \overline{M}_n(\mathbf{s}^{[n]} + (0, \frac{1}{n})) \right|, \\ \left| \overline{M}_n(\mathbf{s}^{[n]}) - \overline{M}_n(\mathbf{s}^{[n]} + (\frac{1}{n}, \frac{1}{n})) \right| \end{array} \right\}. \quad (4.2.23)$$

Let us rewrite the last term, for  $\mathbf{i} \in \mathbb{N}_0^2$ ,

$$\overline{M}_n(\frac{1}{n}\mathbf{i}) - \overline{M}_n(\frac{1}{n}\mathbf{i} + (\frac{1}{n}, \frac{1}{n})) = \frac{1}{n^{3/2}\beta_n^r} \left( \sum_{j=1}^{i_1+1} \zeta_{j, i_2+1} + \sum_{j=1}^{i_2} \zeta_{i_1+1, j} \right).$$

Doing similarly with the two other terms, and using the inequality  $(a + b)^p \leq 2^p(a^p + b^p)$  (recall  $p \in 2\mathbb{N}$ ), we obtain

$$(n^{3/2}\beta_n^2 \|\widetilde{M}_n - \overline{M}_n\|_\infty)^p \leq C_1 \sup_{\mathbf{0} \preceq \mathbf{i} \preceq n\mathbf{t}} \left( \sum_{j=1}^{i_1} \zeta_{(j, i_2+1)} \right)^p + C_1 \sup_{\mathbf{0} \preceq \mathbf{i} \preceq n\mathbf{t}} \left( \sum_{j=1}^{i_2} \zeta_{(i_1+1, j)} \right)^p, \quad (4.2.24)$$

for some  $C_1 > 0$ . Hölder's inequality and a simple Taylor expansion give  $\mathbb{E}[\zeta_{i_1} \cdots \zeta_{i_p}] \leq \mathbb{E}[\zeta_1^p] \leq c_p(\beta_n)^p$ , uniformly in  $n \in \mathbb{N}$  and  $\mathbf{i}_1, \dots, \mathbf{i}_p \in (\mathbb{N}^2)^p$ : this implies

$$\mathbb{E} \left[ \left( \sum_{j=1}^{i_1} \zeta_{(j, i_2+1)} \right)^p \right] \leq c_p (i_1)^p \beta_n^p.$$

Therefore, using that  $\sup_{\mathbf{0} \preceq \mathbf{i} \preceq n\mathbf{t}} u_i \leq \sum_{\mathbf{0} \preceq \mathbf{i} \preceq n\mathbf{t}} u_i$  for any non-negative family  $(u_i)_{i \in \mathbb{N}^2}$ , (4.2.24) gives

$$\mathbb{E} \left[ (n^{3/2}\beta_n^r \|\widetilde{M}_n - \overline{M}_n\|_\infty)^p \right] \leq C_2 n^{p+2} \beta_n^p, \quad (4.2.25)$$

uniformly in  $n$ . We therefore get that

$$\mathbb{E} \left[ \|\widetilde{M}_n - \overline{M}_n\|_\infty^p \right] \leq C_2 n^{2-p/2} \beta_n^{p(1-r)}.$$

If  $r = 1$ , then this goes to 0 as  $n \rightarrow +\infty$ , provided that  $p > 4$ . If  $r = 2$ , one can write the lower bound as  $C_2(n^{1/4}\beta_n)^{-p}n^{2-p/4}$ , which goes to 0 provided  $p > 8$ , since  $n^{1/4}\beta_n \rightarrow +\infty$ .  $\square$

We now have all the required estimates to finish the proof of Theorem 4.1.2.

*Proof of Theorem 4.1.2.* First, let us prove that  $(\widetilde{M}_n)_{n \in \mathbb{N}}$  converges to  $\mathcal{M}$  in distribution. Lemma 4.2.24 ensures us that for  $m \in \mathbb{N}$  and  $\mathbf{s}_1, \dots, \mathbf{s}_m \preccurlyeq \mathbf{t}$ , the vector  $(\widetilde{M}_n(\mathbf{s}_1) - \overline{M}_n(\mathbf{s}_1), \dots, \widetilde{M}_n(\mathbf{s}_m) - \overline{M}_n(\mathbf{s}_m))$  converges to  $(0, \dots, 0) \in \mathbb{R}^m$  in probability. Recalling Proposition 4.2.3 and applying Slutsky's theorem, this implies that  $(\widetilde{M}_n(\mathbf{s}_k))_{k=1}^m$  converges to  $(\mathcal{M}(\mathbf{s}_k))_{k=1}^m$  in distribution, yielding the finite-dimensional convergence. As for the tightness of  $(\widetilde{M}_n)_{n \geq 1}$ , it can be proven with a direct adaptation of Donsker's theorem (see [113, Theorem 4.1.1 and Proposition 4.3.1, Chapter 6] for the multidimensional variant). We do not write the details in this thesis, but they will be added in a later version of this work.

Finally, let  $h : L^\infty([0, \mathbf{t}]) \rightarrow \mathbb{R}$  be a bounded Lipschitz function, and let us prove that  $\lim_{n \rightarrow +\infty} \mathbb{E}[h(\widetilde{M}_n)] = \mathbb{E}[h(\mathcal{M})]$ . We write

$$\left| \mathbb{E}[h(\widetilde{M}_n)] - \mathbb{E}[h(\mathcal{M})] \right| \leq \left| \mathbb{E}[h(\widetilde{M}_n)] - \mathbb{E}[h(\overline{M}_n)] \right| + C \mathbb{E}[\|\widetilde{M}_n - \overline{M}_n\|_\infty], \quad (4.2.26)$$

for some  $C > 0$ . The convergence in distribution of  $(\widetilde{M}_n)_{n \in \mathbb{N}}$  implies that the first term goes to 0 as  $n \rightarrow \infty$ , and Lemma 4.2.5 shows the same for the second term. We conclude with the well-known Portmanteau theorem [29, Theorem 2.1], which proves the convergence in distribution of  $(\widetilde{M}_n)_{n \in \mathbb{N}}$  to  $\mathcal{M}$ .  $\square$

#### 4.2.4 Convergence of the field in $L^p(\mathbb{P})$

Before moving on to the definition of the integral against  $\mathcal{M}$ , let us state a last Claim which strengthens Theorem 4.1.2 on a convenient probability space.

**Claim 4.2.6.** *There exist a probability space  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$  and copies  $\widehat{M}_n, n \geq 1$  (resp.  $\widehat{\mathcal{M}}$ ) of  $\overline{M}_n, n \geq 1$  (resp.  $\mathcal{M}$ ) on that space, such that for  $\mathbf{s} \in [0, \mathbf{t}]$  and  $p \in [1, \infty)$ ,  $\widehat{M}_n(\mathbf{s}) \rightarrow \widehat{\mathcal{M}}(\mathbf{s})$  in  $L^p(\widehat{\mathbb{P}})$  as  $n \rightarrow \infty$ .*

*Proof.* This is a consequence of Skorokhod's representation theorem. Since the set of continuous functions on  $[0, \mathbf{t}]$  with the  $\|\cdot\|_\infty$ -topology is separable, there exist a probability space  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$  and copies  $\widehat{M}'_n, n \geq 1$  (resp.  $\widehat{\mathcal{M}}$ ) of  $\overline{M}_n, n \geq 1$  (resp.  $\mathcal{M}$ ) on that space, such that  $\widehat{M}'_n \rightarrow \widehat{\mathcal{M}}$   $\widehat{\mathbb{P}}$ -almost surely. Using the moment estimates from Lemma 4.2.4 and dominated convergence theorem, it follows that for  $\mathbf{s} \in [0, \mathbf{t}]$ ,  $\widehat{M}'_n(\mathbf{s}) \rightarrow \widehat{\mathcal{M}}(\mathbf{s})$  in  $L^p(\widehat{\mathbb{P}})$ . Finally, let  $\widehat{M}_n$  be a piecewise constant modification of  $\widehat{M}'_n$ , i.e.  $\widehat{M}_n(\mathbf{s}) := \widehat{M}'_n(\mathbf{s}^{[n]})$  for  $\mathbf{s} \in [0, \mathbf{t}]$ ,  $n \in \mathbb{N}$ ; then  $\widehat{M}_n$  has the same law as  $\overline{M}_n, n \in \mathbb{N}$ , and we conclude the proof with Lemma 4.2.5.  $\square$

### 4.3 Covariance measure of $\mathcal{M}$ and stochastic integral

In this section we define a stochastic integral against the field  $\mathcal{M}$ , that is we prove Proposition 4.1.3 and Theorem 4.1.7. Once this is done, we prove the integrability of  $\varphi$  against  $\mathcal{M}$ , that is Proposition 4.1.9.

### 4.3.1 Covariance measure: proof of Proposition 4.1.3

Recall the definitions of the semi-ring of rectangles

$$\mathcal{S}_t := \{[u, v]; \mathbf{0} \preceq u \preceq v \preceq t\} \cup \{\emptyset\},$$

and for a random function  $M : [\mathbf{0}, t] \rightarrow \mathbb{R}$ , the *increment* of  $M$  on  $A = [u, v] \in \mathcal{S}_t$  is given by

$$M(A) := M(v_1, v_2) - M(u_1, v_2) - M(v_1, u_2) + M(u_1, u_2).$$

For  $\mathcal{M}$  a Gaussian field on  $[\mathbf{0}, t]^2$  with covariance  $K$  defined in (4.1.10), the goal is to define a measure  $\nu_{\mathcal{M}}$  on  $\text{Bor}([\mathbf{0}, t]^2)$  such that for  $A, B \in \mathcal{S}_t^2$ ,

$$\nu_{\mathcal{M}}(A, B) := \nu_{\mathcal{M}}(A \times B) = \mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)]. \quad (4.3.1)$$

**Step 1.** Let us start by determining the expression of  $\nu_{\mathcal{M}}(A, B)$  for  $A, B \in \mathcal{S}_t$ . For any real numbers  $x \leq y \leq z$  and  $u \leq v$ , one has

$$\mathcal{M}([u, v] \times [x, z]) = \mathcal{M}([u, v] \times [x, y]) + \mathcal{M}([u, v] \times [y, z]). \quad (4.3.2)$$

Moreover for  $A, B \in \mathcal{S}_t$ , we can decompose them into finite unions of rectangles  $A = \cup_{i=1}^p A_i$  and  $B = \cup_{j=1}^q B_j$  such that for  $1 \leq i \leq p$ ,  $1 \leq j \leq q$  (we can take  $p, q \leq 9$ ), one of the following holds :

- (a)  $A_i = B_j$ .
- (b) There exist  $u_0 \leq u_1$  and  $s_0 \leq s_1 \leq t_0 \leq t_1$  such that either  $A_i = [u_0, u_1] \times [s_0, s_1]$  and  $B_j = [u_0, u_1] \times [t_0, t_1]$  or  $A_i = [s_0, s_1] \times [u_0, u_1]$  and  $B_j = [t_0, t_1] \times [u_0, u_1]$ , or the other way around.
- (c) For  $a \in \{1, 2\}$ , the projections of  $A_i, B_j$  on the  $a$ -th coordinate are disjoint.

This implies that we only have to compute the covariances of increments  $\mathcal{M}(A), \mathcal{M}(B)$  for couples of rectangles  $(A, B)$  satisfying one of the above, and we will obtain covariances of all rectangles thanks to (4.3.2) and the bilinearity of  $(X, Y) \mapsto \mathbb{E}[XY]$ . We do so in the following Lemma.

**Lemma 4.3.1.** *Let  $\mathcal{M}$  be a Gaussian field on  $\mathbb{R}_+^2$  with covariance  $K$  defined in (4.1.10). Let  $u_0 \leq u_1$  and  $s_0 \leq s_1 \leq t_0 \leq t_1$ .*

- (a) *If  $A = B = [u_0, u_1] \times [s_0, s_1]$ , then*

$$\mathbb{E}[\mathcal{M}(A)^2] = (u_1 - u_0)(s_1 - s_0)(u_1 - u_0 + s_1 - s_0).$$

- (b) *If  $A = [u_0, u_1] \times [s_0, s_1]$  and  $B = [u_0, u_1] \times [t_0, t_1]$ , then*

$$\mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)] = (u_1 - u_0)(s_1 - s_0)(t_1 - t_0).$$

- (c) *If the projections of  $A, B$  on the  $a$ -th coordinate are disjoint for  $a \in \{1, 2\}$ , then  $\mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)] = 0$ .*

*Proof.* We only display the proof in the second case  $A = [u_0, u_1) \times [s_0, s_1)$  and  $B = [u_0, u_1) \times [t_0, t_1)$ , since the other two are very similar. Let us first rewrite

$$\mathcal{M}(A) = \sum_{i,j \in \{0,1\}} (-1)^{i+j} \mathcal{M}(u_i, s_j).$$

Thus,

$$\begin{aligned} \mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)] &= \sum_{i,j,k,l \in \{0,1\}} (-1)^{i+j+k+l} \mathbb{E}[\mathcal{M}(u_i, s_j)\mathcal{M}(u_k, t_l)] \\ &= \sum_{i,j,k,l \in \{0,1\}} (-1)^{i+j+k+l} (u_i \wedge u_k)(s_j \wedge t_l)(u_i \vee u_k + s_j \vee t_l) \\ &= \sum_{i,j,k,l \in \{0,1\}} (-1)^{i+j+k+l} (u_i \wedge u_k)(s_j)(u_i \vee u_k + t_l), \end{aligned}$$

where we used  $s_0 \leq s_1 \leq t_0 \leq t_1$ . Let us develop the last factor to rewrite  $\mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)]$  as a sum of two terms: in the first one, we can factorize  $\sum_{l=0}^1 (-1)^l = 0$ , so it remains

$$\mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)] = \left( \sum_{j=0}^1 (-1)^j s_j \right) \left( \sum_{l=0}^1 (-1)^l t_l \right) \left( \sum_{i,k \in \{0,1\}} (-1)^{i+k} (u_i \wedge u_k) \right),$$

and a straightforward computation gives the result.  $\square$

Note that we can rewrite those expressions for any  $A, B \in \mathcal{S}_t$  as

$$\nu_{\mathcal{M}}(A, B) := \mathbb{E}[\mathcal{M}(A)\mathcal{M}(B)] = \int_{[0,t)} \mathbf{1}_A(\mathbf{s}) \left( \int_0^{t_1} \mathbf{1}_B(x, s_2) dx + \int_0^{t_2} \mathbf{1}_B(s_1, y) dy \right) d\mathbf{s}. \quad (4.3.3)$$

where Lemma 4.3.1 proves this identity for  $A, B$  satisfying (a), (b) or (c), and generic couples of rectangles are handled by bilinearity and (4.3.2). This quantity  $\nu_{\mathcal{M}}(A, B)$  can also be rewritten,

$$\begin{aligned} \nu_{\mathcal{M}}(A, B) &= \lambda_3 \left( (x, y, z) \in \mathbb{R}^3; (x, y) \in A \text{ and } (x, z) \in B \right) \\ &\quad + \lambda_3 \left( (x, y, z) \in \mathbb{R}^3; (x, y) \in A \text{ and } (z, y) \in B \right), \end{aligned} \quad (4.3.4)$$

where  $\lambda_3$  denotes the 3-dimensional Lebesgue measure.

**Step 2.** We now extend the definition of  $\nu_{\mathcal{M}}$  to  $\text{Bor}([0, \mathbf{t})^2)$ , and verify that it is a measure. More generally, we give a sufficient condition for the covariance function of a field  $X$  to define a measure on  $\text{Bor}([0, \mathbf{t})^2)$ .

**Proposition 4.3.2.** *Let  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  be a random field. Define the function  $\nu$  on  $\mathcal{S}_t^2$  by*

$$\nu(A \times B) := \mathbb{E}[X(A)X(B)], \quad A, B \in \mathcal{S}_t.$$

*Assume that  $\nu$  satisfies the following:*

- (a)  $\nu(A, B) \geq 0$  for  $A, B \in \mathcal{S}_t$ ,
- (b)  $\nu([\mathbf{0}, \mathbf{t}]^2) < \infty$ ,
- (c)  $\nu$  is additive on  $\mathcal{S}_t^2$ ,
- (d) there exists a finite measure  $\mu$  on  $\text{Bor}([\mathbf{0}, \mathbf{t}]^2)$  such that for  $A, B \in \mathcal{S}_t$ ,  $\nu(A \times B) \leq \mu(A \times B)$ ,

then  $\nu$  can be extended to a finite measure on  $\text{Bor}([\mathbf{0}, \mathbf{t}]^2)$ , which is unique.

We prove this proposition below. Let us recall that additivity on  $\mathcal{S}_t^2$  is a direct consequence of relation (4.3.2).

For the field  $\mathcal{M}$ , the relation (4.3.4) proves that  $\nu_{\mathcal{M}}$  satisfies all necessary assumptions to define a measure on  $\text{Bor}([\mathbf{0}, \mathbf{t}]^2)$ , and even yields an explicit expression; in particular, (4.1.24) directly follows from (4.3.3). Notice that similar observations hold for the Gaussian white noise with (4.1.25), hence it has a well-posed covariance measure on  $\text{Bor}([\mathbf{0}, \mathbf{t}]^2)$ .

*Proof.* By assumption  $\nu$  is additive on  $\mathcal{S}_t^2$  which is a semi-ring of sets, so in order to extend it into a measure on  $\sigma(\mathcal{S}_t^2) = \text{Bor}([\mathbf{0}, \mathbf{t}]^2)$  with Carathéodory's theorem, it only remains to show that it is sigma-additive on  $\mathcal{S}_t^2$  (and the extension will be unique because of assumption (b)).

First, let us prove that  $\nu$  is non-decreasing on  $\mathcal{S}_t^2$ : let  $A_1, A_2 \in \mathcal{S}_t^2$  such that  $A_1 \subset A_2$ . Since  $\mathcal{S}_t^2$  is a semi-ring,  $A_2 \setminus A_1 = \cup_{i=1}^p B_i$  for some  $p \in \mathbb{N}$  and disjoint  $B_i \in \mathcal{S}_t^2$ ,  $1 \leq i \leq p$ . Using that  $\nu$  is additive on  $\mathcal{S}_t^2$  and assumption (a), we write

$$\nu(A_2) = \nu(A_1) + \sum_{i=1}^p \nu(B_i) \geq \nu(A_1),$$

which proves the statement. Moreover, this can straightforwardly be extended to  $A_1, A_2 \in \mathcal{R}(\mathcal{S}_t^2)$  which is the ring generated by  $\mathcal{S}_t^2$  (i.e. all finite unions of rectangles).

Now let  $A_i \in \mathcal{S}_t^2$ ,  $i \in \mathbb{N}$ , such that  $A := \cup_{i \geq 1} A_i \in \mathcal{S}_t^2$ . We may assume  $(A_i)_{i \geq 1}$  disjoint without loss of generality. For  $n \in \mathbb{N}$ , we have

$$\nu(\cup_{i \geq 1} A_i) \geq \nu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \nu(A_i),$$

and taking the limit as  $n \rightarrow \infty$ , we obtain  $\nu(\cup_{i \geq 1} A_i) \geq \sum_{i \geq 1} \nu(A_i)$ . Let us now show that  $\nu(\cup_{i \geq n} A_i)$  vanishes to 0 as  $n \rightarrow \infty$ , which will conclude the proof. For  $n \in \mathbb{N}$ , notice that  $\cup_{i \geq 1} A_i \in \mathcal{S}_t^2$  and  $A_i \in \mathcal{S}_t^2$ ,  $1 \leq i \leq n-1$ . Since  $\mathcal{S}_t^2$  is a semi-ring of sets,  $\cup_{i \geq n} A_i$  can be written as a finite union  $\cup_{j=1}^p B_j$  for some  $B_j \in \mathcal{S}_t^2$  disjoint,  $1 \leq j \leq p$ . Thereby,

$$0 \leq \nu(\cup_{i \geq n} A_i) = \sum_{j=1}^p \nu(B_j) \leq \sum_{j=1}^p \mu(B_j) = \mu(\cup_{i \geq n} A_i),$$

where we used assumption (d). Since  $\mu$  is a genuine finite measure,

$$\lim_{n \rightarrow \infty} \mu(\cup_{i \geq n} A_i) = \mu(\cap_{n \geq 1} \cup_{i \geq n} A_i) = \mu(\emptyset) = 0,$$

which concludes the proof of the proposition. □

### 4.3.2 Construction of the integral: proof of Proposition 4.1.5 and Theorem 4.1.7

Let  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  be a random field on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , and assume that there exists  $\nu$  a finite measure on  $\text{Bor}([0, t]^2)$  such that for  $A, B \in \mathcal{S}_t$ ,

$$\mathbb{E}[X(A)X(B)] = \nu(A \times B).$$

In this section we establish the properties of  $L_X^2$  and  $\langle \cdot, \cdot \rangle_\nu$  (recall (4.1.28-4.1.29)) and define the integral against  $X$ . We divide this section into three parts: first, we need to extend the definition of the integral to  $\{1_B; B \in \text{Bor}([0, t])\}$ , then we use it to prove Proposition 4.1.5, and finally we extend the integral to  $L_X^2$  and prove Theorem 4.1.7.

**Extension to indicator functions.** Let  $\mathcal{R}_t$  be the ring of sets generated by  $\mathcal{S}_t$ , which is given by all finite unions of rectangles (recall that a ring of sets  $\mathcal{R}$  is a non-empty class of sets, stable by finite union and difference). We state a useful lemma from measure theory first.

**Lemma 4.3.3.** *Let  $(E, \mathcal{E}, \mu)$  be a measured space such that  $\mu(E) < \infty$ , and let  $\mathcal{R} \subset \mathcal{P}(E)$  be a ring of sets such that  $E \in \mathcal{R}$  and  $\sigma(\mathcal{R}) = \mathcal{E}$ . Then for any  $B \in \mathcal{E}$  and  $n \in \mathbb{N}$ , there exists  $A_n \in \mathcal{R}$  such that*

$$\mu(B \Delta A_n) \leq 2^{-n}, \quad (4.3.5)$$

where  $\Delta$  denotes the symmetric difference.

This lemma is proven afterwards. Notice that for  $C \in \mathcal{R}_t$ , it can be written as a disjoint union of rectangles  $\cup_{i=1}^k C_i$ , and we can define

$$1_C \diamond X := \sum_{i=1}^k (1_{C_i} \diamond X) = \sum_{i=1}^k X(C_i) =: X(C), \quad (4.3.6)$$

which does not depend on the chosen decomposition (recall (4.3.2)). Moreover (4.1.30) applies to  $1_A, 1_B$  for  $A, B \in \mathcal{R}_t$  by bilinearity.

Let  $B \in \text{Bor}([0, t])$ , and let  $A_n \in \mathcal{R}_t$ ,  $n \geq 1$  be a sequence satisfying (4.3.5) for the measure  $\mu(A) := \nu(A \times [0, t])$  on  $\text{Bor}([0, t])$ . Then  $1_{A_n} \diamond X \in L^2(\mathbb{P})$ ,  $n \geq 1$  is well defined, and for  $p, q \geq 1$ ,

$$\begin{aligned} \|1_{A_p} \diamond X - 1_{A_q} \diamond X\|_{L^2}^2 &= \mathbb{E}[(X(A_p) - X(A_q))^2] \\ &= \int_{[0, t]^2} (1_{A_p} - 1_{A_q})(\mathbf{u}) \times (1_{A_p} - 1_{A_q})(\mathbf{v}) \, d\nu(\mathbf{u}, \mathbf{v}) \\ &\leq \int_{[0, t]^2} 1_{A_p \Delta A_q}(\mathbf{u}) \, d\nu(\mathbf{u}, \mathbf{v}) = \mu(A_p \Delta A_q), \end{aligned}$$

where we used  $|1_{A_p} - 1_{A_q}| = 1_{A_p \Delta A_q} \leq 1$ . Since  $A_p \Delta A_q \subset (A_p \Delta B) \cup (B \Delta A_q)$ , it follows that  $\mu(A_p \Delta A_q) \leq 2^{-p} + 2^{-q}$ , thus  $(1_{A_n} \diamond X)_{n \geq 1}$  is a Cauchy sequence in  $L^2(\mathbb{P})$ . By completeness, it thereby has a limit which we denote  $1_B \diamond X$  and that does not depend



on the chosen sequence  $(A_n)_{n \geq 1}$  such that  $\mu(B \Delta A_n) \rightarrow 0, n \rightarrow \infty$ . Moreover,  $(1_{A_n})_{n \geq 1}$  converges  $\mu$ -a.s. to  $1_B$ , and  $(1_{A_n} \diamond X)_{n \geq 1}$  converges to  $1_B \diamond X$  in  $L^2(\mathbb{P})$ ; thus the functions  $1_B \diamond X, B \in \text{Bor}([0, t])$ , do satisfy (4.1.30) by dominated convergence.

*Proof of Lemma 4.3.3.* Recall that  $(E, \mathcal{E}, \mu)$  is a measured space and  $\mathcal{R}$  a ring of sets such that  $E \in \mathcal{R}, \sigma(\mathcal{R}) = \mathcal{E}$  and  $\mu(E) < \infty$ . Let us define for  $n \in \mathbb{N}$ ,

$$\mathcal{A}_n := \{B \in \mathcal{E}; \exists A \in \mathcal{R}, \mu(B \Delta A) \leq 2^{-n}\},$$

and  $\mathcal{A} = \bigcap_{n \geq 1} \mathcal{A}_n$ . It is clear that  $\mathcal{R} \subset \mathcal{A}_n$  for all  $n \in \mathbb{N}$ , so  $\mathcal{R} \subset \mathcal{A}$ . Let us prove that  $\mathcal{A}$  is a Dynkin system, which will conclude the proof by Dynkin's  $\pi$ - $\lambda$  theorem.

First, we clearly have  $E \in \mathcal{R} \subset \mathcal{A}$ , so  $\mathcal{A}$  is non empty. Let  $B_1, B_2 \in \mathcal{A}$  such that  $B_1 \subset B_2$ , and  $n \in \mathbb{N}$ . By assumption there exist  $A_1, A_2 \in \mathcal{R}$  such that  $\mu(B_a \Delta A_a) \leq 2^{-n-1}, a \in \{1, 2\}$ . Since  $\mathcal{R}$  is a ring, we have  $A_2 \setminus A_1 \in \mathcal{R}$ , and we also notice

$$(B_2 \setminus B_1) \Delta (A_2 \setminus A_1) \subset (B_2 \Delta A_2) \cup (B_1 \Delta A_1).$$

Hence it follows that  $B_2 \setminus B_1 \in \mathcal{A}_n$  for all  $n \in \mathbb{N}$ , thus  $\mathcal{A}$  is stable by difference.

Let  $B_k \in \mathcal{A}$  such that  $B_k \subset B_{k+1}, k \geq 1$ , and let  $B = \bigcup_{k \geq 1} B_k$ . Let  $n \in \mathbb{N}$ . Notice that, since  $\mu(B) < \infty$ , there exists  $k_0 \in \mathbb{N}$  such that  $\mu(B \setminus B_{k_0}) \leq 2^{-n-1}$ . Moreover  $B_{k_0} \in \mathcal{A}$ , so there exists  $A \in \mathcal{R}$  such that  $\mu(B_{k_0} \Delta A) \leq 2^{-n-1}$ . Thus,

$$\mu(B \Delta A) \leq \mu(B \setminus B_{k_0}) + \mu(B_{k_0} \Delta A) \leq 2^{-n},$$

which finishes the proof. □

**Proof of Proposition 4.1.5.** Consider the application

$$(g, h) \mapsto \langle g, h \rangle_\nu = \int_{[0, t]^2} g(\mathbf{u}) h(\mathbf{v}) d\nu(\mathbf{u}, \mathbf{v}),$$

which is well-defined (possibly infinite) on non-negative, measurable functions. We claim that it is bilinear, symmetric and positive semi-definite: indeed, we have proven that the isometry (4.1.30) is satisfied for indicator functions, which implies those properties; and they can be extended to non-negative functions  $f = \sum_{n \geq 1} c_n 1_{B_n}, c_n \geq 0, B_n \in \text{Bor}([0, t]), n \geq 1$  by monotone convergence.

In order to prove that  $L^2_X$  is a vector space (notice that it is not straightforward from definition), we first have to prove a Cauchy-Schwarz inequality for non-negative functions; that is for  $g, h$  non-negative measurable functions on  $[0, t]$ :

$$\langle g, h \rangle_\nu \leq \|g\|_\nu \|h\|_\nu, \quad \forall g, h \geq 0. \quad (4.3.7)$$

To show this, let us define  $\mathcal{G}$  the set of finite linear combinations of indicator functions  $g = \sum_{i=1}^n c_i 1_{A_i}, A_i \in \text{Bor}([0, t]), c_i \in \mathbb{R}, 1 \leq i \leq n$ . Notice that, using the bilinearity of  $\langle \cdot, \cdot \rangle_\nu$ , the application  $g \mapsto g \diamond X$  can be extended to  $\mathcal{G}$ , and satisfies the isometry property (4.1.30) for all  $g, h \in \mathcal{G}$ . With those observations in mind, rewrite  $g = \sum_{i \geq 1} c_i 1_{A_i}$  and

$h = \sum_{j \geq 1} d_j \mathbf{1}_{B_j}$  for some  $c_i, d_j \geq 0$ ,  $A_i, B_j \in \text{Bor}([\mathbf{0}, \mathbf{t}])$ ,  $i, j \geq 1$ . For  $n \in \mathbb{N}$ , define  $g_n = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$  and  $h_n = \sum_{j=1}^n d_j \mathbf{1}_{B_j}$ , and notice  $g_n, h_n \in \mathcal{G}$ ; thereby,

$$\langle g_n, h_n \rangle_\nu = \mathbb{E}[(g_n \diamond X)(h_n \diamond X)] \leq \mathbb{E}[(g_n \diamond X)^2]^{1/2} \mathbb{E}[(h_n \diamond X)^2]^{1/2} = \|g_n\|_\nu \|h_n\|_\nu,$$

where we used (4.1.30) twice on  $\mathcal{G}$ , and Cauchy-Schwarz inequality on  $L^2(\mathbb{P})$ . We conclude the proof of (4.3.7) with the monotone convergence theorem as  $n \rightarrow \infty$ .

Therefore, for  $g, h \in L_X^2$ , inequality (4.3.7) implies that  $\langle |g|, |h| \rangle_\nu < \infty$ , and thus  $g + h \in L_X^2$ ; this finally proves that  $L_X^2$  is a vector space. Moreover, the observation that  $\langle \cdot, \cdot \rangle_\nu$  is bilinear, symmetric and positive semi-definite can also be extended on  $L_X^2$  using (4.1.30) on  $\mathcal{G}$  and dominated convergence. This fully proves Proposition 4.1.5-(i-ii).

Regarding the Cauchy-Schwarz inequality, we may extend it from non-negative functions (recall (4.3.7)) to  $L_X^2$ : indeed, write for  $g, h \in L_X^2$  and  $\lambda \in \mathbb{R}$ ,

$$0 \leq \|g + \lambda h\|_\nu^2 = \|g\|_\nu^2 + 2\lambda \langle g, h \rangle_\nu + \lambda^2 \|h\|_\nu^2,$$

and since  $\langle g, h \rangle_\nu \in \mathbb{R}$  is well-posed, we deduce that the quadratic polynomial in  $\lambda$  above has a non-positive discriminant, which concludes the proof of (iii).

Finally, we deduce that a triangle inequality (iv) holds for  $\|\cdot\|_\nu$ : for  $g, h \in L_X^2$ , write

$$\|g + h\|_\nu^2 = \|g\|_\nu^2 + 2\langle g, h \rangle_\nu + \|h\|_\nu^2 \leq \|g\|_\nu^2 + 2\|g\|_\nu \|h\|_\nu + \|h\|_\nu^2 = (\|g\|_\nu + \|h\|_\nu)^2, \quad (4.3.8)$$

which proves the inequality.

**Proof of Theorem 4.1.7.** We finally extend the definition of the integral to  $L_X^2$ . Recall that  $\mathcal{G}$  is the set of finite linear combinations of indicator functions, that  $g \mapsto g \diamond X$  is well-posed on  $\mathcal{G}$  and satisfies (4.1.30). Finally, notice that  $\mathcal{G}$  is dense in  $L_X^2$ . Therefore, our goal is to extend the integral from  $\mathcal{G}$  to  $L_X^2$  by completeness; however this is not straightforward since  $(L_X^2, \|\cdot\|_\nu)$  is not actually a normed space (because  $\|\cdot\|_\nu$  is not a genuine norm, recall Remark 4.1.6). We circumvent this difficulty with a quotient of vector spaces.

Indeed, define (with an abuse of notation)

$$\begin{aligned} \text{Ker}(X) &:= \{g \in L_X^2; \|g\|_\nu = 0\} \\ &= \{g \in L_X^2; \forall h \in L_X^2, \langle g, h \rangle_\nu = 0\}. \end{aligned} \quad (4.3.9)$$

The equality in (4.3.9) is a direct consequence of the Cauchy-Schwarz inequality: for  $g \in L_X^2$  such that  $\|g\|_\nu = 0$  and  $h \in L_X^2$ , one has  $\langle g, h \rangle_\nu \leq \|g\|_\nu \|h\|_\nu = 0$  and  $-\langle g, h \rangle_\nu \leq \|g\|_\nu \|h\|_\nu = 0$ , so  $\langle g, h \rangle_\nu = 0$ . In particular,  $\text{Ker}(X)$  is a linear subspace of  $L_X^2$ .

For  $g \in L_X^2$ , let us denote  $\bar{g}$  its equivalence class in  $L_X^2/\text{Ker}(X)$ . It is clear that for  $\bar{g}, \bar{h} \in L_X^2/\text{Ker}(X)$  and any representatives  $g_1, g_2 \in \bar{g}$  and  $h_1, h_2 \in \bar{h}$ , then  $\langle g_1, h_1 \rangle_\nu = \langle g_2, h_2 \rangle_\nu$ . Therefore  $\langle \cdot, \cdot \rangle_\nu$  can be defined on  $L_X^2/\text{Ker}(X)$ , on which it is a scalar product; in particular  $\|\cdot\|_\nu$  is a well-defined norm on  $L_X^2/\text{Ker}(X)$ .

Recall that the integral is well-posed on  $\mathcal{G}$  and satisfies (4.1.30). For any  $g \in \mathcal{G}$ , we may define  $\bar{g} \diamond X := g \diamond X$ , which is well-defined almost everywhere on  $(\Omega, \mathcal{F}, \mathbb{P})$ : indeed, for any two representatives  $g_1, g_2 \in \bar{g}$ ,

$$\|g_1 \diamond X - g_2 \diamond X\|_{L^2}^2 = \|g_1 - g_2\|_\nu^2 = 0,$$

so  $g_1 \diamond X = g_2 \diamond X$  a.s.. Therefore, the application  $\bar{g} \mapsto \bar{g} \diamond X$  is an isometry from the normed space  $(\mathcal{G}/\text{Ker}(X), \|\cdot\|_\nu)$  to  $L^2(\mathbb{P})$ , hence it can be extended to the completion of  $(\mathcal{G}/\text{Ker}, \|\cdot\|_\nu)$  which is  $(L^2_X/\text{Ker}(X), \|\cdot\|_\nu)$ . Finally for  $g \in L^2_X$ , define  $g \diamond X := \bar{g} \diamond X$ , which satisfies (4.1.30) and thus concludes the proof of Theorem 4.1.7.  $\square$

**Remark 4.3.4.** Let us stress that Theorem 4.1.7 holds for any  $L^2(\mathbb{P})$ -random field  $X$  on  $\mathcal{S}_t$ , and can actually be extended to any dimension other than 2 straightforwardly. However, the assumption that the covariance function  $\nu$  defines a measure on  $\text{Bor}([0, t]^2)$  is not at all trivial. On the one hand, the field  $X$  cannot have negative correlations (because of (4.1.27)). This is not a very restrictive assumption actually, at least in a statistical mechanics setting, since many disordered systems display non-negative correlations, hence so do discrete fields in a polynomial expansion of the partition function ( $M_n$  in our setting) and their limits ( $\mathcal{M}$  here). Moreover, we expect that a similar construction holds for signed measures.

On the other hand, the covariance function  $\nu$  must be sigma-additive in order to be a measure on  $\text{Bor}([0, t]^2)$ , which is an important restriction. For instance, in dimension 1 this does not hold for the non-random, càdlàg function  $X(s) = 1_{[1/2, 1)}(s)$ ,  $s \in [0, 1)$  and its generated field, since

$$\sum_{n \geq 1} \left( X(1/2 - 2^{-n-1}) - X(1/2 - 2^{-n}) \right) = 0 \neq 1 = X(1/2) - X(0) .$$

Continuity of a random function  $X$  does not seem to be a sufficient condition either. In our setting and for the white noise, we could rewrite the covariance functions in terms of  $d$ -dimensional Lebesgue measure of some sets (recall (4.3.4) and (4.1.25), respectively for  $\lambda_3$  and  $\lambda_2$ ); which allowed us to circumvent this difficulty (since  $\lambda_d$  is a genuine measure, hence sigma-additive). A very interesting issue would be determining sufficient conditions on a random function  $X$  for the covariance function of its increments to be sigma-additive, and effectively define a measure on  $\text{Bor}([0, t]^2)$ .

### 4.3.3 Integrability of $\varphi$ against $\mathcal{M}$ : proof of Proposition 4.1.9

Theorem 4.1.7 defines the *integrals*  $g \diamond \mathcal{M}$ , of functions  $g \in L^2_{\mathcal{M}}$ , where we recall

$$L^2_{\mathcal{M}} := \left\{ g : [0, t] \rightarrow \mathbb{R} \text{ measurable ; } \int_{[0, t]^2} |g(\mathbf{u})g(\mathbf{v})| d\nu_{\mathcal{M}}(\mathbf{u}, \mathbf{v}) < \infty \right\} ,$$

and the measure  $\nu_{\mathcal{M}}$  is given explicitly in (4.1.24).

Let  $g : \mathbb{R}_+^2 \setminus \{0\} \rightarrow \mathbb{R}_+$  be the function defined by  $g(\mathbf{s}) = (\|\mathbf{s}\|_1)^{\alpha-2}$ : recalling (4.1.24), we have that

$$\begin{aligned} \|g\|_\nu^2 &= \int_{[0, t]} g(\mathbf{s}) \left( \int_0^{t_1} g(x, s_2) dx + \int_0^{t_2} g(s_1, y) dy \right) ds \\ &\leq 2C \int_{[0, t]} \|\mathbf{s}\|^{2\alpha-3} ds , \end{aligned} \tag{4.3.10}$$

where we used that  $g(x, s_2) = |x + s_2|^{\alpha-2}$  and that  $\alpha - 2 < -1$ , so there exists a constant  $C$  such that  $\int_0^{t_1} g(x, s_2) dx \leq C|s_2|^{\alpha-1} \leq C\|\mathbf{s}\|_1^{\alpha-1}$ . Then, since  $\alpha > 1/2$ ,

we have that  $2\alpha - 3 > -2$ , and  $\int_{[0,t]} \|s\|^{2\alpha-3} ds < +\infty$ , which concludes the proof of Proposition 4.1.9.  $\square$

Let us now give a corollary of Proposition 4.1.9. Recall the definition of  $\varphi_n$  from (4.1.20), that is

$$\varphi_n(\mathbf{s}) := n^{2-\alpha} L(n) u(\lfloor n\mathbf{s} \rfloor),$$

where  $u(\mathbf{i}) = \mathbf{P}(\mathbf{i} \in \tau)$  is the renewal function. Notice that  $\varphi_n$  is a piecewise constant function, and recall that Williamson [154] (see Theorem 4.1.1) shows that for  $\mathbf{s} \in [0, \mathbf{t}]$ ,  $\varphi_n(\mathbf{s})$  converges to  $\varphi(\mathbf{s})$  as  $n \rightarrow \infty$ .

**Corollary 4.3.5.** *The integrals of  $\varphi_n$  ( $n \geq 1$ ) and  $\varphi$  against  $\mathcal{M}$  are well-defined, that is  $\varphi_n, \varphi \in L^2_{\mathcal{M}}$ . Moreover,  $\|\varphi_n - \varphi\|_{\nu_{\mathcal{M}}} \rightarrow 0$  as  $n \rightarrow \infty$ , and  $\varphi_n \diamond \mathcal{M}$  converges to  $\varphi \diamond \mathcal{M}$  in  $L^2(\mathbb{P})$ .*

*Proof.* Let us recall some uniform bounds on  $u(\mathbf{i}) := \mathbf{P}(\mathbf{i} \in \tau)$ , that can be found in [17] (more precisely they improve Williamson's result [154] when close to the horizontal or vertical axis): for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon$  such that, uniformly in  $m \geq n$

$$\mathbf{P}\left((n, m) \in \tau\right) \leq C_\varepsilon n^{\alpha-2} L(n)^{-1} \left(\frac{m}{n}\right)^{-(1+\alpha)+\varepsilon}. \quad (4.3.11)$$

In particular, by symmetry, we get that  $\varphi_n(\mathbf{s}) \leq C \|\mathbf{s}^{[n]}\|_1^{\alpha-2} = Cg(\mathbf{s})$  uniformly in  $n \in \mathbb{N}$ ,  $\mathbf{s} \in [0, \mathbf{t}]$  (and this bound also holds for  $\varphi$ ).

Therefore, Proposition 4.1.9 shows that  $\varphi_n, \varphi \in L^2_{\mathcal{M}}$  for  $n \in \mathbb{N}$ . Since  $\varphi_n(\mathbf{s}) \rightarrow \varphi(\mathbf{s})$  for any  $\mathbf{s} \in [0, \mathbf{t}]$  as  $n \rightarrow \infty$ , by dominated convergence we get that

$$\|\varphi_n \diamond \mathcal{M} - \varphi \diamond \mathcal{M}\|_{L^2(\mathbb{P})} = \|\varphi_n - \varphi\|_{\nu_{\mathcal{M}}} \xrightarrow{n \rightarrow \infty} 0,$$

which proves the corollary.  $\square$

#### 4.3.4 Integrals of higher rank

In this section we define a multidimensional stochastic integral with respect to  $\mathcal{M}$ , which will be used to express the limits of terms of higher rank from the polynomial expansion of the partition function, i.e. for  $k \geq 1$ ,

$$\sum_{\mathbf{0} = \mathbf{i}_0 < \mathbf{i}_1 < \dots < \mathbf{i}_k \leq n\mathbf{t}} \prod_{l=1}^k \left( \zeta_{\mathbf{i}_l} u(\mathbf{i}_l - \mathbf{i}_{l-1}) \right).$$

Recall that  $\mathcal{S}_t$  denotes the semi-ring of sub-rectangles of  $[0, \mathbf{t}] \subset (\mathbb{R}_+)^2$ , and let  $\mathcal{S}_t^k := \mathcal{S}_t \times \dots \times \mathcal{S}_t$  be the set of sub-rectangles of  $[0, \mathbf{t}]^k \subset (\mathbb{R}_+)^{2k}$ , which is also a semi-ring. For  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$  a random field, define

$$X^{\otimes k}(A) := \prod_{l=1}^k X(A_l), \quad \text{for } A = A_1 \times \dots \times A_k \in \mathcal{S}_t^k,$$

which is a  $L^2(\mathbb{P})$ -random field on  $\mathcal{S}_t^k$  and may be called a *product* field. If  $X$  is a random function, i.e.  $X : [\mathbf{0}, \mathbf{t}) \rightarrow L^2(\mathbb{P})$ , then we may define  $X^{\otimes k}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \prod_{l=1}^k X(\mathbf{s}_l)$  a random function on  $[\mathbf{0}, \mathbf{t})^k$ , and the above definition matches exactly the  $2k$ -dimensional increment of  $X^{\otimes k}$  on  $A = \prod_{i=1}^{2k} [u_0^i, u_1^i) \subset \mathcal{S}_t^k$ , that is

$$X^{\otimes k}(A) = \sum_{\varepsilon_1, \dots, \varepsilon_{2k} \in \{0,1\}} (-1)^{(2k - \sum_{i=1}^{2k} \varepsilon_i)} X(u_{\varepsilon_1}^1, \dots, u_{\varepsilon_{2k}}^{2k}).$$

With this notation, we say that the field  $X$  admits a *covariance measure of rank  $k$*  if there exists a measure  $\nu_{X^{\otimes k}}$  on  $\text{Bor}([\mathbf{0}, \mathbf{t})^{2k})$  such that, for  $A, B \in \mathcal{S}_t^k$ ,

$$\nu_{X^{\otimes k}}(A, B) = \nu_{X^{\otimes k}}(A \times B) := \mathbb{E}[X^{\otimes k}(A) X^{\otimes k}(B)].$$

We state the following.

**Proposition 4.3.6.** *Let  $\mathcal{M}$  be a Gaussian field on  $(\mathbb{R}_+)^2$  with zero-mean and covariance matrix  $K(\mathbf{u}, \mathbf{v})$  given in (4.1.10), and let  $k \geq 1$ . Then  $\mathcal{M}$  admits a (finite, non-negative) covariance measure of rank  $k$  on  $\text{Bor}([\mathbf{0}, \mathbf{t})^{2k})$ , denoted  $\nu_{\mathcal{M}^{\otimes k}}$ , which is unique. Moreover, for any measurable function  $g : [\mathbf{0}, \mathbf{t})^{2k} \rightarrow \mathbb{R}$ , non-negative or bounded, we have*

$$\begin{aligned} & \int_{[\mathbf{0}, \mathbf{t})^{2k}} g(\mathbf{u}_1, \dots, \mathbf{u}_{2k}) d\nu_{\mathcal{M}^{\otimes k}}(\mathbf{u}_1, \dots, \mathbf{u}_{2k}) \\ &= \sum_{\mathcal{J}} \int_{[\mathbf{0}, \mathbf{t})^k} \left( \int_{\mathcal{A}_{u_{i_1}} \times \dots \times \mathcal{A}_{u_{i_k}}} g(\mathbf{u}_1, \dots, \mathbf{u}_{2k}) d\lambda_{u_{i_1}}(\mathbf{u}_{j_1}) \dots d\lambda_{u_{i_k}}(\mathbf{u}_{j_k}) \right) d\mathbf{u}_{i_1} \dots d\mathbf{u}_{i_k}, \end{aligned} \quad (4.3.12)$$

where the sum is over all partitions of  $\{1, \dots, 2k\}$  into pairs  $\mathcal{J} = \{\{i_1, j_1\}, \dots, \{i_k, j_k\}\}$ ; for  $\mathbf{u} \in [\mathbf{0}, \mathbf{t})$ ,  $\mathcal{A}_{\mathbf{u}}$  denotes the set of points in  $[\mathbf{0}, \mathbf{t})$  aligned with  $\mathbf{u}$ , i.e.

$$\mathcal{A}_{\mathbf{u}} := \{(x, u_2), x \in [0, t_1)\} \cup \{(u_1, y), y \in [0, t_2)\},$$

and finally  $\lambda_{\mathbf{u}}$  denotes the (one-dimensional) Lebesgue measure on  $\mathcal{A}_{\mathbf{u}}$ ,  $\mathbf{u} \in [\mathbf{0}, \mathbf{t})$ , i.e. for  $f : [\mathbf{0}, \mathbf{t}) \rightarrow \mathbb{R}$ ,  $\int_{\mathcal{A}_{\mathbf{u}}} f(\mathbf{v}) d\lambda_{\mathbf{u}}(\mathbf{v}) = \int_0^{t_1} f(x, u_2) dx + \int_0^{t_2} f(u_1, y) dy$ .

This claim is a direct analogue of Proposition 4.1.3 for generic  $k \geq 1$ . Formula (4.3.12) can be obtained for simple functions  $g = 1_A$ ,  $A \in \mathcal{S}_t^{2k}$  with Proposition 4.2.3, more precisely with the convergence of moments of  $\overline{M}_n$ ,  $n \geq 1$  to those of  $\mathcal{M}$ , and then be extended to all functions. Details of the proof are left to the reader.

Moreover, even though we stated and proved Theorem 4.1.7 in a 2-dimensional setting, we claim that it holds for random fields on  $d$ -dimensional rectangles for any  $d \geq 1$ , with a direct adaptation of the statement and demonstration. In particular for  $d = 2k$ , we may construct a stochastic integral against any product random field  $X^{\otimes k}$  for some  $X : \mathcal{S}_t \rightarrow L^2(\mathbb{P})$ , which admits a covariance measure of rank  $k$ ,  $\nu_{X^{\otimes k}}$ . Let

$$L_{X^{\otimes k}}^2 := \left\{ f : [\mathbf{0}, \mathbf{t})^k \rightarrow \mathbb{R} \text{ measurable ; } \int_{[\mathbf{0}, \mathbf{t})^{2k}} |f(\mathbf{u}) f(\mathbf{v})| d\nu_{X^{\otimes k}}(\mathbf{u}, \mathbf{v}) < \infty \right\}, \quad (4.3.13)$$

which is a vector space and contains  $L^\infty([\mathbf{0}, \mathbf{t}]^k)$ . For  $A \in \mathcal{S}_t^k$ , define

$$1_A \diamond^k X := 1_A \diamond X^{\otimes k} = X^{\otimes k}(A) \in L^2(\mathbb{P}).$$

Then the application  $g \mapsto g \diamond^k X$  can be extended linearly to  $L^2_{X^{\otimes k}}$ , such that for  $g, h \in L^2_{X^{\otimes k}}$ ,

$$\mathbb{E} \left[ \left( g \diamond^k X \right) \left( h \diamond^k X \right) \right] = \int_{[\mathbf{0}, \mathbf{t}]^{2k}} g(\mathbf{u}) h(\mathbf{v}) d\nu_{X^{\otimes k}}(\mathbf{u}, \mathbf{v}) =: \langle g, h \rangle_{\nu_{X^{\otimes k}}}. \quad (4.3.14)$$

This random variable  $g \diamond^k X \in L^2(\mathbb{P})$  will be called the *integral of rank  $k$*  of  $g \in L^2_{X^{\otimes k}}$  against  $X$ . We will also denote  $\|g\|_{\nu_{X^{\otimes k}}} := \|g \diamond^k X\|_{L^2(\mathbb{P})}$ .

### 4.3.5 Integrability of $\psi_t^{\text{free}}$ against $\mathcal{M}^{\otimes k}$

For any  $k \in \mathbb{N}$ , we define  $\psi_{t,k}^{\text{free}} := \psi_t^{\text{free}}$  as in (4.1.35):  $\psi_{t,k}^{\text{free}}$  is symmetric, for  $\mathbf{0} \prec \mathbf{s}_1 \prec \dots \prec \mathbf{s}_k \prec \mathbf{t}$

$$\psi_{t,k}^{\text{free}}(\mathbf{s}_1, \dots, \mathbf{s}_k) := \varphi(\mathbf{s}_1) \varphi(\mathbf{s}_2 - \mathbf{s}_1) \dots \varphi(\mathbf{s}_k - \mathbf{s}_{k-1}),$$

and  $\psi_{t,k}^{\text{free}}(\mathbf{s}_1, \dots, \mathbf{s}_k) = 0$  if  $\mathbf{s}_1, \dots, \mathbf{s}_k$  cannot be ordered. As in section 4.3.3, we prove here that the integral of  $\psi_{t,k}^{\text{free}}$  against  $\mathcal{M}^{\otimes k}$  is well defined when  $\alpha \in (\frac{1}{2}, 1)$ , and give a bound on its dependence on  $k$ . In this version of this work we do not discuss  $\psi_t^{\text{q}}$  and  $\psi_t^{\text{cond}}$ , but we expect that a very similar claim holds with minor changes in the proof.

**Proposition 4.3.7.** *If  $\alpha \in (\frac{1}{2}, 1)$ , then  $\psi_{t,k}^{\text{free}} \in L^2_{\mathcal{M}^{\otimes k}}$  for all  $k \geq 1$ . More precisely, there is a constant  $C_0 > 0$  such that for  $k \in \mathbb{N}$ , we have  $\|\psi_{t,k}^{\text{free}}\|_{\nu^{(k)}}^2 \leq (C_0 \|\mathbf{t}\|_1)^k k!$ . In particular  $\psi_{t,k}^{\text{free}} \diamond^k \mathcal{M}$  is a well-defined  $L^2(\mathbb{P})$ -random variable.*

*Proof.* Let us define  $g^{(k)}$  similarly to  $\psi_{t,k}^{\text{free}}$ , but with  $g(\mathbf{s}) = \|\mathbf{s}\|^{\alpha-2}$  in place of  $\varphi(\mathbf{s})$ , where  $\|\cdot\| := \|\cdot\|_1$  (notice the analogy with the case  $k = 1$ , more precisely Proposition 4.1.9). We show the proposition for  $g^{(k)}$ , which will imply the result for  $\psi_{t,k}^{\text{free}}$ . Let us warn the reader that the proof is considerably more technical than in the case  $k = 1$ , due to the richer combinatorial structure in the correlation structure of  $\nu_{\mathcal{M}^{\otimes k}}$ , see (4.3.12).

Reordering  $(\mathbf{u}_1, \dots, \mathbf{u}_k)$  when possible, and using that  $g^{(k)}(\mathbf{u}_1, \dots, \mathbf{u}_k)$  equals zero otherwise, we get

$$\begin{aligned} & \|g^{(k)}\|_{\nu_{\mathcal{M}^{\otimes k}}}^2 \\ &= (k!)^2 \int \dots \int_{\substack{\mathbf{0} \prec \mathbf{u}_1 \prec \dots \prec \mathbf{u}_k \prec \mathbf{t} \\ \mathbf{0} \prec \mathbf{v}_1 \prec \dots \prec \mathbf{v}_k \prec \mathbf{t}}} g^{(k)}(\mathbf{u}_1, \dots, \mathbf{u}_k) g^{(k)}(\mathbf{v}_1, \dots, \mathbf{v}_k) d\nu_{\mathcal{M}^{\otimes k}}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_k). \end{aligned} \quad (4.3.15)$$

Now, in view of the expression of  $\nu_{\mathcal{M}^{\otimes k}}$  (recall (4.3.12)), for any fixed  $\mathbf{0} \prec \mathbf{u}_1 \prec \dots \prec \mathbf{u}_k \prec \mathbf{t}$ , the integral over  $\mathbf{v}_1, \dots, \mathbf{v}_k$  is concentrated on the “grid” set

$$\mathcal{G}(\mathbf{u}_1, \dots, \mathbf{u}_k) = \bigcup_{i=1}^k \mathcal{A}_{\mathbf{u}_i},$$

where for  $\mathbf{u} \in [0, \mathbf{t}]$ ,  $\mathcal{A}_{\mathbf{u}}$  is the set of points aligned with  $\mathbf{u}$ , that we write as  $\mathcal{A}_{\mathbf{u}} = \mathcal{L}_{\mathbf{u}}^{(1)} \cup \mathcal{L}_{\mathbf{u}}^{(2)}$  with  $\mathcal{L}_{\mathbf{u}}^{(1)} = \{(x, u_2), x \in [0, t_1]\}$  and  $\mathcal{L}_{\mathbf{u}}^{(2)} = \{(u_1, y), y \in [0, t_2]\}$ . We also let  $\lambda_{(\mathbf{u}_1, \dots, \mathbf{u}_k)}$  be the Lebesgue measure on  $\mathcal{G}(\mathbf{u}_1, \dots, \mathbf{u}_k)$ , i.e. the product of the Lebesgue measures on  $\mathcal{L}_{\mathbf{u}_i}^{(a)}$ ,  $1 \leq i \leq k$ ,  $a = 1, 2$ .

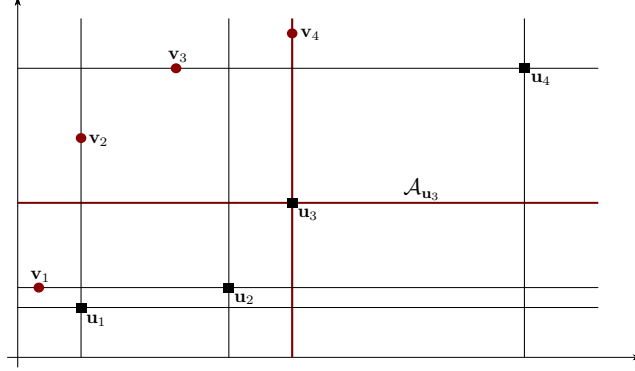


Figure 4.1: Illustration of the grid set  $\mathcal{G}(\mathbf{u}_1, \dots, \mathbf{u}_k)$ , with  $k = 4$  points. The black squares represent the positions of  $\mathbf{u}_1, \dots, \mathbf{u}_4$ , and the set  $\mathcal{A}_{\mathbf{u}_3}$  is represented in a red solid line. The red dots represent the positions of  $\mathbf{v}_1, \dots, \mathbf{v}_4$ : here we have that  $\mathbf{v}_1 \in \mathcal{A}_{\mathbf{u}_2}$  ( $\sigma(1) = 2$ ),  $\mathbf{v}_2 \in \mathcal{A}_{\mathbf{u}_1}$  ( $\sigma(2) = 1$ ),  $\mathbf{v}_3 \in \mathcal{A}_{\mathbf{u}_4}$  ( $\sigma(3) = 4$ ) and  $\mathbf{v}_4 \in \mathcal{A}_{\mathbf{u}_3}$  ( $\sigma(4) = 3$ ).

Moreover, the integral (4.3.15) is concentrated on  $(\mathbf{v}_1, \dots, \mathbf{v}_k) \in \mathcal{G}(\mathbf{u}_1, \dots, \mathbf{u}_k)$  where for any  $1 \leq j \leq k$ , there must be some  $\mathbf{v}_i$  in  $\mathcal{A}_{\mathbf{u}_j}$  (so that all points  $\mathbf{u}_j$  are aligned with one  $\mathbf{v}_i$ ): since  $\mathbf{v}_1 \prec \dots \prec \mathbf{v}_k$ , there is a permutation  $\sigma$  of  $\{1, \dots, k\}$  such that  $\mathbf{v}_i \in \mathcal{A}_{\mathbf{u}_{\sigma(i)}}$  for all  $i$  (using also that the Lebesgue measure of points in  $\mathcal{A}_{\mathbf{u}_j} \cap \mathcal{A}_{\mathbf{v}_i}$  is equal to 0). Then, we prove that there is a constant  $C$  such that, for any fixed  $\mathbf{0} \prec \mathbf{u}_1 \prec \dots \prec \mathbf{u}_k \prec \mathbf{t}$  (with the convention  $\mathbf{u}_0 = 0$ ),

$$\int_{\mathbf{0} \prec \mathbf{v}_1 \prec \dots \prec \mathbf{v}_k \prec \mathbf{t}} \dots \int g^{(k)}(\mathbf{v}_1, \dots, \mathbf{v}_k) \mathbf{1}_{\{\exists \sigma, \mathbf{v}_i \in \mathcal{A}_{\mathbf{u}_{\sigma(i)}} \forall i\}} d\lambda_{(\mathbf{u}_1, \dots, \mathbf{u}_k)}(\mathbf{v}_1, \dots, \mathbf{v}_k) \leq C^k \prod_{i=1}^k \|\mathbf{u}_i - \mathbf{u}_{i-1}\|^{\alpha-1}. \quad (4.3.16)$$

With this at hand, we are able to conclude that

$$\begin{aligned} \|g^{(k)}\|_{\nu_{\mathcal{M}^{\otimes k}}}^2 &\leq C^k (k!)^2 \int_{\mathbf{0} \prec \mathbf{u}_1 \prec \dots \prec \mathbf{u}_k \prec \mathbf{t}} \left( \prod_{i=1}^k \|\mathbf{u}_i - \mathbf{u}_{i-1}\|^{2\alpha-3} \right) ds_1 \dots d\mathbf{u}_k \\ &\leq C^k (k!)^2 \times \frac{C'^k \|\mathbf{t}\|^k}{k!} = (C'' \|\mathbf{t}\|)^k k!. \end{aligned}$$

It therefore remains to prove (4.3.16). We break down the integral on  $\mathcal{G}(\mathbf{u}_1, \dots, \mathbf{u}_k)$  into integrals on  $\mathcal{L}_{\mathbf{u}_1}^{(a_1)} \times \dots \times \mathcal{L}_{\mathbf{u}_k}^{(a_k)}$ , for all possible  $(a_1, \dots, a_k) \in \{1, 2\}^k$ . Recall that  $g^{(k)}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \prod_{i=1}^k \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2}$  if  $\mathbf{0} =: \mathbf{v}_0 \prec \mathbf{v}_1 \prec \dots \prec \mathbf{v}_k \prec \mathbf{t}$ . We prove that

whatever the permutation  $\sigma$  of  $\{1, \dots, k\}$  is, and for any  $(a_1, \dots, a_k) \in \{1, 2\}^k$ , the integral on the set  $\mathbf{v}_i \in \mathcal{L}_i^\sigma := \mathcal{L}_{\mathbf{u}_{\sigma(i)}}^{(a_{\sigma(i)})}$  for all  $1 \leq i \leq k$  verifies

$$\int_{\mathcal{L}_1^\sigma \times \dots \times \mathcal{L}_k^\sigma} \mathbf{1}_{\{\mathbf{0} \prec \mathbf{v}_1 \prec \dots \prec \mathbf{v}_k \prec \mathbf{t}\}} \prod_{i=1}^k \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2} d\lambda_{\sigma(i)}^{a_{\sigma(i)}}(\mathbf{v}_i) \leq C'^k \prod_{i=1}^k \|\mathbf{u}_{\sigma(i)} - \mathbf{u}_{\sigma(i)-1}\|^{\alpha-1}, \quad (4.3.17)$$

where  $\lambda_j^{a_j}$  is the Lebesgue measure on  $\mathcal{L}_{\mathbf{u}_j}^{(a_j)}$ . This gives (4.3.16), with  $C^k = 2^k C'^k$ .

We now prove (4.3.17) by iteration. Let us first give two useful estimates.

Analogously to (4.3.10), we easily establish that there is a constant  $C$  such that, for any fixed  $\mathbf{u}$  and  $\mathbf{v} \in [\mathbf{0}, \mathbf{t}] \setminus \mathcal{L}_{\mathbf{u}}^{(a)}$ , integrating over  $\mathcal{L}_{\mathbf{u}}^{(a)}$  gives

$$\int_{\mathcal{L}_{\mathbf{u}}^{(a)}} \|\mathbf{v}' - \mathbf{v}\|^{\alpha-2} d\lambda(\mathbf{v}') \leq C d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}}^{(a)})^{\alpha-1}, \quad (4.3.18)$$

where  $d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}}^{(a)})$  is the  $L^1$  distance between  $\mathbf{v}$  and the set  $\mathcal{L}_{\mathbf{u}}^{(a)}$ . This proves (4.3.17) for  $k = 1$ , by taking  $\mathbf{v} = \mathbf{0}$ .

We also have that for any fixed  $\mathbf{u} \neq \mathbf{u}'$  and  $a \neq a'$  (i.e.  $\mathcal{L}_{\mathbf{u}}^{(a)}$ ,  $\mathcal{L}_{\mathbf{u}'}^{(a')}$  are orthogonal), and any fixed  $\mathbf{v} \in [\mathbf{0}, \mathbf{t}] \setminus \mathcal{L}_{\mathbf{u}}^{(a)}$ , integrating over  $\mathcal{L}_{\mathbf{u}'}^{(a')}$  we have

$$\int_{\mathcal{L}_{\mathbf{u}'}^{(a')}} \|\mathbf{v}' - \mathbf{v}\|^{\alpha-2} d_1(\mathbf{v}', \mathcal{L}_{\mathbf{u}}^{(a)})^{\alpha-1} d\lambda(\mathbf{v}') \leq C' d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}}^{(a)})^{\alpha-1} d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}'}^{(a')})^{\alpha-1}. \quad (4.3.19)$$

Let us focus on the case  $a' = 1, a = 2$ , that is  $\mathcal{L}_{\mathbf{u}'}^{(a')}$  is horizontal and  $\mathcal{L}_{\mathbf{u}}^{(a)}$  vertical. To prove (4.3.19), we denote  $\mathbf{u}' = (u'_1, u'_2)$  and  $\mathbf{u} = (u_1, u_2)$ , and we let  $\mathbf{h} = (u_1, u'_2)$  be the point of intersection of  $\mathcal{L}_{\mathbf{u}'}^{(a')}$  and  $\mathcal{L}_{\mathbf{u}}^{(a)}$ . We now compute the integral over  $\mathcal{L}_{\mathbf{u}'}^{(a')} = \{(x, u'_2), x \in [0, t_1]\}$ :

$$\int_{\mathcal{L}_{\mathbf{u}'}^{(a')}} \|\mathbf{v}' - \mathbf{v}\|^{\alpha-2} \|\mathbf{v}' - \mathbf{h}\|^{\alpha-1} d\lambda(\mathbf{v}') = \int_0^{t_1} (|x - v_1| + |u'_2 - v_2|)^{\alpha-2} |x - u_1|^{\alpha-1} dx$$

After a change of variable, and setting  $b = |v_2 - u'_2|$ ,  $c = v_1 - u_1$ , we are reduced to estimating the integral (we may assume that  $c > 0$  by symmetry)

$$\int_{\mathbb{R}} (|y| + b)^{\alpha-2} |y - c|^{\alpha-1} dy = \int_0^{+\infty} (y + b)^{\alpha-2} |y - c|^{\alpha-1} dy + \int_0^{+\infty} (y + b)^{\alpha-2} (y + c)^{\alpha-1} dy.$$

The second term is bounded by a constant times  $c^{\alpha-1} b^{\alpha-1}$ , so we focus on the first term. For the first term, we split the sum into three parts:  $y \in [0, c/2)$ ,  $y \in [c/2, 2c)$  and  $y \in [2c, +\infty)$ . We get the upper bound

$$2^{1-\alpha} c^{\alpha-1} \int_0^{c/2} (y + b)^{\alpha-2} dy + (c/2 + b)^{\alpha-2} \int_{c/2}^{2c} |y - c|^{\alpha-1} dy + c^{\alpha-1} \int_{2c}^{+\infty} (y + b)^{\alpha-2} dy.$$



The first and last term are both bounded by a constant times  $c^{\alpha-1}b^{\alpha-1}$ . The second term is bounded by a constant times  $(c+b)^{\alpha-2}c^\alpha$  (recall  $\alpha-1 > -1$ ), which is itself bounded by  $c^{\alpha-1}b^{\alpha-1}$ . All together, we get that

$$\int_{\mathcal{L}_{\mathbf{u}'}^{(a')}} \|\mathbf{v}' - \mathbf{v}\|^{\alpha-2} \|\mathbf{v}' - \mathbf{h}\|^{\alpha-1} d\lambda(\mathbf{v}') \leq C|v_2 - u_2'|^{\alpha-1} |v_1 - u_1|^{\alpha-1}.$$

Using that  $|v_2 - u_2'| \geq d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}'}^{(a')})$  and  $|v_1 - u_1| \geq d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}}^{(a)})$ , we obtain that the integral is bounded by a constant times  $d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}'}^{(a')})^{\alpha-1} d_1(\mathbf{v}, \mathcal{L}_{\mathbf{u}}^{(a)})^{\alpha-1}$ .

We now go back to the proof of (4.3.17). For  $k \geq 2$ , we have thanks to (4.3.18) that for any fixed  $\mathbf{v}_{k-1} \in \mathcal{L}_{\sigma(k-1)}^{(a_{\sigma(k-1)})}$

$$\int_{\mathcal{L}_k^\sigma} \|\mathbf{v}_k - \mathbf{v}_{k-1}\|^{\alpha-2} d\lambda_{\sigma(k)}^{(a_{\sigma(k)})}(\mathbf{v}_k) \leq C d_1(\mathbf{v}_{k-1}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1}.$$

Then, we have two different cases.

If  $a_{\sigma(k)} = a_{\sigma(k-1)}$ , we have that  $\mathcal{L}_k^\sigma$  and  $\mathcal{L}_{k-1}^\sigma$  are parallel (either both vertical or both horizontal): in that case we have that  $d_1(\mathbf{v}_{k-1}, \mathcal{L}_k^\sigma) \geq c \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k-1)}\|$  uniformly for  $\mathbf{v}_{k-1} \in \mathcal{L}_{\sigma(k-1)}^{(a_{\sigma(k-1)})}$ . Moreover, we necessarily have that  $\mathbf{u}_{\sigma(k-1)} \prec \mathbf{u}_{\sigma(k)}$  (otherwise we could not have had  $\mathbf{v}_{k-1} \prec \mathbf{v}_k$ ), so  $\sigma(k) > \sigma(k-1)$ , and we can bound  $\|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k-1)}\| \geq \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k)-1}\|$ . We therefore get that the left-hand side of (4.3.17) is bounded by

$$C \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k)-1}\|^{\alpha-1} \int_{\mathcal{L}_1^\sigma \times \dots \times \mathcal{L}_{k-1}^\sigma} \dots \int 1_{\{0 \prec v_1 \prec \dots \prec v_{k-1} \prec t\}} \prod_{i=1}^{k-1} \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2} d\lambda_{\sigma(i)}^{(a_{\sigma(i)})}(\mathbf{v}_i)$$

and then the induction proceeds.

If  $a_{\sigma(k)} \neq a_{\sigma(k-1)}$ , we have that  $\mathcal{L}_k^\sigma$  and  $\mathcal{L}_{k-1}^\sigma$  are orthogonal. The left-hand side of (4.3.17) is bounded by

$$C \int_{\mathcal{L}_1^\sigma \times \dots \times \mathcal{L}_{k-1}^\sigma} \dots \int 1_{\{0 \prec v_1 \prec \dots \prec v_{k-1} \prec t\}} d_1(\mathbf{v}_{k-1}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1} \prod_{i=1}^{k-1} \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2} d\lambda_{\sigma(i)}^{(a_{\sigma(i)})}(\mathbf{v}_i). \quad (4.3.20)$$

Now, we can integrate over  $\mathcal{L}_{k-1}^\sigma$ : since  $a_{\sigma(k)} \neq a_{\sigma(k-1)}$ , we get thanks to (4.3.19) that for any fixed  $\mathbf{v}_{k-2} \in \mathcal{L}_{\sigma(k-2)}^{(a_{\sigma(k-2)})}$  (with the convention that  $\mathbf{v}_0 = \mathbf{0}$ ),

$$\begin{aligned} \int_{\mathcal{L}_{k-1}^\sigma} \|\mathbf{v}_{k-1} - \mathbf{v}_{k-2}\|^{\alpha-2} d_1(\mathbf{v}_{k-1}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1} d\lambda_{\sigma(k-1)}^{(a_{\sigma(k-1)})}(\mathbf{v}_{k-1}) \\ \leq C d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k-1)}}^{(a_{\sigma(k-1)})})^{\alpha-1} d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1}. \end{aligned} \quad (4.3.21)$$

In the case where  $k-2 = 0$ , this is bounded by a constant times  $\|\mathbf{u}_{\sigma(k-1)}\|^{\alpha-1} \|\mathbf{u}_{\sigma(k)}\|^{\alpha-1}$ , which is itself bounded by  $\|\mathbf{u}_{\sigma(k-1)} - \mathbf{u}_{\sigma(k-1)-1}\|^{\alpha-1} \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k)-1}\|^{\alpha-1}$ . In the case where  $k-2 \geq 1$ , there are again two possibilities.

Either  $a_{\sigma(k-1)} \neq a_{\sigma(k-2)}$ , then  $a_{\sigma(k-2)} = a_{\sigma(k)}$ , and we bound  $d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1}$  by a constant times  $\|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k-2)}\|^{\alpha-1}$ . This can happen only if  $\mathbf{u}_{\sigma(k-2)} \prec \mathbf{u}_{\sigma(k)}$  (otherwise we could not have had  $\mathbf{v}_{k-2} \prec \mathbf{v}_k$ ), and we therefore have that  $\sigma(k) > \sigma(k-2)$  so we bound  $\|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k-2)}\| \geq \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k)-1}\|$ . We then get that (4.3.20) is bounded by a constant times

$$\begin{aligned} & \|\mathbf{u}_{\sigma(k)} - \mathbf{u}_{\sigma(k)-1}\|^{\alpha-1} \\ & \times \int_{\mathcal{L}_1^\sigma \times \cdots \times \mathcal{L}_{k-2}^\sigma} \cdots \int \mathbf{1}_{\{\mathbf{0} \prec \mathbf{v}_1 \prec \cdots \prec \mathbf{v}_{k-2} \prec \mathbf{t}\}} d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k-1)}}^{(a_{\sigma(k-1)})})^{\alpha-1} \prod_{i=1}^{k-2} \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2} d\lambda_{\sigma(i)}^{a_{\sigma(i)}}(\mathbf{v}_i), \end{aligned}$$

and the iteration proceeds for bounding (4.3.20) (recall that we have  $a_{\sigma(k-2)} \neq a_{\sigma(k-1)}$ ).

If  $a_{\sigma(k-1)} = a_{\sigma(k-2)}$  (and  $a_{\sigma(k-2)} \neq a_{\sigma(k)}$ ), and we bound  $d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k-1)}}^{(ra_{\sigma(k-1)})})^{\alpha-1}$  by a constant times  $\|\mathbf{u}_{\sigma(k-1)} - \mathbf{u}_{\sigma(k-2)}\|^{\alpha-1}$ . Again, this can happen only if  $\mathbf{u}_{\sigma(k-2)} \prec \mathbf{u}_{\sigma(k-1)}$ , so  $\sigma(k-1) > \sigma(k-2)$ , and we can bound  $\|\mathbf{u}_{\sigma(k-1)} - \mathbf{u}_{\sigma(k-2)}\| \geq \|\mathbf{u}_{\sigma(k-1)} - \mathbf{u}_{\sigma(k-1)-1}\|$ . We then get that (4.3.20) is bounded by a constant times

$$\begin{aligned} & \|\mathbf{u}_{\sigma(k-1)} - \mathbf{u}_{\sigma(k-1)-1}\|^{\alpha-1} \\ & \times \int_{\mathcal{L}_1^\sigma \times \cdots \times \mathcal{L}_{k-2}^\sigma} \cdots \int \mathbf{1}_{\{\mathbf{0} \prec \mathbf{v}_1 \prec \cdots \prec \mathbf{v}_{k-2} \prec \mathbf{t}\}} d_1(\mathbf{v}_{k-2}, \mathcal{L}_{\mathbf{u}_{\sigma(k)}}^{(a_{\sigma(k)})})^{\alpha-1} \prod_{i=1}^{k-2} \|\mathbf{v}_i - \mathbf{v}_{i-1}\|^{\alpha-2} d\lambda_{\sigma(i)}^{a_{\sigma(i)}}(\mathbf{v}_i), \end{aligned}$$

and then we can proceed by iteration to obtain a bound on (4.3.20) (note that we have  $a_{\sigma(k-2)} \neq a_{\sigma(k)}$ ).  $\square$

## 4.4 Convergence of terms of order $k \geq 1$ to an integral

In this section, we prove that the first term in the polynomial expansion of the partition function (4.1.7) converge to an integral against  $\mathcal{M}$ . We also discuss terms of higher rank  $k > 1$ , and describe what work remains to prove Conjecture 4.1.10.

In this section we assume that  $\mathbb{P} \in \mathfrak{P}_r$  for some  $r \in \{1, 2\}$ ,  $\alpha \in (1/2, 1)$ , and  $\beta_n(n^{\alpha-\frac{1}{2}}L(n)^{-1})^{\frac{1}{r}} \rightarrow \hat{\beta}$  as  $n \rightarrow \infty$  for some  $\hat{\beta} \in [0, +\infty)$ .

### 4.4.1 Convergence for $k = 1$

We first prove the convergence for the term of order 1 in the polynomial expansion, i.e. (4.1.16), since it encapsulates the main ideas of the proof.

**Proposition 4.4.1.** *Under the assumptions of Conjecture 4.1.10, one has*

$$\frac{L(n)}{c_r(\beta_n)^r n^{\alpha-\frac{1}{2}}} \sum_{\mathbf{i} \in \llbracket 1, n \rrbracket} \zeta_{\mathbf{i}} u(\mathbf{i}) \xrightarrow[n \rightarrow \infty]{(d)} \varphi \diamond \mathcal{M}, \quad (4.4.1)$$

and the l.h.s. is bounded in  $L^2(\mathbb{P})$  uniformly in  $n \in \mathbb{N}$ . Moreover on a convenient probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ , this convergence holds in  $L^2(\hat{\mathbb{P}})$ .

Let us give some insight into the proof of this proposition: recalling the idea from (4.1.21), we rewrite the term from the polynomial expansion as some “integral”  $\int \varphi_n d\bar{M}_n$  (see (4.4.2-4.4.3) below). Then, we prove the following convergences for  $n \geq m$ :

$$\int \varphi_m d\bar{M}_n \xrightarrow{n \rightarrow \infty} \int \varphi_m d\mathcal{M} \xrightarrow{m \rightarrow \infty} \int \varphi d\mathcal{M},$$

and we show that  $\int \varphi_n d\bar{M}_n$  is close to  $\int \varphi_m d\bar{M}_n$  for large  $n \geq m$  (see Lemma 4.4.3 below).

**Remark 4.4.2.** Notice that it would be interesting to complete the diagram of convergences, by proving

$$\int \varphi_m d\bar{M}_n \xrightarrow{m \rightarrow \infty} \int \varphi d\bar{M}_n \xrightarrow{n \rightarrow \infty} \int \varphi d\mathcal{M}.$$

However those convergences are not at all trivial, and showing them would burden this paper even more. Moreover they are not needed to prove Conjecture 4.1.10, so we do not further discuss them here.

First we need to introduce some notation. For  $n \in \mathbb{N}$ , let  $D_n := \frac{1}{n}\mathbb{Z}^2 \cap [\mathbf{0}, \mathbf{t}]$ . Recall the definition of  $\bar{M}_n$ ,  $n \geq 1$  from (4.2.6), the definition of increments of a field from (4.1.18), and that  $\Delta_n := [\mathbf{0}, \frac{1}{n}\mathbf{1}]$ . For clarity’s sake, we will write  $\zeta_{i,n} := \zeta_i = e^{\beta_n \omega_i - \lambda(\beta_n)} - 1$  for  $i \in \mathbb{N}^2$ ,  $n \in \mathbb{N}$ .

Using Claim 4.2.6, let us assume that  $\bar{M}_n$ ,  $n \geq 1$  converges  $\mathbb{P}$ -a.s. to  $\bar{M}$ , and that all pointwise convergences hold in  $L^p(\mathbb{P})$  for  $p \geq 1$  (we do not change notation for simplicity’s sake). We will prove the second part of the Proposition under those assumptions, i.e. that the sum converges in  $L^2(\mathbb{P})$ ; then the first statement is a direct consequence.

For  $m \in \mathbb{N}$ , let  $g_m : [\mathbf{0}, \mathbf{t}] \rightarrow \mathbb{R}$  be a piecewise constant function on  $(\frac{1}{m}\mathbf{i} + \Delta_m)_{i \in \mathbb{N}_0^2}$ : i.e. for  $\mathbf{u} \in [\mathbf{0}, \mathbf{t}]$ ,  $g_m(\mathbf{u}) = g_m(\mathbf{u}^{[m]})$ . Then for  $n \in \mathbb{N}$ , let us define

$$\int_{[\mathbf{0}, \mathbf{t}]} g_m d\bar{M}_n := g_m \cdot \bar{M}_n := \sum_{\mathbf{u} \in D_m} g_m(\mathbf{u}) \bar{M}_n \left( \mathbf{u} - \frac{1}{n}\mathbf{1} + \Delta_m \right), \quad (4.4.2)$$

(for conciseness we will mainly use the second notation). The term  $(-\frac{1}{n}\mathbf{1})$  is added to ensure that, if  $n = m$  and  $\frac{1}{n}\mathbf{i} \in D_n = D_m$ , then we get  $\bar{M}_n(\frac{1}{n}(\mathbf{i} - \mathbf{1}) + \Delta_n) = \frac{1}{c_r n^{3/2} \beta_n^r} \zeta_{i,n}$ . Recalling the definition of  $\varphi_n$  from (4.1.20), this implies that for  $n \in \mathbb{N}$ , we can rewrite the sum

$$\frac{L(n)}{c_r(\beta_n)^r n^{\alpha - \frac{1}{2}}} \sum_{i \in \llbracket \mathbf{1}, n\mathbf{t} \rrbracket} \zeta_{i,n} u(\mathbf{i}) = \varphi_n \cdot \bar{M}_n. \quad (4.4.3)$$

*Proof.* We are finally ready to start the proof. Let  $\varepsilon > 0$ , and for  $n \geq m$ , write

$$\|\varphi_n \cdot \bar{M}_n - \varphi \diamond \mathcal{M}\|_{L^2} \leq I_1 + I_2 + I_3, \quad (4.4.4)$$

where

$$\begin{aligned} I_1 &:= \|\varphi_n \cdot \bar{M}_n - \varphi_m \cdot \bar{M}_n\|_{L^2}, \\ I_2 &:= \|\varphi_m \cdot \bar{M}_n - \varphi_m \diamond \mathcal{M}\|_{L^2}, \\ I_3 &:= \|\varphi_m \diamond \mathcal{M} - \varphi \diamond \mathcal{M}\|_{L^2}. \end{aligned} \quad (4.4.5)$$

Corollary 4.3.5 already implies that there exists  $m_0 \in \mathbb{N}$  such that  $I_3 \leq \varepsilon/3$  for any  $m \geq m_0$ . Moreover we claim the following lemma, which is proven afterwards.

**Lemma 4.4.3.** *Let  $\eta > 0$ . There exists  $m_1 \in \mathbb{N}$  such that for  $n \geq m \geq m_1$ ,*

$$\|\varphi_m \cdot \overline{M}_n - \varphi_n \cdot \overline{M}_n\|_{L^2} \leq \eta. \quad (4.4.6)$$

Applying Lemma 4.4.3 with  $\eta = \varepsilon/3$ , we set  $m = \max(m_0, m_1)$ , so that  $I_1 \leq \varepsilon/3$  uniformly in  $n \geq m$ . Let us prove that  $I_2 \rightarrow 0$  as  $n \rightarrow \infty$  (where  $m$  is fixed), which will conclude the proof.

Since  $\varphi_m$  is piecewise-constant on  $(\frac{1}{m}\mathbf{i} + \Delta_m)_{\mathbf{i} \in \mathbb{N}_0^2}$ , there exists a family  $a_{\mathbf{w}} \in \mathbb{R}$ ,  $\mathbf{w} \in D_m$  such that for  $\mathbf{u} \in [\mathbf{0}, \mathbf{t})$ ,

$$\varphi_m(\mathbf{u}) = \sum_{\mathbf{w} \in D_m} a_{\mathbf{w}} \mathbf{1}_{[\mathbf{0}, \mathbf{w})}(\mathbf{u}). \quad (4.4.7)$$

Thus, we may rewrite with an integration by parts,

$$\begin{aligned} \varphi_m \cdot \overline{M}_n &= \sum_{\mathbf{u} \in D_m} \left( \sum_{\mathbf{w} \in D_m} a_{\mathbf{w}} \mathbf{1}_{[\mathbf{0}, \mathbf{w})}(\mathbf{u}) \right) \overline{M}_n \left( \mathbf{u} - \frac{1}{n} \mathbf{1} + \Delta_m \right) \\ &= \sum_{\mathbf{w} \in D_m} a_{\mathbf{w}} \left( \sum_{\mathbf{u} \in D_m, \mathbf{u} \prec \mathbf{w}} \overline{M}_n \left( \mathbf{u} - \frac{1}{n} \mathbf{1} + \Delta_m \right) \right) = \sum_{\mathbf{w} \in D_m} a_{\mathbf{w}} \overline{M}_n \left( \mathbf{w} - \frac{1}{n} \mathbf{1} \right). \end{aligned} \quad (4.4.8)$$

Recall that, by definition of  $(\cdot \diamond \cdot)$ , we have  $\mathbf{1}_{[\mathbf{0}, \mathbf{w})} \diamond \mathcal{M} = \mathcal{M}([\mathbf{0}, \mathbf{w})) = \mathcal{M}(\mathbf{w})$ . Hence,

$$\varphi_m \diamond \mathcal{M} = \sum_{\mathbf{w} \in D_m} a_{\mathbf{w}} \mathcal{M}(\mathbf{w}). \quad (4.4.9)$$

For  $\mathbf{w} \in D_m$ , Claim 4.2.6 gives that  $\overline{M}_n(\mathbf{w}) \rightarrow \mathcal{M}(\mathbf{w})$  in  $L^2$  as  $n \rightarrow \infty$ . We conclude the proof with Lemma 4.A.1, which ensures that for  $\mathbf{w} \in D_m$ ,

$$\left\| \overline{M}_n(\mathbf{w}) - \overline{M}_n \left( \mathbf{w} - \frac{1}{n} \mathbf{1} \right) \right\|_{L^2} \leq \frac{C_1}{\sqrt{n}}, \quad (4.4.10)$$

for some uniform  $C_1 > 0$ ; thus  $\varphi_m \cdot \overline{M}_n \rightarrow \varphi_m \diamond \mathcal{M}$  in  $L^2(\mathbb{P})$  as  $n \rightarrow \infty$ .  $\square$

Let us now prove Lemma 4.4.3. To obtain a uniform control of  $\varphi_m \cdot \overline{M}_n$  in  $n \geq m$ , we use that the convergence  $\varphi_m(\mathbf{u}) \rightarrow \varphi(\mathbf{u})$ ,  $m \rightarrow \infty$ , holds uniformly in  $\mathbf{u} \in [\mathbf{0}, \mathbf{t}) \setminus [0, \delta)^2$  for  $\delta > 0$  (see [154] for the proof). However the function  $\varphi_m$  asymptotically blows up around  $\mathbf{0}$  (recall Theorem 4.1.1), so we handle this bit separately in the following claim, by proving that it gets compensated by the small covariance of the field. For  $K > 0$ ,  $m \in \mathbb{N}$ , define  $\varphi_m^K$  a truncation of  $\varphi_m$ ,

$$\varphi_m^K(\mathbf{u}) := \varphi_m(\mathbf{u}) \vee K, \quad \mathbf{u} \in [\mathbf{0}, \mathbf{t}). \quad (4.4.11)$$

**Claim 4.4.4.** *Let  $\varepsilon > 0$ . There exist  $K > 0$  and  $m_2 > 0$  such that for  $n \geq m \geq m_2$ , one has*

$$\|\varphi_m^K \cdot \overline{M}_n - \varphi_m \cdot \overline{M}_n\|_{L^2} \leq \varepsilon. \quad (4.4.12)$$

*Proof of Claim 4.4.4.* Let  $K > 0$ , and let us denote  $\varphi_m^{>K} := \varphi_m - \varphi_m^K \geq 0$ . A straightforward computation yields

$$\begin{aligned} & \|\varphi_m^K \cdot \bar{M}_n - \varphi_m \cdot \bar{M}_n\|_{L^2}^2 \\ &= \sum_{\mathbf{u}, \mathbf{v} \in D_m} \varphi_m^{>K}(\mathbf{u}) \varphi_m^{>K}(\mathbf{v}) \mathbb{E} \left[ \bar{M}_n(\mathbf{u} - \frac{1}{n} \mathbf{1} + \Delta_m) \bar{M}_n(\mathbf{v} - \frac{1}{n} \mathbf{1} + \Delta_m) \right] \end{aligned}$$

Recall from Corollary 4.3.5 that there exists  $c > 0$  such that  $\varphi_m(\frac{i}{m}, \frac{j}{m}) \leq c(\frac{i \vee j}{m})^{\alpha-2}$  uniformly in  $m \in \mathbb{N}$ , and  $1 \leq i, j \leq m$ . In particular for  $i, j \in \mathbb{N}$ ,  $\varphi_m^{>K}(\frac{i}{m}, \frac{j}{m}) > 0$  implies  $i, j \leq m\delta_K$  where  $\delta_K := (c/K)^{1/(2-\alpha)}$ . Using Lemma 4.A.2 (which is not altered by the translation  $-\frac{1}{n} \mathbf{1}$ ), we obtain

$$\begin{aligned} \|\varphi_m^K \cdot \bar{M}_n - \varphi_m \cdot \bar{M}_n\|_{L^2} &\leq 2C \sum_{i=1}^{m\delta_K} \sum_{j=1}^{m\delta_K} \sum_{j'=1}^{m\delta_K} \left(\frac{i \vee j}{m}\right)^{\alpha-2} \left(\frac{i \vee j'}{m}\right)^{\alpha-2} \frac{1}{m^3} \\ &\quad + C \sum_{i=1}^{m\delta_K} \sum_{j=1}^{m\delta_K} \left(\frac{i \vee j}{m}\right)^{2\alpha-4} \frac{1}{n\beta_n^{2(r-1)}m^2} \\ &\leq 2Cm^{1-2\alpha} \sum_{i=1}^{m\delta_K} \sum_{j=1}^{m\delta_K} \sum_{j'=1}^{m\delta_K} (i \vee j)^{\alpha-2} (i \vee j')^{\alpha-2} \\ &\quad + C \frac{m^{2-2\alpha}}{n\beta_n^{2(r-1)}} \sum_{i=1}^{m\delta_K} \sum_{j=1}^{m\delta_K} (i \vee j)^{2\alpha-4}. \end{aligned}$$

Then, notice that there exists  $C_2 > 0$  such that

$$\sum_{j=1}^{m\delta_K} (i \vee j)^{\alpha-2} \leq i^{\alpha-1} + \sum_{j=i+1}^{\infty} j^{\alpha-2} \leq C_2 i^{\alpha-1},$$

uniformly in  $i \in \mathbb{N}$ , and similarly  $\sum_{j=1}^{m\delta_K} (i \vee j)^{2\alpha-4} \leq C_3 i^{2\alpha-3}$ . Moreover,  $\sum_{i=1}^{m\delta_K} i^{2(\alpha-1)} \sim (m\delta_K)^{2\alpha-1}$  as  $m \rightarrow \infty$ , and  $\sum_{i \geq 1} i^{2\alpha-3} < \infty$ , thus

$$\|\varphi_m^K \cdot \bar{M}_n - \varphi_m \cdot \bar{M}_n\|_{L^2}^2 \leq C_4 \delta_K^{2\alpha-1} + C_5 \frac{m^{2-2\alpha}}{n\beta_n^{2(r-1)}},$$

uniformly in  $K > 0$  and  $n, m \in \mathbb{N}$ . Finally, set  $K > 0$  such that  $C_4 \delta_K^{2\alpha-1} < \varepsilon/2$  (since  $\delta_K \rightarrow 0$  as  $K \rightarrow \infty$ ), and notice that for  $n \geq m$ , the second term is dominated by  $C_5 m^{1-2\alpha}$  if  $r = 1$  and by  $C_6 n^{\frac{1}{2}-\alpha} L(n)^{-1}$  if  $r = 2$ ; recalling that  $\alpha > 1/2$ , we conclude the proof by setting  $m_2$  sufficiently large.  $\square$

*Proof of Lemma 4.4.3.* Applying Claim 4.4.4 with  $\varepsilon = \eta/3$ , the lemma will follow the fact that there exists  $m_1 \in \mathbb{N}$  such that for  $n \geq m \geq m_1$ ,

$$\|\varphi_m^K \cdot \bar{M}_n - \varphi_n^K \cdot \bar{M}_n\|_{L^2} \leq \eta/3, \quad (4.4.13)$$

for some fixed  $K > 0$ , where we recall that  $\varphi_m^K := \varphi_m \vee K$ .

First, notice that for  $n \geq m$  and  $\mathbf{u} \in D_p$ ,  $p \in \{n, m\}$ , we may rewrite

$$\overline{M}_n(\mathbf{u} - \frac{1}{n} + \Delta_p) = \sum_{\mathbf{w} \in (\mathbf{u} + \Delta_p) \cap D_{nm}} \overline{M}_n(\mathbf{w} - \frac{1}{n} + \Delta_{nm}),$$

which gives with (4.4.2),

$$\varphi_m^K \cdot \overline{M}_n - \varphi_n^K \cdot \overline{M}_n = \sum_{\mathbf{w} \in D_{nm}} (\varphi_m^K(\mathbf{w}) - \varphi_n^K(\mathbf{w})) \overline{M}_n(\mathbf{w} - \frac{1}{n} + \Delta_{nm}). \quad (4.4.14)$$

Let  $\eta' > 0$  (which we will fix below). Because of the truncation, we have  $\varphi_m^K \rightarrow \varphi^K$  uniformly on  $[\mathbf{0}, \mathbf{t}]$  as  $m \rightarrow \infty$ , hence there is  $m_3 > 0$  such that  $\|\varphi_m^K - \varphi^K\|_\infty \leq \eta'$  for  $m \geq m_3$ . Therefore, taking the squared  $L^2$ -norm of (4.4.14), we obtain for  $n \geq m \geq m_3$ ,

$$\begin{aligned} \|\varphi_m^K \cdot \overline{M}_n - \varphi_n^K \cdot \overline{M}_n\|_{L^2}^2 &\leq 4(\eta')^2 \sum_{\mathbf{w}, \mathbf{w}' \in D_{nm}} \mathbb{E} \left[ \overline{M}_n(\mathbf{w} - \frac{1}{n} + \Delta_{nm}) \overline{M}_n(\mathbf{w}' - \frac{1}{n} + \Delta_{nm}) \right] \\ &\leq 4(\eta')^2 \mathbb{E} \left[ \overline{M}_n([\mathbf{0}, \mathbf{t}])^2 \right] = 4(\eta')^2 \mathbb{E} \left[ \overline{M}_n(\mathbf{t})^2 \right]. \end{aligned}$$

Recall that  $(\mathbb{E}[\overline{M}_n(\mathbf{t})^2])_{n \geq 1}$  converges, so it is bounded. Thus, we conclude the proof of (4.4.13) by taking

$$\eta' = \frac{1}{6} \eta \left( \sup \left( \mathbb{E} \left[ \overline{M}_n(\mathbf{t})^2 \right], n \geq 1 \right) \right)^{-1}.$$

□

#### 4.4.2 Discussion: convergence for $k \geq 1$ and Conjecture 4.1.10

In this section we display the analogue of Proposition 4.4.1 for all rank  $k \geq 1$ , and discuss what work remains to obtain Conjecture 4.1.10.

Let  $k \geq 1$ , and let us first introduce some notation. For  $m, n \in \mathbb{N}$  and  $g_m : [\mathbf{0}, \mathbf{t}]^k \rightarrow \mathbb{R}$  a function constant on each  $\prod_{l=1}^k (\frac{1}{m} \mathbf{i}_l + \Delta_m)$ ,  $\mathbf{i}_1, \dots, \mathbf{i}_k \in \mathbb{N}_0^2$ , define similarly to (4.4.2),

$$g_m \cdot \overline{M}_n^k := \sum_{\mathbf{u}_1, \dots, \mathbf{u}_k \in D_m} g_m(\mathbf{u}_1, \dots, \mathbf{u}_k) \prod_{l=1}^k \overline{M}_n(\mathbf{u}_l - \frac{1}{n} \mathbf{1} + \Delta_m). \quad (4.4.15)$$

Recall the definitions of  $\varphi$  and  $\psi := \psi_t^{\text{free}}$  from Theorem 4.1.1 and (4.1.35), where we dropped the sub/superscripts of  $\psi_t^{\text{free}}$  for conciseness. Recall also the definitions of  $\varphi_m$  and  $\varphi_m^K$  from (4.1.20) and (4.4.11) respectively, and define  $\psi_m, \psi_m^K$  by replacing  $\varphi$  with the appropriate function in (4.1.35). With those notation, we may rewrite the  $k$ -th term from the polynomial expansion of the partition function (4.1.7),

$$\left( \frac{L(n)}{c_r(\beta_n)^r n^{\alpha - \frac{1}{2}}} \right)^k \sum_{\mathbf{0} = \mathbf{i}_0 < \mathbf{i}_1 < \dots < \mathbf{i}_k \leq n\mathbf{t}} \prod_{l=1}^k \zeta_{\mathbf{i}_l, n} u(\mathbf{i}_l - \mathbf{i}_{l-1}) = \frac{1}{k!} \psi_n \cdot \overline{M}_n^k, \quad (4.4.16)$$

where we dropped the ordering of indices in the r.h.s. at the cost of a factor  $1/k!$ . Then, we conjecture the following results.

**Proposition 4.4.5** (Work in progress). *Let  $k \geq 1$ . Under the assumptions of Conjecture 4.1.10, one has*

$$\psi_n \cdot^k \overline{M}_n \xrightarrow[n \rightarrow \infty]{(d)} \psi \diamond^k \mathcal{M}, \quad (4.4.17)$$

and the squared  $L^2(\mathbb{P})$ -norm of the l.h.s. is bounded by some  $(C_0 \|\mathbf{t}\|_1)^k k!$  uniformly in  $n, k \in \mathbb{N}$ . Moreover on a convenient probability space  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ , this convergence holds in  $L^2(\widehat{\mathbb{P}})$ .

**Claim 4.4.6** (Work in progress). *Let  $\eta > 0$ . There exists  $m_1 \in \mathbb{N}$  such that for  $n \geq m \geq m_1$ ,*

$$\|\psi_m \cdot^k \overline{M}_n - \psi_n \cdot^k \overline{M}_n\|_{L^2} \leq \eta. \quad (4.4.18)$$

**Claim 4.4.7** (Work in progress). *Let  $\varepsilon > 0$ . There exist  $K > 0$  and  $m_2 > 0$  such that for  $n \geq m \geq m_2$ , one has*

$$\|\psi_m^K \cdot^k \overline{M}_n - \psi_m \cdot^k \overline{M}_n\|_{L^2} \leq \varepsilon. \quad (4.4.19)$$

Those results are analogous to Proposition 4.4.1, Lemma 4.4.3 and Claim 4.4.4 respectively, and are obtained similarly. At the current state of this work we have partial proofs of these results, and are working on the last technical estimates required to conclude (notably an adaptation of lemmas from Appendix 4.A).

Once Proposition 4.4.1 is proven, we straightforwardly obtain Conjecture 4.1.10 in the case  $h_n = \widehat{h} = 0$ ,  $n \geq 1$  for the *free* partition function: Proposition 4.3.7 already proves that the series in (4.1.33) is well-defined in  $L^2(\mathbb{P})$ , and Proposition 4.4.1 gives the convergence of any (finite) number of terms from the polynomial expansion of  $Z_{[nt], h_n}^{\beta_n, \omega, \text{free}}$  to terms of the series.

Regarding the more general setting of the conjecture: on the one hand, sequences  $(h_n)_{n \geq 1}$  satisfying (4.1.31) will be handled by properly defining intertwined integrals against the field (that is  $d\mathcal{M}(\cdot)$ ) and Lebesgue measure  $(\widehat{h} d\lambda_2(\cdot))$  in the series (4.1.33) —we already have all required estimates to do so. On the other hand, the main result for *quenched* and *conditioned* partition functions will be obtained via small adaptations to our propositions; most notably replacing  $\psi^{\text{free}}$  with  $\psi^{\text{q}}$ ,  $\psi^{\text{cond}}$  in Proposition 4.3.7, with little technicality added to the current proof —which is already very involved, so we focused on the free variant only in this thesis.

## Appendices to Chapter 4

### 4.A Technical estimates on $\overline{M}_n$ , $n \geq 1$

In this section we provide some technical estimates on the field  $\overline{M}_n$ ,  $n \geq 1$ . Recall its definition from (4.2.6) and recall also the definition of increments of a field from (4.1.18), and that  $\Delta_n := \frac{1}{n}[\mathbf{0}, \mathbf{1}]$ . Set  $\mathbf{t} \in (\mathbb{R}_+^*)^2$ ,  $r \in \{1, 2\}$  and  $\mathbb{P} \in \mathfrak{P}_r$ . For clarity's sake, we will write  $\zeta_{\mathbf{i},n} := \zeta_{\mathbf{i}}(\beta_n) = e^{\beta_n \omega_{\mathbf{i}} - \lambda(\beta_n)} - 1$  for  $\mathbf{i} \in \mathbb{N}^2$ ,  $n \in \mathbb{N}$ .

**Lemma 4.A.1.** *There exists  $C > 0$  such that for  $n \in \mathbb{N}$  and  $\mathbf{u}, \mathbf{v} \in [\mathbf{0}, \mathbf{t}]$ ,*

$$\left\| \overline{M}_n(\mathbf{u}) - \overline{M}_n(\mathbf{v}) \right\|_{L^2}^2 \leq C \|\mathbf{t}\|_1 \left( \|\mathbf{t}\|_1 + \frac{1}{n\beta_n^{2(r-1)}} \right) \|\mathbf{u}^{[n]} - \mathbf{v}^{[n]}\|_1, \quad (4.A.1)$$

*Proof.* For  $n \in \mathbb{N}$ , Lemma 4.2.1 implies that

$$\begin{aligned} \mathbb{E} \left[ \left( \overline{M}_n(\mathbf{u}) - \overline{M}_n(\mathbf{v}) \right)^2 \right] &\leq \frac{1}{n^3 \beta_n^{2r}} \sum_{\mathbf{i}, \mathbf{j} \in [\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}]} \mathbb{E}[\zeta_{\mathbf{i},n} \zeta_{\mathbf{j},n}] \\ &\leq \frac{C}{n^3 \beta_n^{2(r-1)}} \text{Card}([\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}]) + \frac{C}{n^3} \text{Card}(\{\mathbf{i}, \mathbf{j} \in [\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}], \mathbf{i} \leftrightarrow \mathbf{j}\}), \end{aligned}$$

for some  $C > 0$ , where  $\Delta$  denotes the symmetric difference of two sets.

Without loss of generality, let us assume  $u_1 \leq v_1$ . Notice that  $[\mathbf{0}, \mathbf{u}] \Delta [\mathbf{0}, \mathbf{v}]$  can always be written as the union of two (disjoint) rectangles:

$$[\mathbf{0}, \mathbf{u}] \Delta [\mathbf{0}, \mathbf{v}] = \begin{cases} ([0, u_1] \times (u_2, v_2]) \cup ((u_1, v_1] \times [0, v_2]) & \text{if } u_2 \leq v_2, \\ ([0, u_1] \times (v_2, u_2]) \cup ((u_1, v_1] \times [0, v_2]) & \text{if } u_2 > v_2, \end{cases}$$

and  $[\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}] = (n \times [\mathbf{0}, \mathbf{u}] \Delta [\mathbf{0}, \mathbf{v}]) \cap \mathbb{Z}^2$ . Hence for some  $C_1 > 0$ , we have on the one hand,

$$\begin{aligned} \text{Card}([\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}]) &\leq C_1 m^2 (u_1^{[n]} + v_2^{[n]}) \|\mathbf{u}^{[n]} - \mathbf{v}^{[n]}\|_1 \\ &\leq C_1 n^2 \|\mathbf{t}\|_1 \|\mathbf{u}^{[n]} - \mathbf{v}^{[n]}\|_1, \end{aligned}$$

and on the other hand,

$$\text{Card}(\{\mathbf{i}, \mathbf{j} \in [\mathbf{0}, n\mathbf{u}] \Delta [\mathbf{0}, n\mathbf{v}], \mathbf{i} \leftrightarrow \mathbf{j}\}) \leq C_1 n^3 \|\mathbf{t}\|_1^2 \|\mathbf{u}^{[n]} - \mathbf{v}^{[n]}\|_1,$$

(details are left to the reader), which concludes the proof. □



**Lemma 4.A.2.** *There exists  $C > 0$  such that, for  $n \geq m$  and  $\mathbf{u}, \mathbf{v} \in [\mathbf{0}, \mathbf{t}]$ ,*

$$0 \leq \mathbb{E} \left[ \overline{M}_n(\mathbf{u}^{[m]} + \Delta_m) \overline{M}_n(\mathbf{v}^{[m]} + \Delta_m) \right] \leq \frac{C}{m^3} \mathbf{1}_{\{\mathbf{u}^{[m]} \leftrightarrow \mathbf{v}^{[m]}\}} + \frac{C}{n \beta_n^{2(r-1)} m^2} \mathbf{1}_{\{\mathbf{u}^{[m]} = \mathbf{v}^{[m]}\}}. \quad (4.A.2)$$

*Proof.* We write for  $n \geq m$  and  $\mathbf{i}, \mathbf{j} \in \mathbb{N}_0^2$ , such that  $\mathbf{i}, \mathbf{j} \preceq m\mathbf{t}$ ,

$$\mathbb{E} \left[ \overline{M}_n\left(\frac{1}{m}\mathbf{i} + \Delta_m\right) \overline{M}_n\left(\frac{1}{m}\mathbf{j} + \Delta_m\right) \right] = \frac{1}{n^3 \beta_n^{2r}} \sum_{\substack{\mathbf{k} \in \frac{n}{m} \llbracket \mathbf{i}, \mathbf{i}+1 \rrbracket, \\ \mathbf{l} \in \frac{n}{m} \llbracket \mathbf{j}, \mathbf{j}+1 \rrbracket}} \mathbb{E}[\zeta_{\mathbf{k},n} \zeta_{\mathbf{l},n}] \geq 0.$$

For the upper bound, recall estimates on  $\mathbb{E}[\zeta_{\mathbf{k},n} \zeta_{\mathbf{l},n}]$ ,  $\mathbf{k}, \mathbf{l} \in \mathbb{N}^2$  from Lemma 4.2.1. If  $\mathbf{i} \leftrightarrow \mathbf{j}$ , then there is no pair of indices  $(\mathbf{k}, \mathbf{l})$  in the sum such that  $\mathbf{k} \leftrightarrow \mathbf{l}$ , hence the sum is equal to zero. If  $\mathbf{i} \leftrightarrow \mathbf{j}$ ,  $\mathbf{i} \neq \mathbf{j}$ , then there are at most  $O((n/m)^3)$  contributing pairs of indices  $(\mathbf{k}, \mathbf{l})$  in the sum, each yielding a term  $O(\beta_n^{2r})$ . Therefore, there exists  $C > 0$  such that

$$\mathbb{E} \left[ \overline{M}_n\left(\frac{1}{m}\mathbf{i} + \Delta_m\right) \overline{M}_n\left(\frac{1}{m}\mathbf{j} + \Delta_m\right) \right] \leq \frac{1}{n^3 \beta_n^{2r}} \times C \beta_n^{2r} (n/m)^3 = \frac{C}{m^3}.$$

Similarly, if  $\mathbf{i} = \mathbf{j}$ , there are at most  $O((n/m)^3)$  pairs of indices  $\mathbf{k} \leftrightarrow \mathbf{l}$ ,  $\mathbf{k} \neq \mathbf{l}$  yielding a contribution  $O(\beta_n^{2r})$ , and  $O((n/m)^2)$  pairs  $\mathbf{k} = \mathbf{l}$  yielding  $O(\beta_n^2)$ . Therefore, there exists  $C > 0$  such that

$$\mathbb{E} \left[ \overline{M}_n(\mathbf{i} + \Delta_m) \overline{M}_n(\mathbf{j} + \Delta_m) \right] \leq \frac{C}{n^3 \beta_n^{2r}} \times \left( \beta_n^{2r} (n/m)^3 + \beta_n^2 (n/m)^2 \right),$$

which completes the proof. □

# Conclusion and perspectives

## Summary and open questions

Let us bring this thesis to a close by recalling our main observations and results, and presenting some perspectives for future research that come out of them. Our aim was to investigate perturbations of pinning/wetting models, by adding new elements to the system and quantifying their effect on the pinning phase transition. We did so in two different ways, yielding very different consequences.

In Part I we added another interaction to the model, more precisely a repulsion by the solvent (or equivalently a self-attraction). For some values of the energetic parameters, this interaction may prompt the polymer to form a compact globule instead of extending into the solvent. We called this regime the Collapsed phase, and we noticed that the full polymer cannot undergo a pinning transition in that regime; however, its lower envelope does, displaying a phenomenon we called a *surface transition*. This surface pinning transition has very similar properties to the original pinning phase transition, albeit occurring on a much different size scale for the model (i.e. at its outer layer). In other words, the energetic constraints added on trajectories by the interaction with the solvent in the Collapsed phase do not fully blot out the pinning transition, but they restrict it to a smaller part of the system.

In Part II we considered a “generalized” version of the pinning model, called gPS, where we added inhomogeneities (or disorder) in the composition of the polymers, ergo in pinning interaction energies. We proved that the pinning phase transition still occurs in that setting, and for small disorder intensity, we quantified its potential effect on the phase transition in the form of a shift of the critical point and a smoothing of the phase transition. This allowed us to establish criteria for the relevance or irrelevance of disorder, i.e. whether critical properties of the system around its phase transition are altered or not. We then took another approach via a scaling limit of the model, showing that the potential relevance of disorder also manifests in the “intermediate disorder” regime, and supporting both our previous results and the observation that some link seems to relate the questions of disorder (ir)-relevance and of the existence of intermediate disorder non-trivial limits, as was already conjectured in the literature.

In each approach, we believe that our results bring some additional understanding to issues of genuine interest in the probability and statistical mechanics literature. Nonetheless, many open questions remain in each setting: let us list several here (we already mentioned most of them in Chapter “Introduction et vue d’ensemble”), starting with the IPDSAW

---

interacting with a wall.

- We actually proved the surface transition in the “one-bead” restriction of the model: can we also prove it in the non-restricted variant? We may have already displayed all the tools required to do so, but the technical difficulty of handling *all* types of beads is substantial.
- Can we improve our results regarding the lower envelope, that is for a random walk with both wetting and pre-wetting terms? In the Desorbed-Collapsed regime we only have sharp estimates when  $\delta = 0$ , so extending our estimates to the whole regime would be a meaningful improvement. Moreover, pre-wetting phenomena are of particular interest to the literature in general, and sharp estimates would be of use beyond the setting of this thesis.
- Regarding the Extended regime, we conjecture that it can be divided into two phases  $\mathcal{AE}$  and  $\mathcal{DE}$ , each with a different, analytical expression for the free energy. However our approach with upper/lower envelopes fails to encapsulate the extended behavior efficiently, so another method may be required altogether.
- The model undergoes several phase transitions, for instance from  $\mathcal{DC}$  to  $\mathcal{E}$  and from  $\mathcal{AC}$  to  $\mathcal{E}$ : do they have the same critical behavior? In the IPDSAW model (without wall), the order of the phase transition  $\mathcal{E} \leftrightarrow \mathcal{C}$  has been proven to be  $3/2$ . We may conjecture the same for the transition  $\mathcal{E} \leftrightarrow \mathcal{DC}$  in the model with wall, but it is not clear whether it remains the same or not for  $\mathcal{E} \leftrightarrow \mathcal{AC}$ .
- Finally, what about critical regimes? Is there a “surface” pinning transition as  $\delta$  goes to  $\delta_c(\beta_c)$  when  $(\beta, \delta) \in \overline{\mathcal{C}} \cap \overline{\mathcal{E}}$ ? And in the Collapsed phase, what does the lower envelope of the globule look like on the critical curve  $(\beta, \delta) \in \overline{\mathcal{DC}} \cap \overline{\mathcal{AC}}$ ?

And for the gPS model:

- Our results mostly hold for small intensities of disorder, and it is not clear that  $\beta \mapsto h_c(\beta)$  is monotone (because of the normalizing term  $-\lambda(\beta)$ ). Hence, we cannot yet draw conclusions on the critical behavior for larger disorder intensities, and studying the model for generic  $\beta$  might well be an intricate matter.
- Regarding marginal values of  $\alpha$  (that is  $\alpha = 1/2$  if  $\mathbb{P} \in \mathfrak{P}_1 \cup \mathfrak{P}_2$  and  $\alpha = 1$  if  $\mathbb{P} \in \mathfrak{P}_0$ , recalling (4.1.6)), we only have sufficient conditions for disorder irrelevance in Theorem 3.2.1-3.2.2. We conjecture that disorder is otherwise relevant, but proving it may be technically very involved.
- The homogeneous gPS model undergoes other phase transitions, called “condensation” transitions. What is the effect of disorder on those? Do they still occur? In particular, recent results for the pinning model [86] let us conjecture that they actually vanish altogether in the disordered model.
- Regarding the choice of disorder, our results should apply to a generic interaction function  $f$ , since the determining factor seems to be the correlation structure of the field  $(e^{\beta\omega_i})_{i \in \mathbb{N}^2}$ . However, having in mind that the gPS model aims to portray DNA

---

denaturation, where the two strands are constructed symmetrically, a much more appropriate choice of disorder would be

$$\omega_{\mathbf{i}} := f(\widehat{\omega}_{i_1}, \widehat{\omega}_{i_2}), \quad \mathbf{i} = (i_1, i_2) \in \mathbb{N}^2,$$

for some random sequence  $\widehat{\omega} = (\widehat{\omega}_i)_{i \in \mathbb{N}}$ . This setting has much more intricate correlations, in particular it is not clear that the quenched free energy is well-defined, and the annealed model does not match the homogeneous one anymore.

- Finally, determining intermediate disorder regimes for  $\alpha > 1$  seems an interesting issue. On the one hand, the definition of limiting field  $\mathcal{M}$  for the disorder is not obvious. When  $\mathbb{P} \in \mathfrak{P}_0$ , a computation of correlations let us believe that the limit will be a product of independent Brownian motions,  $(\mathcal{B}^1(s_1)\mathcal{B}^2(s_2))_{(s_1, s_2) \in (\mathbb{R}_+^*)^2}$ . Otherwise, the diagonal terms  $\mathbf{i} = \mathbf{j}$  in (4.2.7) may contribute to the limit, yielding a different continuous field. On the other hand, the bi-dimensional renewal process  $\tau$  has a very different behavior for  $\alpha > 1$ , localizing around the diagonal with deviations of order  $n^{1/\min(\alpha, 2)}$  (see [17]). What consequences does it have on the continuous renewal mass function  $\varphi$ , and on its integration against  $\mathcal{M}$ ?

### A few steps further

All things considered, perspectives for further research in each of those settings are numerous. However, let us not restrict ourselves to those separately: we may phrase many interesting problems by taking a step back and broadening our vision, in particular when considering both approaches simultaneously.

Indeed, let us look at the IPDSAW model interacting with a wall, and let us add some disorder to it. The first, simplest way to do so is adding disorder only to the contact interactions with the wall, and with an i.i.d. field. Similarly to the pinning model, we can expect the phase diagram of the annealed model to be a shift in  $\delta$  (by  $\lambda(\gamma) := \log(\mathbb{E}[e^{\gamma\omega_1}])$ , with  $\gamma \geq 0$  the disorder intensity) of the homogeneous variant. Moreover, the behavior (with respect to the wall) of the full polymer in the Extended regime, and of the lower envelope in the Collapsed phase, should be quite similar to that of the pinning model. Therefore, studying disorder (ir)-relevance on both transitions  $\mathcal{AE} \leftrightarrow \mathcal{DE}$  and  $\mathcal{AC} \leftrightarrow \mathcal{DC}$  should prove to be interesting, keeping in mind that disorder is necessarily “irrelevant” on the transition  $\mathcal{AC} \leftrightarrow \mathcal{DC}$  at the level of the free energy (since it is constant), but its effects should appear at the surface order. This should be yet another perturbation of the adsorption transition, combining observations from both approaches we presented in this thesis.

Another way to add disorder to this model is regarding self-touching interactions. Let us consider the model without a wall for simplicity, and let the self-touching between monomers  $i$  and  $j$  yield some energy  $\gamma\omega_{i,j} + \beta$ ,  $0 \leq i < j - 1 < L$  and  $\gamma \geq 0$ . Assuming first that  $\omega$  is an i.i.d. field, the annealed version model should be a shift in  $\beta$  of the homogeneous one (very similarly to what unfolds in the pinning model). Proving the existence of a quenched free energy should be accessible too. However, measuring the effect of the disorder on the collapse phase transition in the quenched model does not seem

---

straightforward at all, and should prove to be an interesting matter. Still, a choice of disorder field better suited to portray the inhomogeneity of the polymer would be  $\omega_{i,j} = f(\hat{\omega}_i, \hat{\omega}_j)$  for some sequence  $(\hat{\omega}_i)_{i \in \mathbb{N}}$  and interaction function  $f$ . Similarly to our discussion above regarding the choice of disorder in the gPS model, this model should be extremely intricate and difficult to handle, yet very pertinent for modeling purposes.

Let us conclude with a last perspective, regarding our construction of integrals against random fields. We introduced this construction with the hope that it may be general, or at least have some degree of application beyond the gPS model. In particular, the (one-dimensional) pinning model with correlated disorder has greatly interested mathematicians, and many questions remain open. We expect that, for Gaussian disorder with long-range correlations (see e.g. [27, 134]), our construction of stochastic integrals will yield an intermediate-disorder scaling limit, thus making progress on the question of correlated disorder (ir)-relevance. Thereby, if we manage to extend the applications of our work to other settings, this would both give new results on existing models, and support our construction as being a pertinent tool to investigate statistical mechanics systems with strong correlations.

The takeaway from this thesis is that the research field of polymer models is both very intricate and vivid. It explores many phenomena of great interest, from a mathematical standpoint as well as to other specialties, and provides numerous challenging and profound issues.

## Appendix: Notation

In order to remain consistent across English and French parts, this thesis uses English conventions for mathematical notation:  $\log$  denotes the natural logarithm (base  $e \approx 2.718\dots$ ),  $\mathbb{N} = \{1, 2, \dots\}$  the set of positive integers,  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ , and for  $a, b \in \overline{\mathbb{R}}$ ,  $a < b$ ,  $(a, b)$  denotes the open interval from  $a$  to  $b$ .

From a chapter to another, similar notation may be used in different models, for instance  $F$  for the free energy and  $Z$  for the partition function, even though their definitions are different. However, what definition the notation refers to is always clear from context and consistent throughout a chapter.

Regarding two-dimensional mathematical objects, unless otherwise specified, we use bold characters ( $\mathbf{n}, \mathbf{t}, \boldsymbol{\tau}_i, \dots$ ) to denote elements of  $\mathbb{N}^2$  or  $\mathbb{R}^2$ , and plain characters ( $n, t, \tau_i, \dots$ ) for elements of  $\mathbb{N}$  or  $\mathbb{R}$ . When there is no risk of confusion, the coordinates of some  $\mathbf{t} \in \mathbb{R}^2$  may be noted  $\mathbf{t} = (t_1, t_2)$ ; otherwise, the projection of  $\mathbf{t}$  on its  $r$ -th coordinate is written  $\mathbf{t}^{(r)}$ ,  $r \in \{1, 2\}$ .

Let  $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$  be real-valued sequences, then we write  $a_n \sim b_n$  if there is  $n_0 \in \mathbb{N}$  such that  $a_n, b_n \neq 0$  for  $n \geq n_0$ , and if  $\frac{a_n}{b_n} \rightarrow 1$  as  $n \rightarrow \infty$ . We also write

$$a_n \asymp b_n, \quad (5.0.1)$$

if there is  $C > 0$  and  $n_0 \in \mathbb{N}$  such that  $\frac{1}{C}a_n \leq b_n \leq Ca_n$  for all  $n \geq n_0$  (if  $a_n, b_n > 0$  for all  $n \geq 1$ , one can assume  $n_0 = 1$  by adjusting  $C > 0$ ).

Some function  $L : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$  is called *slowly varying at  $+\infty$*  if it satisfies

$$\frac{L(tx)}{L(t)} \xrightarrow{t \rightarrow +\infty} 1, \quad \forall x > 0, \quad (5.0.2)$$

and it is *slowly varying at 0* if  $t \mapsto L(1/t)$  is slowly varying at  $+\infty$ . A function  $f : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$  is called *regularly varying at  $+\infty$*  (resp. at 0) if there exists  $\alpha \in \mathbb{R}$  such that  $t \mapsto t^{-\alpha}f(t)$  is slowly varying at  $+\infty$  (resp. at 0). More details on regularly varying functions can be found in [30].

Let  $f, g : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ , then we write

$$f(x) \asymp_{\text{s.v.}}^{+\infty} g(x), \quad (5.0.3)$$

if there exist slowly varying functions at  $+\infty$ ,  $L_1, L_2$  and  $t_0 > 0$  such that  $L_1(t)f(t) \leq g(t) \leq L_2(t)f(t)$  for all  $t \geq t_0$ ; and we write  $f(x) \asymp_{\text{s.v.}}^0 g(x)$  for the analogous claim at

---

0, i.e.  $f(1/x) \asymp_{s.v.}^{+\infty} g(1/x)$ . We may drop the notation at  $+\infty$  or at 0 if it is clear from context.

## References

- [1] A. Johner and J. F. Joanny, « Polymer adsorption in a poor solvent », in: *J. Phys. II France* 1.2 (1991), pp. 181–194.
- [2] T. Alberts, K. Khanin, and J. Quastel, « The continuum directed random polymer », in: *J. Stat. Phys.* 154 (2014), pp. 305–326.
- [3] T. Alberts, K. Khanin, and J. Quastel, « The intermediate disorder regime for directed polymers in dimension  $1 + 1$  », in: *The Annals of Probability* 42.3 (2014), pp. 1212–1256.
- [4] S. Albeverio. and X. Zhou, « A martingale approach to directed polymers in a random environment », in: *J. Theoret. Probab.* 9 (1996), pp. 171–189.
- [5] K. S. Alexander, « Excursions and local limit theorems for Bessel-like random walks », in: *Electron. J. Probab.* 16 (2011), pp. 1–44.
- [6] K. S. Alexander, « The effect of disorder on polymer depinning transitions », in: *Comm. Math. Phys.* 279 (2008), pp. 117–146.
- [7] K. S. Alexander and V. Sidoravicius, « Pinning of polymers and interfaces by random potentials », in: *The Annals of Applied Probability* 16.2 (2006), pp. 636–669.
- [8] K. S. Alexander and N. Zygouras, « Quenched and annealed critical points in polymer pinning models », in: *Commun. Math. Phys.* 291.3 (2009), pp. 659–689.
- [9] K. S. Alexander, N. Zygouras, et al., « Equality of critical points for polymer depinning transitions with loop exponent one », in: *The Annals of Applied Probability* 20.1 (2010), pp. 356–366.
- [10] G. Amir, I. Corwin, and J. Quastel, « Probability distribution of the free energy of the continuum directed random polymer in  $1 + 1$  dimensions », in: *Commun. Pure Appl. Math.* 64 (2011), pp. 466–537.
- [11] S. Asmussen, *Applied probability and queues*, Second, vol. 51, Applications of Mathematics (New York), Stochastic Modelling and Applied Probability, Springer-Verlag, New York, 2003, pp. xii+438, ISBN: 0-387-00211-1.
- [12] A. Auffinger and M. Damron, « The scaling relation  $\chi = 2\xi - 1$  for directed polymers in a random environment », in: *ALEA, Lat. Am. J. Probab. Math. Stat.* 10.2 (2013), pp. 857–880.



- 
- [13] A. Auffinger and O. Louidor, « Directed Polymers in Random Environment with Heavy Tails », in: *Comm. on Pure and Applied Math.* 64 (2011), pp. 183–204.
- [14] J. Baik, P. Deift, and K. Johansson, « On the distribution of the longest increasing subsequence of random permutations », in: *J. Amer. Math. Soc.* 12 (1999), pp. 1119–1178.
- [15] M. Balázs, J. Quastel, and T. Seppäläinen, « Fluctuation exponent of the KPZ /stochastic Burgers equation », in: *J. Amer. Math. Soc.* 24.3 (2011), pp. 683–708.
- [16] Q. Berger, « Polymères en milieu aléatoire : influence d’un désordre corrélé sur le phénomène de localisation », PhD thesis, ENS Lyon, 2012.
- [17] Q. Berger, « Strong renewal theorems and local large deviations for multivariate random walks and renewals », in: *Electron. J. Probab.* 24.46 (2019), pp. 1–47.
- [18] Q. Berger, G. Giacomin, and M. Khatib, « Disorder and denaturation transition in the generalized Poland-Scheraga model », in: *Ann. Henri Lebesgue* (to appear).
- [19] Q. Berger, G. Giacomin, and M. Khatib, « DNA melting structures in the generalized Poland-Scheraga model », in: *ALEA, Lat. Am. J. Probab. Math. Stat.* 15 (2018), pp. 993–1025.
- [20] Q. Berger and H. Lacoïn, « Pinning on a defect line: characterization of marginal disorder relevance and sharp asymptotics for the critical point shift », in: *Journal of the Institute of Mathematics of Jussieu* 17.2 (2018), pp. 305–346.
- [21] Q. Berger and A. Legrand, *Scaling limit of the disordered generalized Poland-Scheraga model*, work in progress.
- [22] Q. Berger and N. Torri, « Beyond Hammersley’s Last-Passage Percolation: a discussion on possible new local and global constraints », in: *arXiv:1802.04046* (2018).
- [23] Q. Berger and N. Torri, « Directed polymers in heavy-tail random environment », in: *arXiv:1802.03355* (2018).
- [24] Q. Berger et al., « The critical curve of the random pinning and copolymer models at weak coupling », in: *Commun. Math. Phys.* 326 (2014), pp. 507–530.
- [25] Q. Berger, « Pinning Model in Random Correlated Environment: Appearance of an Infinite Disorder Regime », in: *Journal of Statistical Physics* 155 (2013).
- [26] Q. Berger and H. Lacoïn, « Sharp critical behavior for pinning models in a random correlated environment », in: *Stochastic Processes and their Applications* 122.4 (2012), pp. 1397–1436, ISSN: 0304-4149.
- [27] Q. Berger and J. Poisat, « On the critical curves of the pinning and copolymer models in correlated Gaussian environment », in: *Electronic Journal of Probability* 20.none (2015), pp. 1–35.

- 
- [28] P. J. Bickel and M. J. Wichura, « Convergence criteria for multiparameter stochastic processes and some applications (article) Convergence criteria for multiparameter stochastic processes and some applications », in: *Ann. Math. Stat.* 42.5 (1971), pp. 1656–1670.
- [29] P. Billingsley, *Convergence of Probability Measures*, Wiley, 1968.
- [30] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1987.
- [31] G. Biroli, J. Bouchaud, and M. Potters, « Extreme value problems in Random Matrix Theory and other disordered systems », in: *J. Stat. Mech.* 7 (2007), P07019, 15 pp.
- [32] R. D. Blake and S. G. Delcourt, « Thermal stability of DNA », in: *Nucleic Acids Research* 26.14 (1998), pp. 3323–3332, ISSN: 0305-1048.
- [33] R. D. Blake et al., « Statistical mechanical simulation of polymeric DNA melting with MELTSIM », in: *Bioinformatics (Oxford, England)* 15 (1999), pp. 370–5.
- [34] E. Bolthausen, « A note on diffusion of directed polymer in a random environment », in: *Commun. Math. Phys.* 123 (1989), pp. 529–534.
- [35] E. Bolthausen and F. d. Hollander, « Localization transition for a polymer near an interface », in: *Ann. Probab.* 25 (1997), pp. 1334–1366.
- [36] E. Bolthausen, « On a Functional Central Limit Theorem for Random Walks Conditioned to Stay Positive », in: *Ann. Probab.* 4.3 (1976), pp. 480–485.
- [37] A. A. Borovkov, *Probability theory*, Universitext, Translated from the 2009 Russian fifth edition by O. B. Borovkova and P. S. Ruzankin, Edited by K. A. Borovkov, Springer, London, 2013, pp. xxviii+733, ISBN: 978-1-4471-5200-2; 978-1-4471-5201-9.
- [38] A. A. Borovkov and A. A. Mogulskii, « On large and superlarge deviations of sums of independent random vectors under Cramér’s condition. I », in: *Theory Probab. Appl.* 51.2 (2007), pp. 227–255.
- [39] A. A. Borovkov and A. A. Mogulskii, « On large and superlarge deviations of sums of independent random vectors under Cramér’s condition. II », in: *Theory Probab. Appl.* 51.4 (2007), pp. 567–594.
- [40] R. Brak, A. J. Guttmann, and S. G. Whittington, « A Collapse Transition in a Directed Walk Model », in: *J. Phys. A: Math Gen* (1992), pp. 2437–46.
- [41] R. Cairoli and J. B. Walsh, « Stochastic integrals in the plane », in: *Acta Mathematica* 134.none (1975), pp. 111–183.
- [42] F. Caravenna and G. Giacomin, « The weak coupling limit of disordered copolymer models », in: *Ann. Probab.* 38.6 (2010), pp. 2322–2378.
- [43] F. Caravenna and F. d. Hollander, « A general smoothing inequality for disordered polymers », in: *Report Eurandom* (2013).

- 
- [44] F. Caravenna, R. Sun, and N. Zygouras, « Polynomial chaos and scaling limits of disordered systems », in: *J. Eur. Math. Soc. (JEMS)* 19 (2017), pp. 1–65.
- [45] F. Caravenna, R. Sun, and N. Zygouras, « The Continuum Disordered Pinning Model », in: *Probab. Theory Related Fields* 164 (2016), pp. 17–59.
- [46] F. Caravenna, « A local limit theorem for random walks conditioned to stay positive », in: *Probability Theory and Related Fields* 133.4 (2005), pp. 508–530.
- [47] F. Caravenna and L. Chaumont, « An invariance principle for random walk bridges conditioned to stay positive », in: *Electron. J. Probab.* 18 (2013), 32 pp.
- [48] F. Caravenna, G. Giacomin, and L. Zambotti, « Sharp asymptotic behavior for wetting models in (1+1)-dimension », in: *Electron. J. Probab.* 11 (2006), pp. 345–362.
- [49] F. Caravenna, R. Sun, and N. Zygouras, *Scaling limits of disordered systems and disorder relevance*, 2016, arXiv: 1602.05825 [math.PR].
- [50] F. Caravenna, F. L. Toninelli, and N. Torri, « Universality for the pinning model in the weak coupling regime », in: *Annals of Probability* 45.4 (2017), pp. 2154–2209.
- [51] P. Carmona and Y. Hu, « Strong disorder implies strong localization for directed polymers in a random environment », in: *ALEA* 2 (2006), pp. 217–229.
- [52] P. Carmona, G. B. Nguyen, and N. Pétrélis, « Interacting partially directed self avoiding walk. From phase transition to the geometry of the collapsed phase », in: *Ann. Probab.* 44.5 (2016), pp. 3234–3290, ISSN: 0091-1798.
- [53] P. Carmona and N. Pétrélis, « Interacting partially directed self avoiding walk: scaling limits », in: *Electron. J. Probab.* 21 (2016), 52 pp.
- [54] P. Carmona et al., « Interacting partially directed self-avoiding walk: a probabilistic perspective », in: *Journal of Physics A: Mathematical and Theoretical* 51.15 (2018), p. 153001.
- [55] D. Cheliotis and F. d. Hollander, « Variational characterization of the critical curve for pinning of random polymers », in: *Ann. Probab.* 41 (2013), pp. 1767–1805.
- [56] D. Cheliotis, Y. Chino, and J. Poisat, « The random pinning model with correlated disorder given by a renewal set », in: *Annales Henri Lebesgue* (2019).
- [57] F. Comets, *Directed Polymers in Random Environments*, vol. 2175, Ecole d’Eté de probabilités de Saint-Flour, Springer International Publishing, 2016, ISBN: 978-3-319-50486-5.
- [58] F. Comets, « Weak disorder for low dimensional polymers: The model of stable laws », in: *Markov Processes and Related Fields* (2006).
- [59] F. Comets, T. Shiga, and N. Yoshida, « Probabilistic analysis of directed polymers in a random environment: a review », in: *Stochastic analysis on large scale interacting systems*, Adv. Stud. Pure Math. 39, Math. Soc. Japan, Tokyo (2004), pp. 115–142.

- 
- [60] F. Comets and N. Yoshida, « Directed polymers in random environment are diffusive at weak disorder », in: *Ann. Probab.* 34 (2006), pp. 1746–1770.
- [61] F. Comets et al., « Limiting results for the free energy of Directed Polymers in random environment with unbounded jumps », in: *J. Stat. Phys.* 161.3 (2015), pp. 577–597.
- [62] F. Cravenna, R. Sun, and N. Zygouras, « Universality in marginally relevant systems », in: *arXiv:1510.06287v2 [math.PR]* (2015).
- [63] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Springer, 1998.
- [64] B. Derrida et al., « Fractional moment bounds and disorder relevance for pinning models », in: *Comm. Math. Phys.* 287 (2009), pp. 867–887.
- [65] R. Dobrushin and O. Hryniv, « Fluctuations of shapes of large areas under paths of random walks », in: *Probab. Theory Related Fields* 105.4 (1996), pp. 423–458, ISSN: 0178-8051.
- [66] R. A. Doney, « An analogue of renewal theorems in higher dimensions », in: *Proc. London Math. Soc.* 16.3 (1966), pp. 669–684.
- [67] R. A. Doney, « One-sided local large deviation and renewal theorems in the case of infinite mean », in: *Probab. Theory Related Fields* 107.4 (1997), pp. 451–465, ISSN: 0178-8051.
- [68] H. Duminil-Copin et al., « On the number of maximal paths in directed last-passage percolation », in: *arXiv:1801.05777 [math.PR]* (2018).
- [69] R. T. Durrett, D. L. Iglehart, and D. R. Miller, « Weak Convergence to Brownian Meander and Brownian Excursion », in: *Ann. Probab.* 5.1 (1977), pp. 117–129.
- [70] K. B. Erickson, « Strong renewal theorems with infinite mean », in: *Transactions of the American mathematical society* 151.1 (1970), pp. 263–291.
- [71] W. Feller, *An Introduction to Probability Theory and its Applications, Vol I*, Wiley series in probability and mathematical statistics, New York-London-Sydney: John Wiley & Sons. Inc., 1966.
- [72] D. S. Fisher and D. A. Huse, « Directed paths in random potential », in: *Phys. Rev. B* 43 (1991), pp. 10728–10742.
- [73] M. E. Fisher, « Walks, walls, wetting, and melting », in: *Journal of Statistical Physics* 34 (1984), pp. 667–729.
- [74] S. Foss, D. Korshunov, and S. Zachary, *An introduction to heavy-tailed and subexponential distributions. 1st ed*, vol. 38, 2011.
- [75] D. P. Foster, « Exact evaluation of the collapse phase boundary for two-dimensional directed polymers », in: *Journal of Physics A: Mathematical and General* 23.21 (1990), pp. L1135–L1138.

- 
- [76] D. P. Foster and J. Yeomans, « Competition between self-attraction and adsorption in directed self-avoiding polymers », in: *Physica A: Statistical Mechanics and its Applications* 177.1 (1991), pp. 443–452, ISSN: 0378-4371.
- [77] T. Garel and H. Orland, « Generalized Poland-Scheraga model for DNA hybridization », in: *Biopolymers* 75 (2004), pp. 453–467.
- [78] T. Garel and H. Orland, *On the role of mismatches in DNA denaturation*, 2003, arXiv: cond-mat/0304080.
- [79] A. Garsia and J. Lamperti, « A discrete renewal theorem with infinite mean », in: *Comm. Math. Helv.* 37 (1963), pp. 221–234.
- [80] A. M. Garsia, « Continuity Properties of Gaussian processes with multidimensional time parameter », in: *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob II* (1972), pp. 369–374.
- [81] A. Garsia, E. Rodemich, and H. J. Rumsey, « A real variable lemma and the continuity of paths of some Gaussian processes », in: *Indiana Univ. Math. Journal* 20 (1970/1971), pp. 565–578.
- [82] T. Geudré et al., « Ground-state statistics of directed polymers with heavy-tailed disorder », in: *Phys. Rev. E* 91 (2016).
- [83] G. Giacomin, *Disorder and Critical Phenomena Through Basic Probability Models*, vol. 2025, Ecole d’Eté de probabilités de Saint-Flour, Springer-Verlag Berlin Heidelberg, 2010.
- [84] G. Giacomin, *Random Polymer Models*, Imperial College Press, World Scientific, 2007.
- [85] G. Giacomin, « Renewal sequences, disordered potentials, and pinning phenomena », in: *Spin Glasses: Statics and Dynamics* (2008), pp. 235–270.
- [86] G. Giacomin and B. Havret, « Localization, Big-Jump Regime and the Effect of Disorder for a Class of Generalized Pinning Models », in: *J. Stat. Phys.* (2020).
- [87] G. Giacomin and M. Khatib, « Generalized Poland-Scheraga denaturation model and two-dimensional renewal processes », in: *Stoch. Proc. Appl.* 127.526-573 (2017).
- [88] G. Giacomin, H. Lacoin, and F. L. Toninelli, « Disorder relevance at marginality and critical point shift », in: *Ann. Inst. H. Poincaré: Prob. Stat.* 47 (2011), pp. 148–175.
- [89] G. Giacomin, H. Lacoin, and F. L. Toninelli, « Marginal relevance of disorder for pinning models », in: *Commun. Pure Appl. Math.* 63 (2010), pp. 233–265.
- [90] G. Giacomin and F. L. Toninelli, « Smoothing effect of quenched disorder on polymer depinning transitions », in: *Comm. Math. Phys.* 266 (2006), pp. 1–16.
- [91] G. Giacomin and F. L. Toninelli, « The localized phase of disordered copolymers with adsorption », in: *ALEA* 1 (2006), pp. 149–180.

- 
- [92] G. Giacomin, *Disorder and critical phenomena through basic probability models*, vol. 2025, Lecture Notes in Mathematics, Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School], Heidelberg: Springer, 2011, pp. xii+130, ISBN: 978-3-642-21155-3.
- [93] G. Giacomin, « Renewal convergence rates and correlation decay for homogeneous pinning models », in: *Electron. J. Probab* (2008), pp. 513–529.
- [94] A. J. Guttmann, « Analysis of series expansions for non-algebraic singularities », in: *Journal of Physics A: Mathematical and Theoretical* 48.4 (2015), p. 045209, ISSN: 1751-8121.
- [95] L. de Haan and S. I. Resnick, « Conjugate  $\Pi$ -variation and process inversion », in: *The Annals of Probability* (1979), pp. 1028–1035.
- [96] B. Hambly and J. B. Martin, « Heavy Tails in Last-Passage Percolation », in: *Probability Theory and Related Fields* 137 (2007), pp. 227–275.
- [97] J. M. Hammersley, « A few seedlings of research », in: *Proc. Sixth Berkeley Symp. Math. Statist. and Probab.* Ed. by U. of California Press, vol. 1, 1972, pp. 345–394.
- [98] A. B. Harris, « Effect of Random Defects on the Critical Behaviour of Ising Models », in: *J. Phys.* (1974), pp. 1671–1692.
- [99] R. van der Hofstad, *Random Graphs and Complex Networks: Volume 1*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press; 1st edition, 2016, p. 375, ISBN: 110717287X - 978-1107172876.
- [100] F. d. Hollander, *Random polymers*, vol. 1974, Lecture Notes in Mathematics, Lectures from the 37th Probability Summer School held in Saint-Flour, 2007, Springer-Verlag, Berlin, 2007, ISBN: 978-3-642-00332-5.
- [101] F. den Hollander, *Large deviations*, vol. 14, Fields Institute Monographs, American Mathematical Society, Providence, RI, 2000, pp. x+143, ISBN: 0-8218-1989-5.
- [102] O. Hryniv and Y. Velenik, « Universality of critical behaviour in a class of recurrent random walks », in: *Probab. Theory Related Fields* 130.2 (2004), pp. 222–258.
- [103] D. A. Huse and C. L. Henley, « Pinning and roughening of domain walls in Ising systems due to random impurities », in: *Phys. Rev. Lett.* 54 (1985), pp. 2708–2711.
- [104] F. Igloi, « Collapse transition of directed polymers: Exact results », in: *Physical review. A* 43 (1991), pp. 3194–3197.
- [105] J. Z. Imbrie and T. Spencer, « Diffusion of directed polymer in a random environment », in: *J. Stat. Phys.* 52 (1988), pp. 608–626.
- [106] D. Ioffe, S. Shlosman, and Y. Velenik, « An invariance principle to Ferrari-Spohn diffusions », in: *Comm. Math. Phys.* 336.2 (2015), pp. 905–932, ISSN: 0010-3616.

- 
- [107] D. Ioffe and Y. Velenik, « Low temperature interfaces: Prewetting, layering, faceting and Ferrari-Spohn diffusions », in: *Mark. Proc. Rel. Fields* 24.1 (2018), pp. 487–537.
- [108] S. Janson, « Brownian excursion area, Wright’s constants in graph enumeration, and other Brownian areas », in: *Probability Surveys* 4 (2007), pp. 80–145.
- [109] K. Johansson, « Transversal fluctuations for increasing subsequences on the plane », in: *Probab. Theory Rel. Fields* 116.4 (1999), pp. 445–456.
- [110] Y. Kafri, D. Mukamel, and L. Peliti, « Why is the DNA Denaturation Transition First Order? », in: *Physical review letters* 85 (2001), pp. 4988–91.
- [111] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed., Springer Probability and its Applications, Springer-Verlag New York, 2002.
- [112] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer-Verlag, 2000, ISBN: 9781468403046.
- [113] D. Khoshnevisa, *Multiparameter Processes: An Introduction to Random Fields*, Springer Monographs in Mathematics, Springer-Verlag New York, 2002, ISBN: 978-0-387-95459-2.
- [114] H. Lacoïn, « New Bounds for the Free Energy of Directed Polymers in Dimension  $1 + 1$  and  $1 + 2$  », in: *Communications in Mathematical Physics* 294 (2009).
- [115] H. Lacoïn, « The martingale approach to disorder irrelevance for pinning models », in: *Electron. Commun. Probab.* 15 (2010), pp. 418–427.
- [116] H. Lacoïn and A. Teixeira, « A mathematical perspective on metastable wetting », in: *Electronic Journal of Probability* 20.17 (2015), pp. 1–23.
- [117] G. F. Lawler and V. Limic, *Random Walk: A Modern Introduction*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2010, p. 376, ISBN: 9780521519182.
- [118] M. Ledoux, *The concentration of measure phenomenon*, vol. 89, American Mathematical Society, 2001.
- [119] A. Legrand, « Influence of disorder on DNA denaturation: the disordered generalized Poland-Scheraga model », in: *Electronic Journal of Probability* 26 (2021), pp. 1–43.
- [120] A. Legrand and N. Pétrélis, *A sharp asymptotics of the partition function for the collapsed interacting partially directed self-avoiding walk*, preprint, 2021, arXiv: 2104.01247.
- [121] A. Legrand and N. Pétrélis, *Surface transition in the collapsed phase of a self-interacting walk adsorbed along a hard wall*, preprint, 2020, arXiv: 2007.11313.
- [122] B. F. Logan and L. A. Shepp, « A variational problem for random Young tableaux », in: *Advances in Math.* 26 (1977), pp. 206–222.

- 
- [123] P. Mishra et al., « Does a surface attached globule phase exist ? », in: *Physica A: Statistical Mechanics and its Applications* 318 (2002), p. 171.
- [124] A. A. Mogulskii and C. Pagma, « Superlarge deviations for sums of random variables with arithmetical super-exponential distributions », in: *Siberian Adv. Math.* 18.3 (2008), pp. 185–208.
- [125] A. V. Nagaev, « Integral limit theorems taking into account large deviations when Cramer’s condition does not hold. I », in: *Theory Probab. Appl.* 14.1 (1969), pp. 51–64.
- [126] A. V. Nagaev, « Large deviations of sums of independent random variables », in: *Ann. Probab.* 7.5 (1979), pp. 745–789.
- [127] S. Nakajima, « Concentration results for directed polymer with unbounded jumps », in: *arXiv:1603.05032v1* (2016).
- [128] R. A. Neher and U. Gerland, « Intermediate phase in DNA melting », in: *Phys. Rev. E* 73 (2006).
- [129] P. Ney, « A refinement of the coupling method in renewal theory », in: *Stochastic Processes and their Applications* 11.1 (1981), pp. 11–26, ISSN: 0304-4149.
- [130] G. B. Nguyen and N. Petrelis, « A Variational Formula for the Free Energy of the Partially Directed Polymer Collapse », in: *Journal of Statistical Physics* 151 (2012).
- [131] A. L. Owczarek, T. Prellberg, and R. Brak, « New scaling form for the collapsed polymer phase », in: *Phys. Rev. Lett.* 70 (7 1993), pp. 951–953.
- [132] E. Perfilev and V. Wachtel, *Local tail asymptotics for the joint distribution of length and of maximum of a random walk excursion*, 2019, arXiv: 1907.02578 [math.PR].
- [133] J. A. Plascak, P. H. L. Martins, and M. Bachmann, « Solvent-dependent critical properties of polymer adsorption », in: *Physical Review E* 95.5 (2017), ISSN: 2470-0053.
- [134] J. Poisat, « Ruelle-Perron-Frobenius operator approach to the annealed pinning model with Gaussian long-range correlated disorder », in: *Markov Processes and Related Fields* 19.3 (2013), pp. 577–606.
- [135] D. Poland and H. A. Scheraga, *Theory of helix-coil transitions in biopolymers; statistical mechanical theory of order-disorder transitions in biological macromolecules*, English, Academic Press, 1970, xvii, 797 p.
- [136] C. J. Preston, « A generalization of the FKG inequalities », in: *Comm. Math. Phys.* 36.3 (1974), pp. 233–241.
- [137] R Rajesh et al., « Adsorption and collapse transitions in a linear polymer chain near an attractive wall », in: *Physical review. E, Statistical, nonlinear, and soft matter physics* 65 (2002), p. 056124.



- 
- [138] R. Resnick, *Extreme values, regular variation and point process*, Series in Operations Research and Financial Engineering, New York, Springer, 1987.
- [139] S. I. Resnick, *Heavy-Tail Phenomena, Probabilistic and Statistical Modeling*, Springer-Verlag New York, 2007, ISBN: 978-0-387-45024-7.
- [140] J. S., L. T., and R. A., *Random Graphs*, vol. 45, Wiley Series in Discrete Mathematics and Optimization, John Wiley and Sons, 2011, p. 348, ISBN: 1118030966, 9781118030967.
- [141] T. Seppäläinen, « Scaling for a one-dimensional directed polymer with boundary conditions », in: *Ann. Probab.* 40.1 (2012), pp. 19–73.
- [142] J. Sohler, « Finite size scaling for homogeneous pinning models », in: *Alea* 6 (2009), pp. 163–177.
- [143] J. Sohler, « A functional limit convergence towards brownian excursion. », working paper or preprint, 2010.
- [144] C. Stone, « On local and ratio limit theorems », in: *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob. (Univ. of Calif. Press, 1967)* 2 (1967), pp. 217–224.
- [145] L. Takács, « Limit Distributions for the Bernoulli Meander », in: *Journal of Applied Probabilities* 32.2 (1995), pp. 375–395.
- [146] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Third, Amer. Math. Soc., 2015.
- [147] F. L. Toninelli, « A replica-coupling approach to disordered pinning models », in: *Communications in Mathematical Physics* 280.2 (2008), pp. 389–401.
- [148] F. L. Toninelli, « Localization transition in disordered pinning models. Effect of randomness on the critical properties », in: *Methods of Contemporary Mathematical Statistical Physics, Lecture Notes in Mathematics* 1970 (2009), pp. 129–176.
- [149] F. L. Toninelli, « Disordered pinning models and copolymers: Beyond annealed bounds », in: *Ann. Appl. Probab.* 18.4 (2008), pp. 1569–1587.
- [150] N. Torri, « Pinning Model with Heavy Tailed Disorder », in: *Stochastic Processes and their Applications* 126 (2016), pp. 542–571.
- [151] S. M. Ulam, « Monte Carlo calculations in problems of mathematical physics », in: *Modern mathematics for the engineer: Second series*, McGraw-Hill, New York, 1961, pp. 261–281.
- [152] A. M. Vershik and S. V. Kerov, « Asymptotics of the plancherel measure of the symmetric group and the limiting form of Young tables », in: *Soviet Math. Dokl.* 18 (1977), pp. 527–531.
- [153] R. Wei, « On the long-range directed polymer model », in: *Journal of Statistical Physics* 165.2 (2016), pp. 320–350.

- 
- [154] J. A. Williamson, « Random walks and Riesz kernels », in: *Pacific Journal of mathematics* 25.2 (1968).
- [155] R. Zwanzig and J. Lauritzen, « Exact calculation of the partition function for a model of two dimensional polymer crystallization by chain folding », in: *J. Chem. Phys.* 48.8 (1968), p. 3351.





---

## **Titre : Perturbations de la transition d'adsorption dans des modèles de polymères**

**Mots clés :** Modèles de polymères, transition de phase, adsorption, effondrement, transition de surface, désordre corrélé, pertinence du désordre, régime de désordre intermédiaire

**Resumé :** Cette thèse étudie deux modèles de polymères, chacun présentant un phénomène de transition de phase d'adsorption. Le premier modèle concerne un polymère interagissant avec son solvant et avec un mur dur. Si l'affinité chimique du polymère avec le solvant est suffisamment faible, le polymère s'effondre sur lui-même pour former un globule compact. Dans ce régime, nous étudions l'interaction entre le globule et le mur, et explicitons un phénomène de transition de surface. Nous donnons également un équivalent asymptotique exact de la fonction de partition du polymère effondré dans le cas où il n'y a pas de mur. Le second est une modélisation du phénomène de dénaturation des brins d'ADN, ou plus généralement de deux polymères pouvant s'accrocher entre eux. Nous étudions l'effet que des inhomogénéités dans la composition des polymères ont sur la transition de dénaturation, en terme de (non)-pertinence d'un champ de désordre fortement corrélé. Nous étudions également la limite d'échelle de désordre intermédiaire de ce modèle.

---

## **Title : Perturbations of adsorption transition in polymer models**

**Keywords :** Polymer models, phase transition, adsorption, collapsed polymer, surface transition, correlated disorder, disorder relevance, intermediate disorder regime

**Abstract :** This thesis investigates two polymer models that undergo an adsorption (or pinning) phase transition. The first model is a polymer interacting repulsively with its solvent and attractively with a hard wall. When solvent repulsion is sufficiently strong, the polymer collapses into a compact globule. In this regime we study the interaction between the globule and the wall, and prove that it undergoes a surface transition. We also display an exact equivalent of the partition function for the collapsed polymer when there is no wall. The second model represents the denaturation phenomenon of DNA strands, or more generally two polymers which may pin to each other. We investigate the influence of inhomogeneities in the composition of polymers on this denaturation transition, in terms of disorder (ir)-relevance for a strongly correlated field. We also study the intermediate disorder regime of this model.