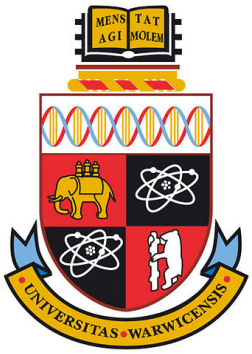


CRITICAL SCALING LIMITS AND SINGULAR SPDES

Simon Gabriel

Thesis submitted for the degree of
Doctor of Philosophy in Mathematics (Research)

University of Warwick
Mathematics Institute

September 2023

Contents

Tables and Illustrated Material iii

Acknowledgements v

Declaration vii

Abstract ix

Symbols xi

1 Introduction 1

1.1 Interacting particle systems . . . 2
1.2 Disordered systems . . . 8
1.3 Singular stochastic PDEs . . . 14
1.4 Outline of the thesis . . . 22

2 Preliminaries 23

2.1 A primer on Wiener chaoses . . . 23
2.1.1 Gaussian Hilbert spaces and white noise . . . 23
2.1.2 The Wiener chaos decomposition . . . 27
2.1.3 Contractions and Feynman diagrams . . . 31
2.2 On rooted trees and the Butcher series . . . 35
2.2.1 Rooted trees . . . 35
2.2.2 Associating trees to nested derivatives . . . 39
2.2.3 The Butcher series . . . 44
2.2.4 A brief note on rough paths . . . 47
2.3 Wild expansion of SPDEs with additive noise . . . 48
2.3.1 Iterated stochastic integrals, trees and Wiener chaoses . 53
2.3.2 From the Wild expansion to a Butcher series . . . 62
2.3.3 A brief note on regularity structures . . . 66

3 The Allen–Cahn equation with critical initial datum 69

3.1 Proof steps for the main result . . . 74
3.1.1 From the Wild expansion to single-tree estimates . . . 75
3.1.2 Proof of Theorem 3.0.1 . . . 79

3.2 Paired trees and their structure 86
3.2.1 Pairings between trees and $L^2(\mathbb{P})$ estimates 87
3.2.2 1-cycles and their removal 89
3.2.3 Existence of v-cycles in paired trees 94
3.2.4 Removal of v-cycles of arbitrary length 99
3.3 Contributing and non-contributing trees 109
3.3.1 Contributing contractions 109
3.3.2 Non-contributing trees 119
3.3.3 Proof of Proposition 3.1.2 126

4 Central limit theorems for the 2D random directed polymer 129
4.1 Background and outline of the proof 134
4.2 Approximating point-to-point partition functions 140
4.2.1 A single jump factorises the term 145
4.2.2 Multiple exceptional jumps are negligible 150
4.3 Decay of remainders in polynomial chaoses 157
4.4 The annealed polymer measure 159
4.5 The quenched polymer measure 166
4.5.1 Finite-dimensional distributions 166
4.5.2 An invariance principle for polymer paths 173
4.5.3 Local limit theorem for the polymer marginals 177

5 Size-biased diffusion limits and the inclusion process 181
5.1 Results and comparison to the literature 185
5.1.1 Scaling limits of the inclusion process 185
5.1.2 A size-biased viewpoint on the PD-diffusion 189
5.1.3 The Trotter–Kurtz approximation theorem 194
5.2 Scaling limits in the case $\mathrm{d}L \to \theta < \infty$ 196
5.2.1 The measure-valued process 196
5.2.2 Equivalence of the superprocess with PD-diffusion . . 206
5.2.3 The advantage of a size-biased evolution 209
5.2.4 Convergence to a Fleming-Viot process 212
5.3 The diffusion limit in the case $\mathrm{d}L \to \infty$ 214
5.3.1 Deriving the diffusion limit 215
5.3.2 Duality and the hydrodynamic limit 220
5.3.3 A natural extension of the PD-diffusion 225
5.4 Discussion and outlook 226

Bibliography 243

Tables and Illustrated Material

List of tables

1.1 Examples of interacting particle systems 3
1.2 Harris criterion for disorder relevance 10
1.3 Classical choices of Gaussian forcing/noise for SPDEs. 14
1.4 Conditions for (scaling) sub-/super-/criticality of SPDEs. . . . 18

2.1 Examples of rooted trees 38

List of figures

1.1 Formation of a Bose-Einstein condensate (by NASA/JPL-Caltech) 4
1.2 Example configuration for the inclusion process 6

4.1 Existence of large jumps for the 2D-DRPM 146
4.2 Existence of a large enough jump for the 2D-DRPM 150

5.1 Size-biased sampling of a particle configuration 183
5.2 Simulation: inclusion process and size-biased jump diffusion . 224
5.3 Condensation in the inclusion process in the case $\mathtt{d}(L)L \to \infty$ 227

Throughout the thesis, various plots of trees have been generated using TikZ for LaTeX. Credit for the codebase is due to Martin Hairer, see e.g. `https://arxiv.org/e-print/2201.03487` under the license CC-BY 4.0.

Acknowledgements

First, I want to express my gratitude towards Nikos Zygouras and Paul Chleboun for their supervision throughout my PhD, and also towards Stefan Grosskinsky for his initial involvement as my supervisor during his last year at Warwick and guidance ever since. Countless discussions and interactions with them, as well as with Giuseppe Cannizzaro and Tommaso Rosati, shaped me as a mathematician. Throughout my PhD, I was supported by the Warwick Mathematics Institute Centre for Doctoral Training, and I acknowledge funding from the University of Warwick and EPSRC through grant EP/R513374/1.

Second, I'd like to thank Giuseppe Cannizzaro, Roger Tribe and Lorenzo Zambotti for agreeing, and taking their time, to examine this thesis.

Throughout my graduate studies, I had the opportunity to spend several research visits abroad, hosted by Stefan Grosskinsky (Augsburg), Rongfeng Sun (Singapore) and Fabio Toninelli (Vienna). I'd like thank them as well as the institutions for their hospitality. Moreover, I am grateful for helpful conversations with Sigurd Assing, Francesco Caravenna, Ajay Chandra, Stefan Junk, Felix Otto and Daniel Ueltschi. A special mention to Hendrik Weber, whose lectures on singular stochastic PDEs, delivered through Oxford's Tought Course Center, had a great impact on my research.

My time at Warwick would not have been the same without my PhD colleagues Andrew, David, Diana, Dimitris, Harry, Pierre, Sotiris and Ziyang. Not just for mathematical discussions, but for making these years memorable.

At last, I am indebted to my friends and family. In particular, to *Faye*.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of **Doctor of Philosophy in Mathematics**. It has been composed by myself and has not been submitted in any previous application for any degree. The material presented (in Chapters 3 – 5, and partially in Chapter 2) is based on (joint) published research:

- [GRZ23] S. Gabriel, T. Rosati, N. Zygouras. The Allen–Cahn equation with weakly critical random initial datum (submitted [after defense]).
- [Gab23] S. Gabriel. Central limit theorems for the (2+1)-dimensional directed polymer in the weak disorder limit. Published in Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.
- [CGG23] P. Chleboun, S. Gabriel, S. Grosskinsky. Size-biased diffusion limits and the inclusion process (submitted).

Abstract

In this thesis, we study scaling limits of idealised models arising in statistical physics. We present three works that, in different directions, explore critical behaviour of such systems.

In the first part, we study the 2D Allen–Cahn equation with white noise initial datum. This differential equation falls into the class of critical singular stochastic partial differential equations (SPDEs), for which no solution theory exists due to the roughness of the data. To give meaning to the equation, we consider the weak coupling scaling and establish non-trivial Gaussian fluctuations of the solution, by treating an infinite perturbative series expansion in terms of iterated stochastic integrals. To our best knowledge, this is the first time such an approach has been implemented for a (non-linear) critical SPDE and possibly a first step towards a general theory of SPDEs in this regime.

Closely connected, the second part of the thesis comprises the large-scale behaviour of the 2D directed random polymer model, describing the trajectory of a random walk in a random potential. In the weak disorder limit, we derive an invariance principle for the polymer paths. As a consequence, the random potential has no effect on large scales, which is due to a self-averaging effect. Similar results were previously only obtained for all but the (critical) two dimensional case.

Last, we study a system of particles with attractive interactions. Clustering of particles occurs when tuning the system's parameters, with growing system size. This describes a simplified model of (Bose–Einstein) condensation. A careful analysis of the infinitesimal dynamics, using the Trotter–Kurtz approximation theorem, allows us to identify the limiting evolution in terms of a measure-valued Markov process. Moreover, we establish a link between the derived dynamics and the infinitely-many-neutral-alleles model in population genetics. The presented approach covers all interesting scaling regimes of the system parameters.

Symbols

$\mathbf{N}, \mathbf{Z}, \mathbf{R}$	Natural numbers, integers, reals
$D(X, \mathbf{R})$	Skorokhod space on X
$C(X, \mathbf{R})$	Continuous functions on X
$\mathcal{M}_1(X)$	Probability measures on X
$\lambda^{(d)}$	Lebesgue measure on $\mathbf{R}^d$
$\xrightarrow{d}$	Weak convergence of probability measures
$(W_t)_{t\geq 0}$	Brownian motion
$p_t^{(\mathfrak{m})}(x)$	Heat kernel (with potential $\mathfrak{m}$) (2.51)
$P_t^{(\mathfrak{m})}$	Heat semigroup (with potential $\mathfrak{m}$) (2.51)

Chapters 2 and 3

ξ	Space-time white noise, Table 1.3
η	Space white noise, Table 1.3
$\star$	Space convolution operator
$*$	Space-time convolution operator
$\mathcal{T}$	Set of finite rooted unordered trees (2.22)
$\mathcal{T}_n$	n-ary trees (2.27)
$\mathcal{T}_{\leq n}$	Sub-n-ary trees (2.26)
$\mathcal{L}(\tau)$	Set of leaves of tree τ
$\mathcal{I}(\tau)$	Set of inner nodes of tree τ
$\mathfrak{o}_\tau$	Root of tree τ
$\ell(\tau)$	Total number of leaves of tree τ
$i(\tau)$	Total number of inner nodes of tree τ
$s(\tau)$	Symmetry factor (2.28)
$\tau!$	Tree factorial (2.31)

$h^{(\tau)}$	Tree/elementary differential with respect to a function h, (2.35)
$[\cdots]$	Grafting operator (2.24)
$\mathscr{T}$	Trimming operator (2.82)
$\mathcal{K}(\tau)$	Set of contractions (2.68)
$\mathcal{Y}(\tau,\tau)$	Set of pairings, see Definition 3.2.3

Chapter 4

$\mathbf{P}_N$	Simple random walk reference measure
$q_n(z)$	Simple random walk transition probability $\mathbf{P}_N(S_n = z)$, $n \leqslant N$
$\overline{\pi}$	Diffusive rescaling and embedding of nearest neighbour paths into $C([0,1],\mathbf{R})$, see (4.2)
ω	Random disorder
β	Disorder strength
$\mathbf{P}_{\beta,N}$	Polymer measure (1.6)
$Z_{\beta,N}$	Partition function (1.7)

Chapter 5

N	Number of particles
L	Number of locations/positions/sites
$\Omega_{L,N}$	Space of particle configurations (1.1)
η,ξ	Particle configurations in $\Omega_{L,N}$
$\mathfrak{L}_{L,N}$	Generator of particle system, (1.2) and (5.1)
$\mathfrak{d}$	Particle diffusivity
$\pi_{L,N}$	Canonical distribution
$\overline{\nabla}$	Kingman simplex (5.2)
E	Space of size-biased probability measures (5.5)

We will write $a_N \sim b_N$ for sequences if $\lim_{N\to\infty} a_N/b_N = 1$. This is not to be confused with $X \sim \mu$, describing a random variable X with law μ. However, the meaning should be clear from the context. Moreover, we denote various constants by C, and sometimes write $C(a)$ to emphasize dependency on a parameter a.

Chapter 1

Introduction

This thesis comprises topics related to scaling limits, particularly focusing on some explicit models arising in statistical physics. The objective of statistical physics (or statistical mechanics) is the explanation of large scale (macroscopic) phenomena occurring in nature, by only postulating interaction rules on a microscopic level. The most classical example of such a scaling limit is Donsker's invariance principle, see for example [MP01, Theorem 5.22]. It states that a simple random walk $(S_n)_{1\leqslant n\leqslant N}$ converges in distribution to a Brownian motion $(W_t)_{t\in[0,1]}$ when diffusively rescaled, i.e.

$$\left(\tfrac{1}{\sqrt{N}}\overline{S}_{tN}\right)_{t\in[0,1]} \xrightarrow{d} (W_t)_{t\in[0,1]}\,, \qquad \text{in } C([0,1],\mathbf{R})\,,$$

with $\overline{S}$ denoting the linear interpolation of S in $C([0,N],\mathbf{R})$. Therefore, macroscopic observables of a particle can be expressed in terms of Brownian motion, assuming its small scale dynamics are those of a simple random walk.

More generally, one aims to explain how non-trivial large-scale behaviour (which sometimes is known from experimental data) emerges from a small-scale interaction rule. For a better understanding of such a large-scale observation and its origin, we want to

- find microscopic dynamics, as simple as possible, that give rise to the desired large-scale behaviour, and
- study the effect of microscopic modifications/perturbations on large-scale observables.

A plethora of models in modern probability theory arose with this goal. In

this thesis, we will address specific models in the following three fields: **Interacting particle systems (IPS), disordered systems** and **stochastic partial differential equations (SPDEs)**. Below, we give a brief introduction into each of the three topics and present our main findings. Throughout the introduction, we will trade in mathematical rigor for a lighter presentation, thus, only presenting slimmed versions of the main theorems. We will point the reader to the full statements of the results whenever this is the case.

1.1 Interacting particle systems

A discrete interacting particle system is a finite system consisting of N particles on L sites indexed by the set Λ, where $|\Lambda| = L$. For simplicity we take $\Lambda = \{1, \dots, L\}$. The space of particle configurations $\eta = (\eta_1, \dots, \eta_L)$, with $\eta_x \in \mathbf{N}_0$, is then given by

$$\Omega_{L,N} := \left\{ \eta \in \mathbf{N}_0^L : \sum_{x=1}^{L} \eta_x = N \right\} , \tag{1.1}$$

with η_x denoting the number of particles located at site x. We equip $\Omega_{L,N}$ with the discrete topology, i.e. the finest possible topology on $\Omega_{L,N}$, with respect to which every function is continuous.

The microscopic dynamics of particles are then described by a continuous-time Markov jump process $(\eta^{(L,N)}(t))_{t \geqslant 0}$ with state space $\Omega_{L,N}$, which is characterised by an infinitesimal generator of the form

$$\mathfrak{L}_{L,N} f(\eta) = \sum_{x,y=1}^{L} p(x,y) u(\eta_x, \eta_y) \left[f(\eta^{xy}) - f(\eta) \right] , \tag{1.2}$$

where f is a bounded function on $\Omega_{L,N}$. The configuration $\eta^{x,y}$ denotes the new configuration η after one particle moved from x to y, i.e. $\eta^{x,y} = \eta + e^y - e^x$, with $(e^x)_z := \mathbb{1}_{x=z}$, provided $\eta_x > 0$. The jump rates $u(n,m) \geq 0$ are given by a non-negative function of the occupation numbers n on the departing site and the occupation number m on the target site. To avoid degeneracies, we furthermore assume that $u(n,m) = 0$ if and only if $n = 0$. On the other hand, $p(x,y)$ denotes irreducible transition rates on $\{1, \dots, L\}$ and models the geometry of the underlying lattice Λ. Notice that the Markov process conserves the total number of particles $\sum_{x=1}^{L} \eta_x = N$.

We are interested in the macroscopic behaviour of these interacting sys-

tems when taking both the total number of particles N and the size of the system L to infinity, while keeping the density of the system approximately constant, i.e. $N, L \to \infty$ such that

$$\lim_{N,L\to\infty} \frac{N}{L} =: \rho > 0\,.$$

This is known as the **thermodynamic limit** and will be abbreviated subsequently by simply writing $\frac{N}{L} \to \rho$.

A large class of models of the form (1.2) have been introduced in the past century, see for example [Spi70, CT85], to describe particle movements in physics or population dynamics in genetics. Typical examples consist of the

exclusion process		$u(n,m) = n(1-m)\,,$
inclusion process	**with**	$u(n,m) = n(\mathtt{d}+m), \mathtt{d} \in [0,\infty)\,,$
zero-range process		$u(n,m) = n\,.$

Table 1.1: Examples of interacting particle systems

In the case of exclusion process, one also restricts the state space to particle configurations η satisfying the exclusion rule, i.e. $\eta_x \in \{0,1\}$ for all $x = 1, \ldots, L$.

The macroscopic observables of interest depend on the particle system at hand, and its modeling purpose. For example, the exclusion process is a simplified model of mass transport and one is interested in the mass distribution when "zooming out". In the thermodynamic limit one then observes a smooth interface which is described by the evolution of a PDE, see for example [KL99, Chapter 4]. On the other hand, both in the inclusion process and zero-range process, the absence of the exclusion rule allows particles to cluster. Hence, when taking the thermodynamic limit, a diverging number of particles may occupy a single site; the reader may think of "sharp peaks" in the macroscopic interface. This phenomenon is known as **condensation** and will be our main interest in the context of IPS.

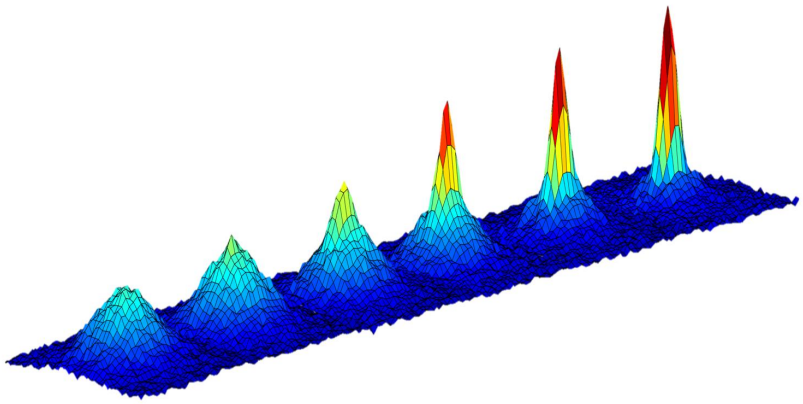


Figure 1.1: Formation of a Bose-Einstein condensate when cooling down a cloud of atoms closer and closer to absolute zero temperature, from left to right. NASA/JPL-Caltech, 2018. (Online). Available from: `https://www.jpl.nasa.gov/images/pia22561-bose-einstein-condensate-graph` (Accessed 8 August 2023)

We want to introduce the notion of condensation rigorously in the setting of IPS. Because the underlying state space $\Omega_{L,N}$ is finite and we assumed the underlying dynamics to be irreducible, there exists a unique invariant distribution $\pi_{L,N} \in \mathcal{M}_1(\Omega_{L,N})$ with respect to $\mathfrak{L}_{L,N}$. We call $\pi_{L,N}$ a **canonical distribution**. For convenience, let us also assume that $\pi_{L,N}$ is spatially homogeneous in the sense that

$$\pi_{L,N}(\eta_x \in \cdot\,) = \pi_{L,N}(\eta_y \in \cdot\,), \qquad \text{for all } \; x, y = 1, \dots, L\,.$$

For example, spatial homogeneity holds if the jump rates u satisfy a detailed balance equation and the transition rates p are symmetric [CG14]. Let us furthermore assume that the weak limit of the single-site marginals of $\pi_{L,N}$ exists for all $\rho \geq 0$, namely

$$\pi_{L,N}(\eta_x \in \cdot\,) \xrightarrow{d} \nu_\rho(\,\cdot\,), \qquad \text{as } \; \frac{N}{L} \to \rho\,, \tag{1.3}$$

for some probability measure ν_ρ on $\mathbf{N}_0$. This convergence is usually implemented in the context of *equivalence of ensembles*, see for example [CGG22, Proposition A.1].

Now, comparing the average particle occupation under $\pi_{L,N}$ and ν_ρ, we

can determine if a system exhibits condensation. First, due to spatial homogeneity and conservation of mass, we have

$$\pi_{L,N}(\eta_x) = \frac{1}{L}\sum_{y=1}^{L} \pi_{L,N}(\eta_y) = \frac{N}{L} \to \rho\,.$$

Here we wrote $\pi_{L,N}(f)$ for the expectation of f under the measure $\pi_{L,N}$. On the other hand, since the identity $f(n) = n$ is an unbounded function, we cannot guarantee that the particle density of the system is conserved in the limit, and $\nu_\rho(\eta_x)$ may be strictly smaller than ρ. Thus, a positive fraction of particles must be clustered in chunks of diverging size, which we shall call the condensate [CG14, JCG19, CGG22].

Definition 1.1.1 (Condensation)**.** *A system characterised by spatially homogeneous canonical distributions $(\pi_{L,N})_{L,N}$ exhibits **condensation** in the thermodynamic limit $\frac{N}{L} \to \rho$, if ν_ρ in (1.3) exists and*

$$\nu_\rho(\eta_x) < \rho = \lim_{N/L\to\rho} \pi_{L,N}(\eta_x)\,.$$

*Furthermore, we say that the system has a **condensation transition** with **critical density** $\rho_c \geq 0$ if*

$$\nu_\rho(\eta_x) \begin{cases} = \rho & \text{if } \rho < \rho_c\,, \\ < \rho & \text{if } \rho > \rho_c\,. \end{cases}$$

We established that the "missing" mass $\rho - \nu_\rho(\eta_x)$ must concentrate on sites with diverging occupation numbers, we call this the **condensed phase**. The total number of such sites has a vanishing volume fraction and does not contribute to the weak limit ν_ρ, describing the distribution of the **background** or **bulk phase**.

After establishing condensation of a given model, one is interested in the finer structure of the condensate. For example, the condensed phase could consist of a single site which accommodates all particles, which is known as **complete condensation**. Alternatively, the condensate may be distributed between various sites, finite of infinite in number. The occurrence of condensation and the statistics of the condensate have been studied for various particle systems, see e.g. [EH05, GL12, CG14]. References for the inclusion process and related models include [WE12, CCG14, CCG15, CGG22] and [BDG17, KS21] in the context of metastable dynamics of the condensate.

Up to now, the condensation phenomenon is only expressed in terms of stationary measures. Studying the limiting dynamics of the condensed phase is a natural next step. However, the condensed phase is only observed macroscopically and cannot be captured using the microscopic state space. Instead, the dynamics imposed by (1.2) have to be translated onto a macroscopic level.

In Chapter 5, we will study the particular case of the inclusion process on a complete graph, i.e. $p(\cdot,\cdot) = 1$, for which (1.2) takes the form

$$\mathfrak{L}_{L,N} f(\eta) = \sum_{\substack{x,y=1 \\ x \neq y}}^{L} \eta_x (\mathtt{d} + \eta_y)[f(\eta^{x,y}) - f(\eta)]\,.$$

The dynamics of the inclusion process consists of two parts. Each particle performs a continuous-time random walk with rate $\mathtt{d}$, and in addition particles attract each other at unit rate.

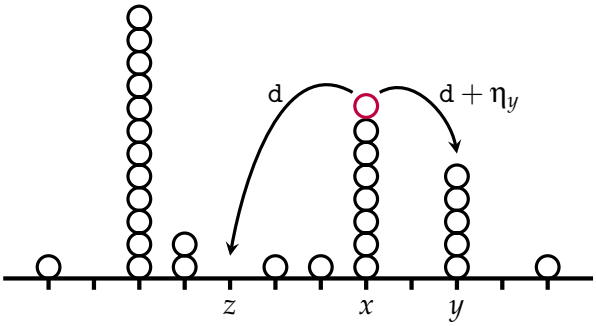


Figure 1.2: An example configuration η for the inclusion process. The purple particle located at x will jump at rate $\mathtt{d} + 5$ onto site y and at rate $\mathtt{d}$ onto site z.

The inclusion process was introduced in [GKR07] as the dual of a model of energy transfer and is the natural counterpart of the exclusion process, since interactions are attractive rather than repulsive. Moreover, it can be interpreted in the context of population genetics, describing a population of Moran type [Mor58], where $\mathtt{d}$ corresponds to the mutation rate and resampling occurs at rate 1. In the thermodynamic limit, condensation of the inclusion process was studied heuristically in [CCG14], with vanishing diffusivity $\mathtt{d} = \mathtt{d}(L) \to 0$. Its finer structure was then analysed rigorously in [JCG19]: At the **critical scaling**

$$\mathtt{d}L \to \theta\,, \qquad \text{for some } \theta \in (0,\infty)\,,$$

the condensate is supported on a diverging number of sites that are each occupied by $O(N)$ many particles. In particular, the authors of [JCG19] were able to identify the limiting distribution of the condensate (under a suitable embedding of particle configurations) as the **Poisson-Dirichlet distribution** with parameter θ. The case $\theta = 0$ corresponds to complete condensation. On the other hand, if $\theta = \infty$, then sites in the condensed phase only have occupation numbers of order $O(\mathtt{d}^{-1})$, i.e. the condensate is located on **mesoscopic scales**. We refer to Chapter 5 for a more detailed review of the results in [JCG19].

Determining the limiting dynamics of the inclusion process, requires a careful analysis of the generator $\mathfrak{L}_{L,N}$ on macroscopic scales. In the remainder of the section, we give a brief outline of the result without going into the technical details, which are provided in Chapter 5.

First, we have to choose an appropriate common state space that describes the macroscopic dynamics of the system. There are various choices, but it will turn out useful to embed particle configurations η into the space of probability measures in a **size-biased** fashion, using the map

$$\eta \mapsto \mu_{\#}\eta := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{\eta_x}{N}} \in E\,, \tag{1.4}$$

where E is the subspace of $\mathcal{M}_1([0,1])$ characterised as the closure of particle configurations under $\mu_{\#}$. With regards to IPS, our first main result determines the dynamics of the inclusion process on macroscopic scales, which in particular describes the condensed phase.

Theorem A ([CGG23]). *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L)$ such that $\mathtt{d}L \to \theta \in [0,\infty)$. If $\eta(0) = \eta^{(L,N)}(0) \in \Omega_{L,N}$ such that $\mu_{\#}\eta(0) \xrightarrow{d} \mu_0 \in E$ as $\frac{N}{L} \to \rho$, then*

$$\big(\mu_{\#}\eta(t)\big)_{t\geq 0} \xrightarrow{d} (\mu_t)_{t\geq 0}\,, \quad \text{in } D([0,\infty), E)\,, \quad \text{as } \frac{N}{L} \to \rho\,.$$

Here $(\mu_t)_{t\geq 0}$ denotes a measure-valued process on E with initial value μ_0, which can be identified by an explicit infinitesimal generator $\mathcal{L}_\theta$.

A detailed version of the result, including the definition of the generator $\mathcal{L}_\theta$, can be found in Theorem 5.1.1. Moreover, we prove that the limiting process derived in Theorem A is equivalent to the infinitely-many-neutral-alleles model, also known as the Poisson-Dirichlet diffusion, with mutation

parameter θ [EK81]. We will introduce this process in Chapter 5.

The above result does not cover the case $\mathtt{d}L \to \infty$. Recall from above that the condensate then concencrates on mesoscopic scales of order $O(\mathtt{d}^{-1})$ [JCG19]. This suggests to consider the embedding of particle configurations η via

$$\eta \mapsto \hat{\mu}_{\#}\eta := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\mathtt{d}L\frac{\eta_x}{N}} \in \mathcal{M}_1(\mathbf{R}_+)\,, \tag{1.5}$$

into the space of probability measures. Because we also want to observe diverging rescaled masses, we consider the statespace $\mathcal{M}_1(\overline{\mathbf{R}}_+)$, with $\overline{\mathbf{R}}_+ = [0,\infty]$. Again, we can derive non-trivial limiting dynamics of the condensate, when slowing down time appropriately.

Theorem B ([CGG23]). *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L) \to 0$ such that $\mathtt{d}L \to \infty$. If $\eta(0) \in \Omega_{L,N}$ such that $\hat{\mu}_{\#}\eta(0) \xrightarrow{d} \hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$, then*

$$\Big(\hat{\mu}_{\#}\eta\big(\tfrac{t}{\mathtt{d}L}\big)\Big)_{t\geq 0} \xrightarrow{d} (\hat{\mu}_t)_{t\geq 0}\,, \qquad \text{in } D([0,\infty), \mathcal{M}_1(\overline{\mathbf{R}}_+))\,, \quad \text{as } \frac{N}{L} \to \rho\,.$$

Here $(\hat{\mu}_t)_{t\geq 0}$ denotes a measure-valued process on $\mathcal{M}(\overline{\mathbf{R}}_+)$ with initial value $\hat{\mu}_0$, whose infinitesimal generator $\hat{\mathcal{L}}$ can be identified explicitly.

A detailed version of the result, including the definition of the generator $\hat{\mathcal{L}}$, can be found in Theorem 5.1.2. Notably, the alternative viewpoint on macroscopic dynamics via (1.4), presented in [CGG23] and this thesis, allows the derivation of a meaningful limit in the case $\theta = \infty$, which can be viewed as a natural boundary case of the infinitely-many-neutral-alleles model with infinite mutation rate $\theta = \infty$.

1.2 Disordered systems

In the previous section, we discussed an IPS for which microscopic particle interactions lead to non-trivial behaviour on macroscopic scales. However, models of this type do not take into account the environment the particles are inhabiting. In nature, the environment usually consists of impurities or a small scale potential, which can be modelled in terms of additional randomness, locally affecting the microscopic dynamics.

One of the prime examples for such a disordered system is the **directed random polymer model (DRPM)**. It describes a random walk whose law is

exponentially tilted by a random environment. The strength of the environment is described by a non-negative parameter $\beta \geq 0$, corresponding to the inverse temperature, which we will refer to as **disorder strength**. Individual models may vary but the most common definition is as follows: Consider the law $\mathbf{P}_N$ of a nearest neighbour random walk of length N starting at the origin. Furthermore, let $\omega = (\omega_{n,z})_{(n,z)\in\mathbb{N}\times\mathbb{Z}^d}$ be a family of mean-zero and unit variance random variables with law $\mathbb{P}$ (independent of $\mathbf{P}_N$) with sufficiently many moments. For a fixed realisation of ω, the directed polymer measure of length N and disorder strength $\beta \geq 0$ is then defined using the following change of measure

$$\mathbf{P}^{\omega}_{\beta,N}(\,\mathrm{d}S) := \frac{1}{Z_{\beta,N}} \exp\left(\sum_{n=1}^{N}(\beta\,\omega_{n,S_n} - \lambda(\beta))\right)\mathbf{P}_N(\,\mathrm{d}S), \tag{1.6}$$

where $\lambda(\beta)$ is a positive constant, such that $Z_{\beta,N}$ has unit mean. The denominator

$$Z_{\beta,N} := \mathbf{E}_N\left[\exp\left(\sum_{n=1}^{N}(\beta\,\omega_{n,S_n} - \lambda(\beta))\right)\right] \tag{1.7}$$

is a (random) normalising constant, called the **partition function**, making $\mathbf{P}^{\omega}_{\beta,N}$ a probability measure. The original model goes back to the physics literature [HH85] where directed random polymers were introduced to study the interface in two-dimensional Ising models with random interactions. Subsequently, the model was studied by the mathematical community [IS88, Bol89] and attracted attention because of its application to SPDEs, see for example [BC95, BG97] and the discussion at the end of this section. But even on its own, the DRPM remains an interesting mathematical model. We refer to [CSY04, Com17] and references therein for an overview of the polymer literature.

The weak disorder limit

When studying disordered systems, one is interested whether the implementation of disorder has an effect on the large-scale behaviour of the model. If a small amount of disorder does not change the statistical properties of the system, we say the model is **disorder irrelevant**. On the other hand, the model is **disorder relevant**, if an arbitrary weak disorder does indeed change the model's statistics.

In the 1970's Harris proposed a heuristic criterion, formulated in the con-

text of a ferromagnetic Ising model with random impurities, to determine if the model is disorder (ir)relevant [Har74]. The proposed heuristic can be extended to a larger class of models. It compares the effective dimension $d_{\text{eff}} := d + 2$ of space-time to **a** correlation length exponent ν (for the simple random walk $\nu = \frac{1}{2}$) of the underlying pure dynamics. The heuristic yields the following classification:

	$\nu > \frac{2}{d_{\text{eff}}}$		disorder irrelevant
if	$\nu = \frac{2}{d_{\text{eff}}}$	**then**	marginal (inconclusive)
	$\nu < \frac{2}{d_{\text{eff}}}$		disorder relevant

Table 1.2: Harris criterion for disorder relevance

In the case $\nu = \frac{2}{d_{\text{eff}}}$, the Harris criterion is inconclusive and the model is said to be marginal, meaning that the large-scale behaviour depends on the finer structure of the model. We then speak of either **marginal relevance** or **marginal irrelevance** of the disorder. See also [CSZ17a] for an alternative characterisation of disorder relevance, which agrees with Harris heuristic. According to Harris criterion, the DRPM is disorder relevant if $d = 1$ and disorder irrelevant if $d \geqslant 3$, while the case $d = 2$ remains marginal.

For the DRPM specifically, it is more natural to identify the effect of the disorder in terms of the partition function. It was shown in [Bol89], that the partition function's large N-limit is either positive or equal zero $\mathbb{P}$-almost surely. On the other hand, it is easy to see that $Z_{0,N} = 1$ for the pure model ($\beta = 0$). Hence, we expect the partition function to vanish if the disorder is relevant and remain positive if the disorder is irrelevant. Clearly, once β is large enough such that the disorder is relevant, this should continue to hold for any larger β [CY06]: For arbitrary dimension d there exists a $\beta_c = \beta_c(d)$ such that $\mathbb{P}$-almost surely

$$\lim_{N\to\infty} Z_{\beta,N} \begin{cases} > 0 & \text{if } \beta \in \{0\} \cup (0, \beta_c), \\ = 0 & \text{if } \beta > \beta_c. \end{cases} \tag{1.8}$$

The subcritical phase is referred to as **weak disorder regime**, whereas the supercritical phase is known as the **strong disorder regime**. In particular, it was established that $\beta_c = 0$ when $d \leq 2$. Notably, this agrees with the prediction of Harris criterion.

The strong disorder regime in $d \leqslant 2$ remains difficult to study due to

the degeneracy of the partition function. Instead, it is common practice to interpolate between the pure model ($\beta = 0$) and the strong disorder regime ($\beta > 0$) by tuning the disorder strength $\beta = \hat{\beta} \cdot o(1)$, $\hat{\beta} \in (0, \infty)$, with the size of the system. This is known as the **intermediate disorder regime** [AKQ14b, CSZ17a]. We should then see a non-trivial limit of $Z_{\beta,N}$ in $d \leq 2$, if β is rescaled appropriately as a function of N. For $d = 1$, the limiting partition function is strictly positive when $\beta = \beta_N \sim \widehat{\beta} N^{-1/4}$ for arbitrary $\widehat{\beta} > 0$, see [AKQ14b]. The case of $d = 2$ is special: After rescaling

$$\beta = \beta_N \sim \widehat{\beta} \sqrt{\frac{\pi}{\log N}}, \qquad \text{with } \widehat{\beta} > 0, \tag{1.9}$$

we see a phase transition in $\widehat{\beta}$ [CSZ17b]. More precisely, the limit $\lim_{N\to\infty} Z_{\beta_N,N}$ is strictly positive if $\widehat{\beta} \in (0, 1)$ and vanishes if $\widehat{\beta} \geq 1$. The log-scaling of β_N, but not the corresponding phase transition in $\widehat{\beta}$, was observed in [BC98]. In the one-dimensional case no such phase transition exists.

The concept of considering polymers when the disorder strength is scaled as a function of N appeared already earlier in the physics literature [BD00, CDR10]. We will denote the appropriate rescaling of the disorder strength with the size of the system as **weak disorder limit**.

Despite $\lim_{N\to\infty} Z_{\beta_N,N}$ being of order one, the limiting random variable has non-trivial fluctuations and it remains to investigate if these fluctuations influence the large scale behaviour of the polymer. With regards to the DRPM, this thesis main result is the identification of the limiting polymer path distribution for $d = 2$, when taking the weak disorder limit.

Theorem C ([Gab23])**.** *Let $\widehat{\beta} \in (0, 1)$ and β_N be as in (1.9), and assume that the given disorder is either bounded or normal.*[1] *Then*

$$\overline{\pi}_{\#} \mathbf{P}^{\omega}_{\beta_N,N} \xrightarrow{d} \mathbf{P}\big(\tfrac{1}{\sqrt{2}} W \in \cdot\big), \qquad \text{as } N \to \infty, \quad \text{in } \mathbb{P}\text{-probability},$$

where $\mathbf{P}$ denotes the Wiener measure on $C([0, 1], \mathbf{R}^2)$. Here $\overline{\pi}_{\#} \mathbf{P}^{\omega}_{\beta_N,N}$ denotes the pushforward measure of polymer paths on the space of continuous functions, using a diffusive rescaling of space-time and linear interpolation.

Notably, we see that the non-trivial fluctuations of $\lim_{N\to\infty} Z_{\beta_N,N}$ have no influence on the large-scale statistics of the polymer, underlining the *marginal*

[1] We will prove Theorem C for a more general class of environments (Theorem 4.0.1), however, for the sake of presentation we simplified the statement here.

irrelevance in the regime considered. For comparision, let us point out that for $d \geqslant 3$ the limiting law of the polymer paths (in the weak disorder regime) was identified as the law of Brownian motion with dimension-dependent diffusion matrix [AZ96, SZ96, CY06]. On the other hand, when $d = 1$ and considering the weak disorder limit, the polymer measure is singular with respect to the Wiener measure [AKQ14a]. The two dimensional case remained open for the past years, Theorem C closes this gap in the literature.

The multiplicative stochastic heat equation

Besides being an interesting model by itself, the DRPM is closely related to a particular SPDE, namely the **multiplicative stochastic heat equation (mSHE)**. This random PDE is described in terms of an evolution equation on $\mathbf{R}^d$:

$$\partial_t u = \frac{1}{2}\Delta u + \beta\, u \cdot \xi\,, \quad u(0,\cdot) = 1\,, \tag{1.10}$$

where ξ denotes space-time **white noise** on $\mathbf{R} \times \mathbf{R}^d$, which is a zero-mean Gaussian distribution (generalised function) ξ with covariance structure

$$\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(t-s)\delta(x-y) \qquad \text{for all} \quad (t,x),(s,y) \in \mathbf{R} \times \mathbf{R}^d\,. \tag{1.11}$$

Because white noise is a random distribution, pointwise evaluation for its covariance does not make sense, but only holds on the level of a formal description. We will give a rigorous definition of white noise in Section 2.1 below.

The relation of mSHE (1.10) with the DRPM is most traceable when representing its solution by a Feynman-Kac formula [KS91, Theorem 4.2]. Disregarding the roughness of ξ for a moment, and instead treating it as a smooth function, one could write

$$u(t,x) = \mathbf{E}_x\left[\exp\left(\beta \int_0^t \xi(t-s, W_s)\,\mathrm{d}s\right)\right]\,, \tag{1.12}$$

with $\mathbf{E}_x$ being the expectation with respect to a Brownian motion $(W_s)_{s\geqslant 0}$ started in x. This representation indeed reminds of the partition function of the DRPM (1.7). However, the stochastic integral in the exponent is not well defined when $d \neq 1$. For $d = 1$ the integral can be defined in an Itô–sense, since the heat kernel is L^2-integrable in space-time, see for example

[BCJL94] or Section 2.1 below. But for $d \geqslant 2$ this argument breaks down. This corresponds to (1.10) being a **singular** SPDE when $d \geqslant 2$, a concept we introduce in the next section.

The connection between (1.12) and the partition function (1.7) can be made more precise: Instead of working with ξ directly, we consider a space smoothened version ξ^ε defined as

$$\xi^\varepsilon(t,x) := (\xi \star j_\varepsilon)(t,x) = \int_{\mathbf{R}^d} j_\varepsilon(y-x)\xi(t,y)\,\mathrm{d}y\,, \tag{1.13}$$

where $j_\varepsilon = \varepsilon^{-d} j(\cdot/\varepsilon)$ for some symmetric $j \in C_c^\infty(\mathbf{R}^d)$. Now, consider the solution of the mollified mSHE

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \beta\, u_\varepsilon \cdot \xi^\varepsilon\,, \qquad u_\varepsilon(0,\cdot) = 1\,, \tag{1.14}$$

which can be represented in terms of

$$u_\varepsilon(t,x) = \mathbf{E}_x\left[\exp\left(\beta\int_0^t \xi^\varepsilon(t-s,W_s)\,\mathrm{d}s - \frac{1}{2}\beta^2\|j\|_{L^2}^2\,\frac{t}{\varepsilon^d}\right)\right]\,, \tag{1.15}$$

using a generalised Feynman-Kac formula [BC95]. The additional correction term in the exponent is due to the roughness of ξ^ε in time. Again, the stochastic integral in (1.15) should be read with care: Despite space-mollification, ξ^ε can not be evaluated pointwise in time. However, the stochastic integral is well-defined in an Itô–sense. Thus, for fixed $t > 0$, the solution of the mollified mSHE (1.14) is the partition function (up to a time-reversal) of a *continuous directed random polymer model* started from x

$$\mathbf{P}^\xi_{\beta,\varepsilon,x}(\,\mathrm{d}X) := \frac{1}{u_\varepsilon(t,x)}\exp\left(\beta\int_0^t \bar{\xi}^\varepsilon(s,X_s)ds - \frac{1}{2}\beta^2\|j\|_{L^2}^2\,\frac{t}{\varepsilon^d}\right)\mathbf{P}_x(\,\mathrm{d}X), \tag{1.16}$$

with $\bar{\xi}^\varepsilon(s,y) = \xi^\varepsilon(t-s,y) \overset{d}{=} \xi^\varepsilon(s,y)$ and $\mathbf{P}_x$ the law of Brownian motion started at x. For a detailed comparison of $u_\varepsilon(t,x)$ and $Z_{\beta,N}$, we refer the reader to [CSZ17b].

Similar to the discrete 2D-DRPM, studying non-trivial scaling limits of (1.14) for $d = 2$ requires to choose $\beta = \beta_\varepsilon$ appropriately. This is known as weak coupling, which we will introduce below.

1.3 Singular stochastic PDEs

In the previous section, we got a first glimpse of a singular SPDE, namely the mSHE (1.10), which cannot be made sense of when $d \geqslant 2$. Subsequently, we will discuss such SPDEs in more detail. Generally, we are interested in stochastic evolution equations of the form

$$\partial_t u = \frac{1}{2}\Delta u + H(u, \nabla u) + G(u)\Xi\,, \quad u(0,\cdot) = u_0\,, \tag{1.17}$$

with G, H nice enough functions, say analytic, and Ξ a placeholder for some Gaussian noise. Throughout, $\Xi(t,x)$ represents either of the following three choices:

space-time white noise	$\xi(t,x)$	$\mathcal{C}^{-\frac{d}{2}-1-}(\mathbf{R}\times\mathbf{R}^d)$
space white noise	$\eta(x)$	$\mathcal{C}^{-\frac{d}{2}-}(\mathbf{R}^d)$
space white noise initial datum	$\delta_0(t)\,\eta(x)$	$\mathcal{C}^{-\frac{d}{2}-2-}(\mathbf{R}\times\mathbf{R}^d)$

Table 1.3: Classical choices of Gaussian forcing/noise for SPDEs.

Recall (1.11) for a formal description of space-time white noise ξ. Similarly, η is a random Gaussian distribution with covariance $\mathbb{E}[\eta(x)\eta(y)] = \delta(x-y)$. For a rigorous definition of white noise, we refer the reader to Section 2.1. Moreover, the table includes each noise's Besov regularity, which we introduce in (1.21) below.

SPDEs of the form (1.17) are expected to describe a wide range of physical phenomena such as interface growth, particle diffusion or dynamical Euclidean quantum field theories, where the random noise Ξ plays the role of impurities or small scale randomness. However, the roughness of the noise makes the study of (1.17) difficult, since certain terms in the equation cannot necessarily be made sense of. Whenever this is the case we call the SPDE **singular**. Thus, even giving meaning to expressions like (1.17) can be difficult. For the most parts, this makes the analysis of (singular) SPDEs challenging, but also mathematically intriguing. For an overview on singular SPDEs, see [CW17, Gub18, CS19].

For the moment, let us consider space-time white noise $\Xi = \xi$ and return to some explicit examples. For $G(u) = \beta\, u$, $H = 0$, we recover the mSHE (1.10). In fact, it will be more natural to study the fluctuations of the solution of (1.10) around the stable, constant field 1 (the initial datum), given by

$\overline{u}(t,x) := u(t,x) - 1$ which solves the

$$\textbf{(recentred) mSHE:} \qquad \partial_t \overline{u} = \frac{1}{2}\Delta \overline{u} + \beta\, \overline{u} \cdot \xi + \beta\, \xi\,, \tag{1.18}$$

with vanishing initial datum $\overline{u}(0,\cdot) = 0$. Other examples of commonly studied SPDEs are the

$$\textbf{KPZ equation:} \qquad \partial_t h = \frac{1}{2}\Delta h + \frac{1}{2}|\nabla h|^2 + \beta\, \xi\,, \tag{1.19}$$

$$\textbf{dynamic } \varphi^4\textbf{-model:} \qquad \partial_t \varphi = \frac{1}{2}\Delta \varphi - \mathfrak{m}^2\, \varphi - \varphi^3 + \xi\,, \tag{1.20}$$

with vanishing initial data $h(0,\cdot) = \varphi(0,\cdot) = 0$ and $\mathfrak{m} \in \mathbf{R}$. At this point, it will be useful to keep track of various objects' regularity or roughness, in order to classify when a SPDE is singular. Therefore, let us introduce (rather informally) the notion of Besov spaces $\mathcal{C}^\alpha$, $\alpha \in (-\infty, 1) \setminus \{0\}$, which "measures" the regularity of a distribution. More precisely, $\mathcal{C}^\alpha(\mathbf{R} \times \mathbf{R}^d)$ contains distributions $F \in \mathcal{S}'(\mathbf{R} \times \mathbf{R}^d)$ that satisfy for any compact set $\mathrm{K} \subset \mathbf{R} \times \mathbf{R}^d$

$$\sup_{(t,x)\in \mathrm{K}} \sup_{\lambda \in (0,1]} \frac{\big|F\big(\varphi^{(\lambda)}_{t,x}\big)\big|}{\lambda^\alpha} < \infty\,, \tag{1.21}$$

where $\varphi \in C_c^\infty(\mathbf{R} \times \mathbf{R}^d)$ with $1(\varphi) \neq 0$, and

$$\varphi^{(\lambda)}_{t,x}(s,y) := \lambda^{-d-2}\big(\lambda^{-2}(s-t), \lambda^{-1}(y-x)\big)\,,$$

which approximates a Dirac-δ at (t,x) as $\lambda \to 0$. In other words, the regularity α of a distribution $F \in \mathcal{C}^\alpha$ denotes the exponent of the worst divergence when trying to evaluate F pointwise. If $\alpha \in (0,1)$ the elements are actually functions and $\mathcal{C}^\alpha$ agrees with the usual Hölder spaces. Instead of considering a time-space domain, one can similarly define Besov spaces $\mathcal{C}^\alpha(\mathbf{R}^d)$. For a rigorous introduction of Besov spaces we refer the reader to [CW17, Definition 2.3] and [CZ21, Definition 12.1].

When considering random distributions, such as white noise, we require (1.21) to hold in expectation. In Table 1.3, we summarised in which Besov spaces the various choices of Ξ lie in. There, $\mathcal{C}^{\alpha-}$ should be read as $\mathcal{C}^{\alpha-\delta}$ for any $\delta > 0$.

Having introduced a way to keep track of each term's regularity, we can

be slightly more detailed on why and when a SPDE is singular. As a first step, we look at a very crude approximation of (1.17) by considering the **additive stochastic heat equation (aSHE)**

$$\partial_t \overline{v} = \frac{1}{2}\Delta \overline{v} + \Xi\,, \qquad \overline{v}(0,\cdot) = u_0\,. \tag{1.22}$$

Its solution $\overline{v}$ is a well-defined random distribution satisfying

$$\overline{v} = \left(\partial_t - \tfrac{1}{2}\Delta\right)^{-1}\Xi \in \mathcal{C}^{\alpha+2-}(\mathbf{R}\times\mathbf{R}^d)\,,$$

by a Schauder-like estimate [CW17, Theorem 2.8], i.e. space-time convolution with the heat kernel increases regularity by 2 (the reader may think of inverting $(-\Delta)$ and taking two anti-deriviatives). Here α corresponds to the Besov regularity of the noise Ξ, cf. Table 1.3. In the case of space-time white noise, i.e. $\Xi = \xi$, we conclude $\overline{v} \in \mathcal{C}^{-\frac{d}{2}+1-}(\mathbf{R}\times\mathbf{R}^d)$, which is a function if $d = 1$ and a distribution otherwise.

Now, it is not too far fetched that the solution of (1.17) will (at best) have the same regularity as $\overline{v}$: Assume for simplicity that $H(u,\nabla u) = H(u)$, then a 0^{th}-order Taylor approximation of both G and H yields the formulation

$$\partial_t v = \frac{1}{2}\Delta v + H(0)\,v + G(1)\,\Xi\,, \qquad v(0,\cdot) = u_0\,, \tag{1.23}$$

which again can be explicitly solved and has the same regularity as $\overline{v}$. The solution of (1.23) is the first term in a Picard-like iteration of (1.17), exploiting the analyticity of both G and H. Therefore, for any of the three SPDEs mentioned above, we can assume that their solution lies in the same regularity class as $\overline{v}$, i.e. $\mathcal{C}^{-\frac{d}{2}+1-}(\mathbf{R}\times\mathbf{R}^d)$. This leads to the following problems when stating the corresponding dynamics: Considering the

- **mSHE** (1.10), the product $u\cdot\xi$ does not make sense. For $d \geqslant 2$, it is ill-defined as a product of two distributions. Even in $d = 1$, the product cannot be made sense of classically, however, it can be interpreted in an Itô-sense [BCJL94].

- **KPZ equation** (1.19), the components of ∇h lie in $\mathcal{C}^{-\frac{d}{2}-}(\mathbf{R}\times\mathbf{R}^d)$ and are distributions. Hence, $|\nabla h|^2$ is not well defined for any $d \geqslant 1$.

- **dynamic φ^4-model** (1.20), the cubic non-linearity does not make sense when $d \geqslant 2$, since φ is a distribution.

Overall, we conclude that (1.17) is generally ill-posed for d large enough, and even giving sense to the equation as stated is a non-trivial task. In the past decade, multiple breakthroughs in this direction sparked the interest of a large mathematical community, such as the concepts of regularity structures [Hai14], energy solutions [GJ14], paracontrolled distributions [GIP15] and renormalisation groups [Kup16]. At this day, there have been many works analysing singular SPDEs, building on the above mentioned techniques and push them to their boundaries. However, all approaches mentioned have the disadvantage to only apply in the so called **(scaling) subcritical regime**, which allows to treat equations with a noise that is not "too rough".

Scaling criticality

Despite singular SPDEs being a priori ill-defined, there are heuristic arguments to guess their small and large-scale behaviour. In the following, we give meaning to the so called **scaling (sub)critical** regime, while closely following [CW17], see also [Hai14, Assumption 8.3].

Under the assumption that the general SPDE (1.17) has a solution u, let us define

$$u^\lambda(t,x) := \lambda^{-\alpha} u(\lambda^2 t, \lambda x)\,, \qquad \text{for some } \lambda > 0\,, \tag{1.24}$$

describing the small-scale ($\lambda \ll 1$) and large-scale ($\lambda \gg 1$) behaviour of u when choosing α appropriately.[2] Since u admits the SPDE (1.17), also u^λ will solve a SPDE with slightly altered coefficients

$$\partial_t u^\lambda = \frac{1}{2}\Delta u^\lambda + \lambda^{-\alpha+2} H(\lambda^\alpha u^\lambda, \lambda^{\alpha-1}\nabla u^\lambda) + \lambda^{-\alpha+2} G(\lambda^\alpha u^\lambda)\Xi(\lambda^2\cdot, \lambda\cdot)\,, \tag{1.25}$$

and $u^\lambda(0,\cdot) = \lambda^{-\alpha} u_0(\lambda\,\cdot)$. In the case of our main examples, which are driven by space-time white noise ξ, we additionally make use of the scaling invariance

$$\tilde{\xi}(t,x) := \lambda^{\frac{d}{2}+1}\xi(\lambda^2 t, \lambda x) \overset{d}{=} \xi(t,x)\,, \tag{1.26}$$

i.e. $\tilde{\xi}$ is also a space-time white noise. Because we want to keep the noise invariant on both small and large scales, (1.26) in combination with (1.25) suggests to choose α as $\alpha = -\frac{d}{2} + 1$. Moreover, with this choice of α we expect u^λ to remain of order one, since the solution u lies in $\mathcal{C}^{-\frac{d}{2}+1-}(\mathbf{R}\times\mathbf{R}^d)$.

[2] Again, we stress that the pointwise evaluation is mathematically not correct, as u will be a distribution in most cases.

Thus, for $\lambda \ll 1$ (1.21) formally yields

$$|u^\lambda(t,x)| = \frac{|u(\lambda^2 t, \lambda x)|}{\lambda^{-\frac{d}{2}+1}} < \infty\,.$$

Overall, we have derived the following list of scaled equations:

$$\begin{aligned}
\textbf{(recentred) mSHE:} \qquad & \partial_t \overline{u}^\lambda = \frac{1}{2}\Delta \overline{u}^\lambda + \beta\, \lambda^{1-\frac{d}{2}}\, \overline{u}^\lambda \cdot \tilde{\xi} + \beta \tilde{\xi}\,, \\
\textbf{KPZ:} \qquad & \partial_t h^\lambda = \frac{1}{2}\Delta h^\lambda + \lambda^{1-\frac{d}{2}} |\nabla h^\lambda|^2 + \beta \tilde{\xi}\,, \\
\textbf{dynamic } \varphi^4\textbf{:} \qquad & \partial_t \varphi^\lambda = \frac{1}{2}\Delta \varphi^\lambda - \lambda^2 \mathfrak{m}^2 \varphi^\lambda - \lambda^{4-d}(\varphi^\lambda)^3 + \tilde{\xi}\,,
\end{aligned} \tag{1.27}$$

with $\overline{u}^\lambda(0,\cdot) = h^\lambda(0,\cdot) = \varphi^\lambda(0,\cdot) = 0$. Notably, in all examples above, the coefficients in front of the ill-defined terms depend only on the spatial dimension d. Hence, subject to the choice of d, the coefficient will vanish, diverge or stay of order one, as $\lambda \to 0$. This leads to the following classification of SPDEs with respect to their small-scale behaviour:

	vanishes		subcritical
if the coefficient	remains constant	**then SPDE is**	critical
	diverges		supercritical

Table 1.4: Conditions for (scaling) sub-/super-/criticality of SPDEs.

In particular, both (recentred) mSHE (1.18) and the KPZ equation (1.19) are scaling critical for $d = 2$, and the dynamic φ^4-model (1.20) is critical when $d = 4$. In all cases, the dimensions below this threshold correspond to the subcritical regime, whereas larger dimensions correspond to the supercritical regime. Hence, on a heuristic level, in subcritical dimensions the effect of the non-linearity can be ignored on small scales. Therefore, the small scale behaviour of the KPZ equation (1.19), $d = 1$, and the dynamic φ^4-model (1.20), $d \leqslant 3$, is described by the linearised equation (1.23). Similarly, for mSHE (1.18) the singular product should have no effect on small scales when $d = 1$.

On the other hand, the large scale behaviour is reversed: We expect the (recentred) mSHE (1.18) to be described by the additive stochastic heat equation when $d \geqslant 3$, since $\lambda^{1-\frac{d}{2}} \to 0$ as $\lambda \to \infty$. Note that this is in accordance with the disorder (ir)relevance of the DRPM and Harris criterion,

which we discussed in Section 1.2, since the corresponding fluctuations are not visible on large scales. To see this in more detail, one should repeat the above steps for $\overline{u}$ (or u as in (1.10)) with $\alpha = 0$ in (1.24), in which case both the multiplicative and additive noise in (1.18) vanish as λ diverges.

Preceeding the introduction of scaling criticality, we mentioned recent solution theories for singular SPDEs [Hai14, GJ14, GIP15, Kup16]. All such theories restrict to subcritical SPDEs. In particular, [Hai14, GIP15] write the solution of a given SPDE in terms of a perturbative expansion around the linearised equation, e.g. around aSHE (1.22) in the cases of KPZ (1.19) and the dynamic φ^4-model (1.20). At this point, the reader may think about the perturbative expansion in terms of an abstract Taylor expansion, around the solution of the linearised equation (1.23). To control the inflicted error, they require (finitely many) higher order terms of this series expansion. We refer to Section 2.3.3 for a brief outline of this method. In turn, this argument breaks down when considering a (scaling) critical SPDEs, since finitely many terms will **not** suffice in the analysis. Following a similar perturbative approach in the critical dimension, requires the control of the whole series.

The weak coupling scaling

The treatment of scaling critical SPDEs, in contrast to subcritical equations, is more delicate. Even when considering (1.17) with a mollified noise Ξ^ε, it is not possible to guarantee convergence in the small ε-limit. Instead, it is common practice to introduce an artificial scaling of the noise (or the nonlinearity) to zero, as the mollification is turned off $\varepsilon \to 0$. The idea, similar to the weak disorder limit for DRPM, is to interpolate between the linearised equation and the actual SPDE (1.17), while still observing non-trivial behaviour. This is known as the **weak coupling scaling**, the limit is referred to as **the weak coupling limit**.

Contrary to subcritical SPDEs, there is no general solution theory for critical equations. Instead, critical SPDEs in the weak coupling regime have only been treated on a case-by-case basis. Let us give a brief overview of the literature, starting with the easiest critical SPDE, namely the 2D-mSHE

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \beta_\varepsilon\, u_\varepsilon \cdot \xi^\varepsilon\,, \quad u_\varepsilon(0,\cdot) = 1\,, \tag{1.28}$$

with the choice $\beta_\varepsilon = \hat{\beta}(\log\frac{1}{\varepsilon})^{-\frac{1}{2}}$, $\hat{\beta} > 0$. Here ξ^ε denotes a mollified white

noise, cf. (1.13). The linearity of the equation (1.28) allows for an explicit treatment via an infinite expansion. This was successfully implemented in [CSZ17b] by approximating the solution of (1.28) with the partition function of the 2D-DRPM, and making use of a polynomial chaos expansion in combination with the fourth-moment theorem. A non-linear mSHE has been studied in [DG22, Tao22]. Solutions of the mSHE (1.28) and the KPZ equation

$$\partial_t h_\varepsilon = \frac{1}{2}\Delta h_\varepsilon + \frac{1}{2}|\nabla h_\varepsilon|^2 + \beta_\varepsilon\, \xi^\varepsilon - C_\varepsilon\,, \quad h_\varepsilon(0,\cdot) = 0\,, \tag{1.29}$$

with C_ε being suitable constants such that $C_\varepsilon \to \infty$, are formally related by the Cole–Hopf transformation

$$h_\varepsilon = \log u_\varepsilon\,.$$

This relation was used in [CD20, Gu19, CSZ20] to express fluctuations of the 2D-KPZ equation in terms of the mSHE. At this point, both the mSHE (1.28) and the KPZ equation (1.29) can only by analysed for $\hat\beta$ below a critical point, thus, limiting the interpolation between the linear model and the strongly coupled model to merely one side of the spectrum. In turn, [CSZ23] analyses the 2D-mSHE at the critical point, which lead to the introduction of **the critical stochastic heat flow**. Lastly, let us mention some examples that were treated without linearisation of the equation. Instead, stationary fluctuations of certain SPDEs with an explicit Gaussian invariant measure can be determined by analysing the corresponding (linear) infinite-dimensional generator. This has been successfully implemented for the anisotropic KPZ equation [CES21, CET21], where $|\nabla h|^2$ in (1.19) is replaced by $(\partial_1 h)^2 - (\partial_2 h)^2$, and the stochastic Burgers equation [CGT23].

Remarkably, weakly coupled critical SPDEs usually exhibit Gaussian fluctuations. More precisely, their weak coupling limit can be represented in terms of the linear Edwards-Wilkinson equation

$$\partial_t v = \frac{\gamma}{2}\Delta v + \mathfrak{m}\, v + \sigma\, \Xi\,, \qquad v(0,\cdot) = 0\,, \tag{1.30}$$

for some $\mathfrak{m}, \sigma, \gamma \in \mathbf{R}$. At this point, the reader might suspect that a vanishing coefficient, such as β_ε in front of the noise in (1.28) and (1.29), trivialises the study of weakly coupled SPDEs and fluctuations are simply given by the equation where the non-linearity was dropped, cf. (1.22) and (1.23). However, this is not the case: In all examples mentioned, **the non-linearity**

is relevant and contributes to the limit in terms of altering the coefficients $\mathfrak{m}, \sigma, \gamma$ in (1.30). Thus, when studying the weak coupling limit, we are mainly interested in determining the exact coefficients $\mathfrak{m}_{\hat{\beta}}, \sigma_{\hat{\beta}}, \gamma_{\hat{\beta}} \in \mathbf{R}$ as functions of the scaling parameter $\hat{\beta}$.

The study of critical SPDEs in the weak coupling limit remains a challenging task and their treatment vastly differs from solution theories regarding subcritical SPDEs. Nevertheless, expansions in the fashion of [Hai14, GIP15] remain a natural approach, despite the theories themselves not being applicable in the case of scaling criticality. With regards to critical singular SPDEs, the main result of this thesis is the treatment of a weakly coupled SPDE following a Picard-like expansion. Notably, this requires the control and analysis of infinitely many stochastic terms. To the author's best knowledge, this is the first time the infinite expansion of a (non-linear) SPDE has been treated.

In detail, we consider the weak coupling limit of the **2D Allen–Cahn equation** with random critical forcing at zero-time:

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \mathfrak{m}\, u_\varepsilon - u_\varepsilon^3 + \frac{\hat{\lambda}}{\sqrt{\log \varepsilon^{-1}}}\delta_0 \eta^\varepsilon\,, \qquad u_\varepsilon(0,\cdot) = 0\,, \tag{1.31}$$

for some $\mathfrak{m} \in \mathbf{R}$. Here $\delta_0\eta^\varepsilon$ denotes a mollified space white noise forcing at $t = 0$, cf. Table 1.3. As a consequence of the additive forcing being scaled to zero (rather than the non-linearity), we have $u_\varepsilon \to 0$. Thus, our focus lies on the non-trivial fluctuations around the trivial solution.

Theorem D ([GRZ23]). *Let $\mathfrak{m} \in \mathbf{R}$. Then there exists a $\hat{\lambda}_{\mathrm{fin}} > 0$ such that for $\hat{\lambda} \in (0, \hat{\lambda}_{\mathrm{fin}})$ and $T \in (0, \infty)$ satisfying*

$$\overline{\mathfrak{m}}\, T \leqslant \log \frac{\hat{\lambda}_{\mathrm{fin}}}{\hat{\lambda}}\,, \tag{1.32}$$

where $\overline{\mathfrak{m}} = \mathfrak{m} \vee 0$, we have for all $(t,x) \in (0,T] \times \mathbf{R}^2$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left|\sqrt{\log \tfrac{1}{\varepsilon}} \cdot u_\varepsilon(t,x) - v(t,x)\right|^2\right] = 0\,.$$

Here $v = v^{(\hat{\lambda})}$ denotes the solution of the linear equation

$$\partial_t v = \frac{1}{2}\Delta v + \mathfrak{m}\, v + \sigma_{\hat{\lambda}} \cdot \delta_0 \eta\,, \quad v(0,\cdot) = 0\,, \qquad \textit{with} \quad \sigma_{\hat{\lambda}} = \frac{\hat{\lambda}}{\sqrt{1 + \frac{3}{\pi}\hat{\lambda}^2}}\,.$$

The novelty of our approach to (1.31) is that we are able to treat the **full** (infinite) expansion of a critical singular SPDE. We stress that our result only covers short times when $\mathfrak{m} > 0$, thus, falls short of analysing the large time behaviour of the Allen–Cahn equation and its front-formation according to mean-curvature flow. However, the limit in Theorem D is well-defined for an arbitrary choice of parameters and we expect the restrictions on $\hat{\lambda}$ and T to be solely a consequence of our approach's technical restrictions. In the author's eyes, the explicit treatment of (1.31), by an infinite power series expansion in $\hat{\lambda}$, is possibly a first step towards a more general framework of the treatment of critical SPDEs in the weak coupling limit.

1.4 Outline of the thesis

The remainder of this thesis is structures as follows.

- In Chapter 2, we begin with a general introduction into Wiener chaoses and white noise. Thereafter, we make the reader familiar with rooted trees, their use in (stochastic) differential equations, and in particular Wild expansions of additive SPDEs. This can be viewed as the technical setup for Chapter 3.

Each of the remaining chapters consists of one of three the topics introduced in the sections above, with the aim of proving the results presented. The chapters are independent of each other and can be enjoyed in arbitrary order.

- In Chapter 3, we derive the weak coupling limit of the 2D Allen–Cahn equation with white noise initial data, cf. Theorem D, by treating an infinite Wild expansion of the equation. The chapter is based on the work [GRZ23].

- Chapter 4 consists of the analysis of the path measure of the 2D-DRPM, cf. Theorem C. We prove an invariance principle for the polymer paths in the full subcritical weak disorder regime. The chapter is based on the article [Gab23].

- In Chapter 5, we derive dynamical scaling limits of the condensed inclusion process, presented in Theorems A and B. Moreover, we link the derived limits to a well-studied model in population genetics. The chapter is based on the paper [CGG23].

Chapter 2

Preliminaries

The aim of this chapter is to give a self-contained exposition on tree expansions to SPDEs, which is the basis for Chapter 3. We will start by giving a basic introduction to white noise and the theory of Wiener chaoses. In the second part, we shed light on how rooted trees appeared in the study of ordinary differential equations (ODEs) and the Butcher series. Finally, we set up the Wild (or tree) expansion for SPDEs with additive noise. Most notably, we explain how the Wild expansion of such SPDEs links to corresponding Butcher series.

2.1 A primer on Wiener chaoses

In this section, we introduce Gaussian white noise, Wiener chaoses, the Wiener chaos decomposition and contractions. The exposition is in particular aimed at readers unfamiliar with these concepts. Throughout, we will implement generalised definitions of the Itô and the Stratonovich integral, and relations between the two. As our main reference, we are following the monographs [Jan97, Nua06].

2.1.1 Gaussian Hilbert spaces and white noise

The space of random distributions associated with white noise is a special case of a Gaussian Hilbert space, which can be set up in a very general manner.

Definition 2.1.1 (Gaussian Hilbert space)**.** *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. A Gaussian Hilbert space $\mathcal{H}$ is a real, linear and complete subspace of $L^2(\mu) = L^2(\Omega, \mathcal{F}, \mu)$ which only consists of centred Gaussian random variables.*

A simple, but non-trivial, example of a Gaussian Hilbert space $\mathcal{H}$ is the span of $\{W_t\}_{t\geqslant 0}$, where W is a standard Brownian motion, cf. [Jan97, Example 1.10]. Because linear combinations with respect to $\{W_t\}_{t\geqslant 0}$ approximate Itô integrals, one has the explicit representation

$$\mathcal{H} = \left\{ \int_0^\infty h(t)\, \mathrm{d}W_t \,:\, h \in L^2([0,\infty)) \right\} . \tag{2.1}$$

On the other hand, the Gaussian Hilbert space that will be central throughout the thesis is the one induced by white noise, which we introduced informally in (1.17). Our goal in the remainder of the section is to make the previous discussion on white noise rigorous. For this, we denote by $\mathcal{S}'(\mathbf{R}^d)$ the space of real tempered distributions, which is the continuous (or topological) dual of the Schwartz space (rapidly decreasing test functions) [Str03]

$$\mathcal{S}(\mathbf{R}^d) := \left\{ f \in C^\infty(\mathbf{R}^d) \,:\, \forall \mathbf{n}, \mathbf{m} \in \mathbf{N}^d \ \sup_{x\in\mathbf{R}^d} |x^{\mathbf{n}}\, \partial^{\mathbf{m}} f(x)| < \infty \right\} ,$$

where we used multi-index notation. White noise η on $\mathbf{R}^d$ can then be made sense of as a random distribution $\mathcal{S}'(\mathbf{R}^d)$.

We want to give a brief outline on the construction of white noise in the following, which requires the introduction of some technical preliminaries which we shall state without a proof and without going into further details. Instead, we refer the interested reader to [Jan97, Example 1.13].

Definition 2.1.2 (Gaussian measure)**.** *Let $\mathcal{X}$ be a real locally convex topological vector space and μ a Borel probability measure on $\mathcal{X}$. We call μ a (symmetric) Gaussian measure if*

$$x^*_\# \mu \quad \text{is a centred Gaussian on } \mathbf{R}\,, \qquad \text{for all } x^* \in \mathcal{X}^* .$$

Here $f_\#\mu$ denotes the pushforward measure of μ under a suitable function f.

Lemma 2.1.3. *Let $\mathcal{X}$ be a real locally convex topological vector space and μ a symmetric Gaussian measure on $\mathcal{X}$, then the completion of $\mathcal{X}^* \subset L^2(\mathcal{X}, \mathcal{B}, \mu)$ is a Gaussian Hilbert space. Here, $\mathcal{B}$ denotes the Borel-σ-algebra on X.*

Having the above lemma at hand, we can prove the existence of white noise. When defining white noise, the goal is to construct a Gaussian measure $\mathbb{P}$ on $\mathcal{S}'(\mathbf{R}^d)$ such that for $\eta \sim \mathbb{P}$

$$\eta(\varphi) \sim \mathcal{N}(0, \|\varphi\|^2_{L^2(\mathbf{R}^d)}) \quad \text{and} \quad \mathbb{E}[\eta(\varphi)\,\eta(\varphi')] = \int_{\mathbf{R}^d} \varphi(x)\varphi'(x)\,\mathrm{d}x\,, \quad (2.2)$$

for all $\varphi, \varphi' \in \mathcal{S}(\mathbf{R}^d)$, where we wrote $\mathbb{E}$ for the expectation with respect to $\mathbb{P}$.

Lemma 2.1.4 (White noise)**.** *There exists a Gaussian measure $\mathbb{P}$ on $\mathcal{S}'(\mathbf{R}^d)$ satisfying (2.2). We call $\mathbb{P}$ the **white noise measure** and the random distribution $\eta \sim \mathbb{P}$ **white noise**.*

The construction of such a measure $\mathbb{P}$, and its corresponding Gaussian Hilbert space, is rather technical, however very interesting. The linchpin of the proof is the (Bochner-)Minlos theorem, which states that any continuous semi-definite symmetric bilinear form on $\mathcal{S}(\mathbf{R}^d)$ corresponds to a unique Gaussian measure on $\mathcal{S}'(\mathbf{R}^d)$ [GV64, Chapter III/IV]. Given this result, we don't want to shy away from the construction of white noise, as it sheds light on some of its properties which may be interesting to the unfamiliar reader. We follow the argument presented in [Jan97, Example 1.16], however, omitting the technical details.

Proof. In order to avoid confusion throughout the proof, we will write $\mathcal{S}^*$ instead of $\mathcal{S}'$, which in our case both denotes the continuous (or topological) dual.

We start by assuming that the Gaussian measure $\mathbb{P}$ on $\mathcal{S}^*(\mathbf{R}^d)$ with the desired properties (2.2) exists. For convenience, let us define the random variable $\eta \sim \mathbb{P}$ (which the reader should think of as white noise). Let $\mathfrak{J} : \mathcal{S}(\mathbf{R}^d) \to \mathcal{S}^{**}(\mathbf{R}^d)$ be the evaluation map, which for every $\varphi \in \mathcal{S}(\mathbf{R}^d)$ yields a functional

$$\mathfrak{J}(\varphi)(F) := F(\varphi)\,, \quad \forall F \in \mathcal{S}^*(\mathbf{R}^d)\,.$$

Note that $\mathcal{S}(\mathbf{R}^d)$ is reflexive, and $\mathfrak{J}$ is an isometry to its bidual $\mathcal{S}^{**}(\mathbf{R}^d)$. Because $\mathbb{P}$ is assumed to be a Gaussian measure on $\mathcal{S}^*(\mathbf{R}^d)$, the push-forward measure $\mathfrak{J}(\varphi)_\#\mathbb{P}$ is normal, i.e. $\eta(\varphi)$ is a centred Gaussian, cf. Definition 2.1.2. Furthermore, by Lemma 2.1.3, the completion of $\mathcal{S}^{**}(\mathbf{R}^d)$ is a Gaussian Hilbert space, and so is the space

$$\mathcal{H} := \overline{\{\mathfrak{J}(\varphi) \in \mathcal{S}^{**}(\mathbf{R}^d) \,:\, \varphi \in \mathcal{S}(\mathbf{R}^d)\}} \subset L^2(\mathcal{S}^*(\mathbf{R}^d), \mathcal{B}, \mathbb{P})\,,$$

as it is a closed subspace. Let us also define the symmetric semi-definite bilinear form $\mathcal{E}$ on $\mathcal{S}(\mathbf{R}^d)$

$$\mathcal{E}(\varphi, \varphi') := \mathbb{E}[\eta(\varphi)\eta(\varphi')]\,. \tag{2.3}$$

Then, the completion $S_\mathcal{E}$ of $(\mathcal{S}(\mathbf{R}^d), \mathcal{E})$ is a Hilbert space. Lastly, it can be shown that $\mathfrak{J}$ is a continuous map from $\mathcal{S}(\mathbf{R}^d)$ to $\mathcal{H}$, which also implies continuity of $\mathcal{E}$. Thus, $\mathfrak{J}$ extends to an isometry $\mathrm{I} : S_\mathcal{E} \to \mathcal{H}$.

Now, the crucial step in the construction is due to Minlos's theorem, which allows to interpret the definition in (2.3) in the reverse direction. Namely, it links any choice of a continuous semi-definite symmetric bilinear form $\mathcal{E}$ on $\mathcal{S}(\mathbf{R}^d)$ to a Gaussian measure with covariance structure induced by $\mathcal{E}$. In particular, this yields the existence of a Gaussian measure $\mathbb{P}$ for the classical choice $\mathcal{E}(\varphi, \varphi') = \langle \varphi, \varphi' \rangle_{L^2(\mathbf{R}^d)}$. □

In the above proof, we saw that for white noise $S_\mathcal{E} = L^2(\mathbf{R}^d)$ with the usual inner product. Hence, white noise as a random (tempered) distribution $\eta \sim \mathbb{P}$ can be naturally extended to test functions in $L^2(\mathbf{R}^d)$, using the extension $\mathrm{I} = \overline{\mathfrak{J}}$ of the evaluation map.

Corollary 2.1.5. *Let $\mathbb{P}$ be the white noise measure on $\mathcal{S}'(\mathbf{R}^d)$. Then the corresponding Gaussian Hilbert space*

$$\mathcal{H} := \overline{\{\eta(\varphi) \,:\, \varphi \in \mathcal{S}(\mathbf{R}^d)\}} \subset L^2(\mathbb{P})$$

is isometric to $L^2(\mathbf{R}^d)$. In other words, white noise η can be extended from $\mathcal{S}(\mathbf{R}^d)$ to $L^2(\mathbf{R}^d)$ using the isometry I.

Corollary 2.1.5 states that $\mathcal{H}$ can be indexed by $L^2(\mathbf{R}^d)$, cf. [Jan97, Definition 1.18]. In [Nua06], this property is part of the definition of an "isonormal Gaussian process". Notably, Corollary 2.1.5 allows to interpret $\mathrm{I}(f) \in \mathcal{H}$, $f \in L^2(\mathbf{R}^d)$, as a stochastic integral satisfying Itô's isometry

$$\mathbb{E}[|\mathrm{I}(f)|^2] = \|f\|^2_{L^2(\mathbf{R}^d)}\,.$$

Recall also the case of Brownian motion (2.1), which can be viewed as a special case of white noise $\eta = \dot{W}$ for $d = 1$. The integral representation will be advantageous throughout later sections, thus, we will often write

$$\int_{\mathbf{R}^d} f(z)\,\eta(\,\mathrm{d}z) = \eta(f) := \mathrm{I}(f)\,, \qquad \forall f \in L^2(\mathbf{R}^d)\,, \tag{2.4}$$

while being aware that η is not defined pointwise.

2.1.2 The Wiener chaos decomposition

From now on, we shall only consider the Gaussian Hilbert space $\mathcal{H}$ generated by the white noise measure $\mathbb{P}$. However, most of this section holds for general Gaussian Hilbert spaces, see [Jan97, Chapter 2].

First, note that elements $Z_1, \dots, Z_m \in \mathcal{H}$ are Gaussian random variables, which have moments of arbitrary order. In particular,

$$\left\| \prod_{i=1}^{m} Z_i \right\|_{L^2(\mathbb{P})} \leqslant \prod_{i=1}^{m} \|Z_i\|_{L^{2m}(\mathbb{P})} = \sqrt{(2m-1)!!} \prod_{i=1}^{m} \|Z_i\|^2_{L^2(\mathbb{P})} < \infty ,$$

by an m-fold application of Hölder's inequality and Wick's theorem

$$\|Z_i\|^{2m}_{L^{2m}(\mathbb{P})} = \mathbb{E}[|Z_i|^{2m}] = (2m-1)!! \|Z_i\|^{2m}_{L^2(\mathbb{P})} . \tag{2.5}$$

Notably, we constructed the product $Z_1 \cdots Z_m \in L^2(\mathbb{P})$ outside of $\mathcal{H}$ by simple algebraic means. In the present section, we will see that at its furthest, this idea can be pushed to reconstruct (almost all) of the space $L^2(\mathbb{P})$.

We start by defining the building blocks of this theory:

Definition 2.1.6 (Wiener chaoses)**.** *Let $n \in \mathbf{N}$. The **homogeneous chaos of order** n is the space*

$$\mathcal{H}_n := \overline{\mathcal{P}}_n(\mathcal{H}) \cap \overline{\mathcal{P}}_{n-1}(\mathcal{H})^{\perp} , \tag{2.6}$$

where $\overline{\mathcal{P}}_n(\mathcal{H})$ is the $L^2(\mathbb{P})$-closure of the linear space

$$\overline{\mathcal{P}}_n(\mathcal{H}) := \big\{ p(Z_1, \dots, Z_m) \ : p \text{ a polynomial of degree } \leqslant n , \\ Z_1, \dots, Z_m \in \mathcal{H} , \ m \in \mathbf{N} \big\} . \tag{2.7}$$

*Components of the increasing sequence of subspaces $\overline{\mathcal{P}}_n(\mathcal{H})$ are commonly referred to as **inhomogeneous chaos of order** n. We will also write $\mathcal{H}_{\leqslant n} := \overline{\mathcal{P}}_n(\mathcal{H})$.*

The spaces $\mathcal{H}_n$ are mutually orthogonal by definition, which yields

$$\mathcal{H}_{\leqslant n} = \overline{\mathcal{P}}_n(\mathcal{H}) = \bigoplus_{k=0}^{n} \mathcal{H}_k . \tag{2.8}$$

Recalling (2.5), higher moments of normal variables can be estimated in terms of second-moments. Because elements of Wiener chaoses are represented in

terms of polynomials of such Gaussian random variables, this observation can be generalised to Wiener chaoses.

Proposition 2.1.7 (Hypercontractivity)**.** *Let $\mathcal{H} \subset L^2(\mathbb{P})$ be a Gaussian Hilbert space and $n \in \mathbf{N}$. Then for every $p, q < \infty$ there exists a constant $c_{p,q} < \infty$ such that*

$$\|X\|_{L^q(\mathbb{P})} \leqslant (c_{p,q})^n \|X\|_{L^p(\mathbb{P})}\,, \quad \forall X \in \mathcal{H}_{\leqslant n}\,.$$

For the case $p = 2$ and $q \geqslant 2$, one has $c_{2,q} = \sqrt{q-1}$.

For a proof of this proposition, we refer the reader to [Jan97, Theorem 5.10], and [Jan97, Remark 5.11] for the explicit constant $c_{2,q}$.

Next, we demystify the abstract definition of Wiener chaoses by expressing elements explicitly in terms of stochastic iterated integrals. First, we note that the notion of the stochastic integral $\mathrm{I}(f)$ from (2.4) can be extended to stochastic iterated integrals of the form

$$\mathrm{I}_n(f) = \int_{(\mathbf{R}^d)^n} f(z_1, \dots, z_n)\eta^{\otimes n}(\,\mathrm{d}z_{1,\dots,n})\,, \tag{2.9}$$

which again do not make sense as displayed. Instead, one can define linear operators $\mathrm{I}_n : L^2((\mathbf{R}^d)^n) \to L^2(\mathbb{P})$ to satisfy the properties of iterated stochastic integrals. The construction of I_n is similar to the classical Itô integral with respect to Brownian motion by defining the operator on simple test functions and extending it to all of $L^2((\mathbf{R}^d)^n)$ by continuity. We refer to [Nua06, Section 1.1.2] for a detailed construction.

Lemma 2.1.8. *Let $\mathbb{P}$ be the white noise measure and $\mathcal{H}$ as in Corollary 2.1.5. Then for every $n \in \mathbf{N}$,*

$$\mathrm{I}_n : L^2_{\mathrm{sym}}((\mathbf{R}^d)^n) \to \mathcal{H}_n$$

is an isomorphism[1], with $L^2_{\mathrm{sym}}(\cdot)$ being the subspace of $L^2(\cdot)$-functions that are symmetric in their arguments. Recall that I_n denotes the (abstract) iterated integral operator, cf. (2.9). In particular, I_n can be extended to $f \in L^2((\mathbf{R}^d)^n)$ by symmetrisation

$$\mathrm{I}_n(f) := \mathrm{I}_n(\tilde{f})\,, \qquad \text{with} \quad \tilde{f}(z_1, \dots, z_n) := \frac{1}{n!}\sum_{\sigma \in S_n} f(z_{\sigma(1)}, \dots, z_{\sigma(n)})\,, \tag{2.10}$$

where S_n is the symmetric group of size n. We say $\mathrm{I}_n(f)$ is a ***Wiener–Itô stochastic integral****.*

[1] Multiplying I_n with $\sqrt{n!}^{-1}$ yields an isometry.

To avoid further technical details, we refer the curious reader to [Jan97, Section 7.2], and in particular to [Jan97, Theorem 7.26] for an exposition of the statement above. It is interesting to note that the construction of $\mathrm{I}_n(f)$ does only depend on f up to symmetries, cf. (2.10), which can be formally argued to be a consequence of rotationally invariance (in law) of white noise. At this point, let us stress the difference of two notions for stochastic integrals used throughout: $\eta^{\otimes n}(\,\mathrm{d}z_{1,\dots,n})$ will denote Itô integration, whereas $\prod_i^n \eta(\,\mathrm{d}z_i)$ will denote Stratonovich integration, introduced in Definition 2.1.14 below.

An important property of the stochastic integrals' construction is the following orthogonality, which can be found in [Nua06, Section 1.1.2] alongside other properties of the operators I_n, see also [Jan97, Section 7.2]. For $m, n \in \mathbf{N}$ and $f \in L^2((\mathbf{R}^d)^n)$, $g \in L^2((\mathbf{R}^d)^m)$

$$\mathbb{E}[\mathrm{I}_n(f)] = 0 \quad \text{and} \quad \mathbb{E}\left[\mathrm{I}_n(f)\,\mathrm{I}_m(g)\right] = n!\,\mathbb{1}_{n=m}\,\langle \tilde{f}, \tilde{g}\rangle_{L^2((\mathbf{R}^d)^n)}\,, \tag{2.11}$$

where $\tilde{\cdot}$ denotes the symmetrisation operator defined in (2.10). In particular, we have

$$\mathbb{E}\left[\left|\mathrm{I}_n(f)\right|^2\right] = n!\,\|\tilde{f}\|^2_{L^2((\mathbf{R}^d)^n)} \leqslant n!\,\|f\|^2_{L^2((\mathbf{R}^d)^n)}\,. \tag{2.12}$$

Also note that (2.11) is in accordance with the orthogonality of homogeneous Wiener chaoses (2.8).

Example 2.1.9. Consider functions $h_1, h_2 \in L^2(\mathbf{R}^d)$ and the product

$$\mathrm{I}_1(h_1)\mathrm{I}_1(h_2) = \int_{(\mathbf{R}^2)^2} h_1(z_1)h_2(z_2)\eta(\,\mathrm{d}z_1)\eta(\,\mathrm{d}z_2)\,,$$

which lies in $\mathcal{H}_{\leqslant 2}$. Naturally, we would like to express the above product in terms of I_2, by writing the product of integrals in terms of

$$\int_{(\mathbf{R}^2)^2} h_1(z_1)h_2(z_2)\eta^{\otimes 2}(\,\mathrm{d}z_{1,2}) = \mathrm{I}_2(h_1 \otimes h_2)\,.$$

However, as it is the case for the classical Itô integral, we have to take a correction term into account:

$$\mathrm{I}_1(h_1)\mathrm{I}_1(h_2) = \mathrm{I}_2(h_1 \otimes h_2) + \mathbb{E}\left[\mathrm{I}_1(h_1)\mathrm{I}_1(h_2)\right]\,,$$

see [Nua06, Proposition 1.1.2], where we used the notation

$$(h_1 \otimes \cdots \otimes h_n)(z_1, \dots, z_n) := h_1(z_1) \cdots h_n(z_n)\,. \tag{2.13}$$

Because $\mathrm{I}_2(h_1 \otimes h_2)$ has zero mean, cf. (2.11), it is orthogonal with respect to $\mathbb{E}\left[\mathrm{I}_1(h_1)\mathrm{I}_1(h_2)\right]$. Moreover, due to orthogonality (2.11), $\mathrm{I}_2(h_1 \otimes h_2)$ is uncorrelated to any element in $\mathcal{H}_1 = \{\mathrm{I}_1(h) : h \in L^2(\mathbf{R}^d)\}$. In other words, we have discovered an expansion of $\mathrm{I}_1(h_1)\mathrm{I}_1(h_2) \in L^2(\mathbb{P})$ into its homogeneous chaoses:

$$\mathbb{E}\left[\mathrm{I}_1(h_1)\mathrm{I}_1(h_2)\right] \in \mathbf{R} = \mathcal{H}_0 \quad \text{and} \quad \mathrm{I}_2(h_1 \otimes h_2) \in \mathcal{H}_2\,,$$

with all other chaos components being equal to zero.

We can finally state the Wiener chaos decomposition, which allows to represent $L^2(\mathbb{P})$ random variables in terms of series of stochastic iterated integrals, generalising the decomposition derived in Example 2.1.9.

Proposition 2.1.10 (Wiener–Itô chaos decomposition)**.** *Let $\mathbb{P}$ be the white noise measure on $(\mathcal{S}'(\mathbf{R}^d), \sigma(\mathcal{H}))$, where $\sigma(\mathcal{H})$ denotes the σ-algebra generated by $\mathcal{H}$. Then for every $X \in L^2(\mathcal{S}'(\mathbf{R}^d), \sigma(\mathcal{H}), \mathbb{P})$, there exist $f_n \in L^2((\mathbf{R}^d)^n)$, $n \in \mathbf{N}_0$, such that*

$$X = \sum_{n=0}^{\infty} \mathrm{I}_n(f_n)\,.$$

Moreover, the functions f_n are uniquely determined when restricted to $L^2_{\mathrm{sym}}((\mathbf{R}^d)^n)$.

For a detailed overview of this result, we refer the reader to [Nua06, Section 1.1] with the above proposition being stated in [Nua06, Theorem 1.1.2]. See also [Jan97, Theorem 7.26].

Let us conclude this section with a brief discussion on integrals of the form

$$\int_{\mathbf{R}^{d'}} \mathrm{I}_n(f_x)\,\mathrm{d}x\,, \tag{2.14}$$

where $\{f_x\}_{x\in\mathbf{R}^{d'}}$ be a family of $L^2((\mathbf{R}^d)^n)$ functions, which will appear naturally in Section 2.3. A priori, the integral (2.14) does not have any meaning. However, if $\{f_x\}_{x\in\mathbf{R}^{d'}}$ are such that

$$F := \int_{\mathbf{R}^{d'}} f_x(\cdot)\,\mathrm{d}x \in L^2((\mathbf{R}^d)^n), \tag{2.15}$$

then we can define

$$\int_{\mathbf{R}^{d'}} \mathrm{I}_n(f_x)\,\mathrm{d}x := \mathrm{I}_n(F)\,, \tag{2.16}$$

which reminds of a Fubini theorem. Once again, this follows from an extension argument: Starting with a family $\{f_x\}_{x\in\mathbf{R}^{d'}}$ that is constant on finitely

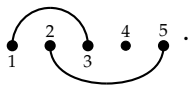
many subsets of $\mathbf{R}^{d'}$, (2.16) holds as an identity by linearity of the operator I_n. Now, approximating F of the form (2.15) by such elementary functions, we can show that the definition in (2.16) is indeed appropriate.

2.1.3 Contractions and Feynman diagrams

In Example 2.1.9, we saw that the product $\mathrm{I}(h_1)\mathrm{I}(h_2)$ is given in terms of a Stratonovich integral, and has an explicit Wiener chaos decomposition. In this section, we introduce the Stratonovich integral in terms of corrected Itô integrals, using the notion of **contractions**.

Let A be a finite set, then a **contraction** κ **of** A is an element in $\binom{A}{2}$, which denotes the set of unordered pairs of A, such that any $i \in A$ lies in at most one element of κ. The set of all contractions of A is denoted by $\mathcal{K}(A)$. We say the pair (A, κ) is the set A **with contraction** κ or simply call A a **contracted set**, and sometimes write A_κ .

Example 2.1.11. Let $A = \{1, \dots, 5\}$. Then $\kappa \in \mathcal{K}(A)$ is of the form



. (2.17)

Here we represented elements of A as vertices and elements of $\kappa = \{\{1,3\},\{2,5\}\}$ as edges connecting them. In the literature such contraction diagrams are sometimes referred to as **Feynman diagrams**, see for example [Jan97, Section 1.5].

Now, consider a finite set A of size $n \in \mathbf{N}$, a function $f \in L^2((\mathbf{R}^d)^A)$ and a contraction $\kappa \in \mathcal{K}(A)$. For convenience let us also define $m := n - 2|\kappa|$ and write $A \setminus \kappa$ for the elements in A not contained in κ. We can then **contract** f **with respect to** κ, which is the function

$$f_\kappa(\hat{z}_{A\setminus\kappa}) := \int_{(\mathbf{R}^d)^A} f(z_A) \prod_{i\in A\setminus\kappa} \delta_{\hat{z}_i}(z_i) \prod_{\{j,j'\}\in\kappa} \delta_{z_j}(z_{j'})\, \mathrm{d}z_A\,, \tag{2.18}$$

where we identified variables according to the pairs in κ. Here and throughout the thesis, we will write $z_A = (z_i)_{i\in A} \in (\mathbf{R}^d)^A$, or simply $z_{1,\dots,n}$ if $A = \{1, \dots, n\}$.

Example 2.1.12. Consider the contraction κ from (2.17) and $f \in L^2((\mathbf{R}^d)^5)$,

then $f_\kappa : \mathbf{R}^d \to \mathbf{R}$ with

$$f_\kappa(\cdot) = \int_{(\mathbf{R}^d)^2} f(z_1, z_2, z_1, \cdot, z_2)\, \mathrm{d}z_{1,2}\,.$$

The single free variable corresponds to the unmatched vertex in (2.17). The remaining pairs have been identified and integrated, according to κ.

However, there is no reason why f_κ should lie in $L^2((\mathbf{R}^d)^{A\setminus\kappa})$ since the contraction operator $(\cdot)_\kappa$ is unbounded and not necessarily well-defined [Jan97, Remark 7.32]. For example, consider the function $f(z_1, z_2) := z_1^{-1}\mathbb{1}_{z_1=z_2\in(0,1]}$ that lies in $L^2((\mathbf{R})^2)$ and the contraction $\kappa = \{\{1,2\}\}$. Then

$$f_\kappa = \int_{\mathbf{R}} f(z,z)\,\mathrm{d}z = \int_0^1 \frac{1}{z}\,\mathrm{d}z = \infty\,.$$

To avoid this stepping stone, we restrict ourselves to functions f such that any contracted version remains L^2-integrable.

Definition 2.1.13 (Contractable functions)**.** *We say a function $f \in L^2((\mathbf{R}^d)^A)$ is **contractable** if*

$$\|f_\kappa\|_{L^2((\mathbf{R}^d)^{A\setminus\kappa})} < \infty\,, \qquad \forall \kappa \in \mathcal{K}(A)\,.$$

It is worth mentioning that functions of the form $f = \bigotimes_{i\in A} h_i$, cf. (2.13), are contractable whenever $h_i \in L^2(\mathbf{R}^d)$: First, note that for every $\kappa \in \mathcal{K}(A)$

$$f_\kappa = \prod_{i\in A\setminus\kappa} h_i(\cdot) \prod_{\{j,j'\}\in\kappa} \langle h_j, h_{j'}\rangle_{L^2(\mathbf{R}^d)}\,,$$

hence, by the Cauchy-Schwartz inequality

$$\begin{aligned}\|f_\kappa\|_{L^2((\mathbf{R}^d)^{A\setminus\kappa})} &\leqslant \prod_{i\in A\setminus\kappa} \|h_i\|_{L^2(\mathbf{R}^d)} \prod_{\{j,j'\}\in\kappa} \|h_j\|_{L^2(\mathbf{R}^d)} \|h_{j'}\|_{L^2(\mathbf{R}^d)} \\ &= \prod_{i=1}^n \|h_i\|_{L^2(\mathbf{R}^d)}\,.\end{aligned}$$

See also [Jan97, Lemma 7.31].

In Example 2.1.9, we derived the decomposition

$$\mathrm{I}_1(h_1)\mathrm{I}_1(h_2) = \mathrm{I}_2(h_1\otimes h_2) + \mathbb{E}\left[\mathrm{I}_1(h_1)\mathrm{I}_1(h_2)\right]\,, \qquad h_1, h_2 \in L^2(\mathbf{R}^d)\,,$$

as for the classical Itô integral. This identity can also be viewed as definition

of the Stratonovich integral $\mathring{\mathrm{I}}_2(h_1 \otimes h_2)$. Building up on this idea, we define Wiener–Stratonovich integrals as follows.

Definition 2.1.14 (Wiener–Stratonovich integral)**.** *Let A be a finite set of size $n \in \mathbf{N}$, and $f \in L^2((\mathbf{R}^d)^A)$ be contractable, cf. Definition 2.1.13. Then the **Stratonovich integral** with respect to f is defined as*

$$\mathring{\mathrm{I}}_n(f) = \int_{(\mathbf{R}^d)^n} f(z_1, \dots, z_n) \prod_{i=1}^n \eta(\,\mathrm{d}z_i) := \sum_{\kappa \in \mathcal{K}(A)} \mathrm{I}_{n-2|\kappa|}(f_\kappa)\,,$$

with f_κ as in (2.18).

For example, the function $f = h \otimes h \otimes h \in L^2((\mathbf{R}^d)^3)$, with $h \in L^2(\mathbf{R}^d)$, is contractable and we have

$$\mathring{\mathrm{I}}_3(f) = \mathrm{I}_1(h)^3 = \mathrm{I}_3(f) + 3\,\mathbb{E}[|\mathrm{I}_1(h)|^2]\,\mathrm{I}_1(h)\,.$$

In particular, the following product rule holds:

Lemma 2.1.15 (Product formula for Stratonovich integral)**.** *Let $m \in \mathbf{N}$, $\{A_i\}_{i=1}^m$ finite sets with $n_i = |A_i|$ and $f_i \in L^2((\mathbf{R}^d)^{A_i})$, then*

$$\prod_{i=1}^m \mathring{\mathrm{I}}_{n_i}(f_i) = \mathring{\mathrm{I}}_{|\mathbf{n}|}(f_1 \otimes \cdots \otimes f_m)\,,$$

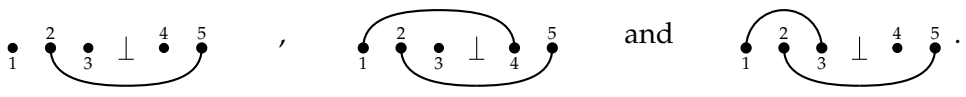
where we used the multi-index notation $|\mathbf{n}| = n_1 + \cdots n_m$.

The proof of this product formula requires a slightly extended notion of contractions. When considering finite sets $\{A_i\}_{i=1}^m$, we define the set of contractions between them

$$\mathcal{K}_\perp(A_1, \dots, A_m) := \left\{ \kappa \in \mathcal{K}(A_1 \cup \cdots \cup A_m) \,:\, \kappa \cap \binom{A_i}{2} = \emptyset \;\; \forall\, 1 \leqslant i \leqslant m \right\},$$

i.e. κ does not contain any pairs connecting elements within the A_i's.

Example 2.1.16. Let $A_1 = \{1, 2, 3\}$ and $A_2 = \{4, 5\}$, and consider the contractions



The first two contractions lie in $\mathcal{K}_\perp(A_1, A_2)$, however, the third one only lies in $\mathcal{K}(A_1 \cup A_2)$ because there is a pairing within A_1.

We can now state a generalisation of Wick's theorem, referring the reader to [Jan97, Lemma 7.33] for a proof.

Lemma 2.1.17. *Let $m \in \mathbf{N}$, $\{A_i\}_{i=1}^m$ finite sets with $n_i = |A_i|$ and $f_i \in L^2((\mathbf{R}^d)^{A_i})$, then*

$$\prod_{i=1}^m \mathrm{I}_{n_i}(f_i) = \sum_{\kappa \in \mathcal{K}_\perp} \mathrm{I}_{|\mathbf{n}|-2|\kappa|}\big((f_1 \otimes \cdots \otimes f_m)_\kappa\big)\,,$$

where we wrote $\mathcal{K}_\perp$ as a shorthand for $\mathcal{K}_\perp(A_1, \dots, A_m)$.

Now we have all the ingredients to provide a proof of the product formula for Stratonovich integrals.

Proof of Lemma 2.1.15. Let $m \in \mathbf{N}$, $\{A_i\}_{i=1}^m$ finite sets with $n_i = |A_i|$ and $f_i \in L^2((\mathbf{R}^d)^{A_i})$. By Definition 2.1.14 of the Stratonovich integral, we can write

$$\begin{aligned}\prod_{i=1}^m \mathring{\mathrm{I}}_{n_i}(f_i) &= \prod_{i=1}^m \sum_{\kappa_i \in \mathcal{K}(A_i)} \mathrm{I}_{n_i-2|\kappa_i|}\big((f_i)_{\kappa_i}\big) \\ &= \sum_{\substack{\kappa_i \in \mathcal{K}(A_i) \\ 1 \leqslant i \leqslant m}} \prod_{i=1}^m \mathrm{I}_{n_i-2|\kappa_i|}\big((f_i)_{\kappa_i}\big)\,.\end{aligned} \tag{2.19}$$

Now, we apply Lemma 2.1.17 for $\{A_i \setminus \kappa_i\}_{i=1}^m$ and $\{(f_i)_\kappa\}_{i=1}^m$ which yields

$$\prod_{i=1}^m \mathrm{I}_{n_i-2|\kappa_i|}\big((f_i)_{\kappa_i}\big) = \sum_{\overline{\kappa} \in \mathcal{K}_\perp} \mathrm{I}_{\overline{n}}\big(\,((f_1)_{\kappa_1} \otimes \cdots \otimes (f_m)_{\kappa_m})_{\overline{\kappa}}\,\big)\,, \tag{2.20}$$

with $\mathcal{K}_\perp = \mathcal{K}_\perp(A_1 \setminus \kappa_1, \dots, A_m \setminus \kappa_m)$ and $\overline{n} = \sum_{i=1}^m (n_i - 2|\kappa_i|) - 2|\overline{\kappa}|$. Note that any tupel $(\overline{\kappa}, \kappa_1, \dots, \kappa_m)$ can be mapped to its naturally induced contraction $\kappa = \overline{\kappa} \cup \kappa_1 \cup \cdots \cup \kappa_m$ and we write

$$\mathrm{I}_{\overline{n}}\big(\,((f_1)_{\kappa_1} \otimes \cdots \otimes (f_m)_{\kappa_m})_{\overline{\kappa}}\,\big) = \mathrm{I}_{|\mathbf{n}|-2|\kappa|}\big(\,(f_1 \otimes \cdots \otimes f_m)_\kappa\,\big)\,. \tag{2.21}$$

In particular, one can check that the above mapping defines a bijection between $\mathcal{K}_\perp(A_1, \dots, A_m) \times \mathcal{K}(A_1) \times \cdots \times \mathcal{K}(A_m)$ and $\mathcal{K}(A_1 \cup \cdots \cup A_m)$. Lastly, combining (2.19), (2.20) and (2.21) we have

$$\prod_{i=1}^m \mathring{\mathrm{I}}_{n_i}(f_i) = \sum_{\kappa \in \mathcal{K}(A_1 \cup \cdots \cup A_m)} \mathrm{I}_{|\mathbf{n}|-2|\kappa|}\big(\,(f_1 \otimes \cdots \otimes f_m)_\kappa\,\big) = \mathring{\mathrm{I}}_{|\mathbf{n}|}(f_1 \otimes \cdots \otimes f_m)\,,$$

which finishes the proof. □

We conclude the section with a Fubini-like theorem for the Stratonovich integral, which is an immediate consequence of Definition 2.1.14 and (2.16).

Lemma 2.1.18 (Fubini-like for Stratonovich)**.** *Let A be a finite set of size n and $\{f_x\}_{x\in\mathbf{R}^{d'}}$ be a family of $L^2((\mathbf{R}^d)^A)$ functions that are contractable, cf. Definition 2.1.13, such that also*

$$F := \int_{\mathbf{R}^{d'}} f_x(\cdot)\,\mathrm{d}x \in L^2((\mathbf{R}^d)^A)$$

is contractable. Then

$$\int_{\mathbf{R}^{d'}} \mathring{\mathrm{I}}_n(f_x)\,\mathrm{d}x = \mathring{\mathrm{I}}_n(F)\,.$$

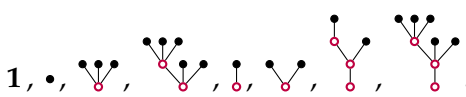
Remark 2.1.19. *Throughout the section, we restricted ourselves to white noise on $\mathbf{R}^d$. However, all results presented should hold generally with with $\mathbf{R}^d$ being replaced by a σ-finite non-atomic measure space $(E,\mathcal{E},\nu)$, as for example considered in [Jan97]. In particular, this allows to also consider the time-space domain $\mathbf{R}_+\times\mathbf{R}^d$.*

2.2 On rooted trees and the Butcher series

In this section, we will review how trees are used to index series expansions of solutions to ODEs. In particular, the kind of expansion that we are interested in goes under the name of **Butcher series** (or **B-series** for short) [But63, HW74], which originates from numerical analysis. Since its introduction, approaches of the same flavor were successfully applied to SDEs [Gub10] and SPDEs [Hai13, Hai14].

2.2.1 Rooted trees

First, we introduce the central objects of this section: finite, rooted trees. We denote the set of finite, rooted, unordered trees by $\mathcal{T}$, containing trees τ



(2.22)

where **1** denotes the empty tree. In the pictorial description, the bottom-most node describes the **root**, whereas **leaves** are indicated by thick dots •. The remaining **inner vertices** are represented by ◦. Given a tree $\tau\in\mathcal{T}$, we denote the set of vertices by $\mathcal{V}(\tau)$ and the set of edges by $\mathcal{E}(\tau)$. Moreover, we partition $\mathcal{V}(\tau)$ into two categories:

- The set of **leaves**, which we write as $\mathcal{L}(\tau)$ with cardinality $\ell(\tau) = |\mathcal{L}(\tau)|$.

- The set of **inner vertices**, denoted by $\mathcal{I}(\tau)$. We also write $i(\tau) = |\mathcal{I}(\tau)|$.

We denote the total number of vertices by $|\tau| = \ell(\tau) + i(\tau)$. The **root** of τ is denoted by $\mathfrak{o}_\tau$, or simply $\mathfrak{o}$, and is considered an inner vertex in all cases but one. The exception is the tree $\tau = \bullet$, for which the single vertex, and root, is considered to be an element of $\mathcal{L}(\tau)$. Lastly, for $v \in \mathcal{V}(\tau) \setminus \mathfrak{o}$ the closest vertex $\mathfrak{p}(v) \in \mathcal{V}(\tau)$, in direction of the root, is called the **parent** of v. Every vertex u satisfying $\mathfrak{p}(u) = v$ is called a (direct) **descendant** of v.

As we will be working with *unordered* trees, the three trees below, for example, will be considered identical:

(2.23)

Trees in $\mathcal{T}$ can be constructed recursively as follows: Suppose $\tau_1, ..., \tau_n \in \mathcal{T}$, we then construct the tree $\tau := [\tau_1 \ \cdots \ \tau_n]$ by connecting the trees $\tau_1, ..., \tau_n \in \mathcal{T}$ to a common vertex, which acts as the root of the new tree τ. We use the depiction

$$[\tau_1 \cdots \tau_n] := \quad \overset{\tau_1 \ \tau_2 \ \dots \ \tau_n}{\diagup\!\!\diagup\!\!\diagup\!\!\diagup} \tag{2.24}$$

and call $[\cdots]$ the **grafting operator**. For example,

$$[\mathbf{1}] = \bullet, \qquad [\bullet] = \, \overset{\bullet}{|}, \qquad [\overset{\bullet}{|}\bullet] = \, \overset{\bullet}{\wedge}, \qquad \text{or} \qquad [\overset{\bullet}{\wedge}\bullet\bullet] = \, \overset{\bullet}{\wedge}.$$

We observe that the empty tree $\mathbf{1}$ can be ignored in a grafting (unless it is the only tree): $[\mathbf{1} \ \tau_1 \cdots \tau_n] = [\tau_1 \cdots \tau_n]$. In this notation, the space of finite, unordered rooted trees can be obtained from the space of finite rooted trees through the equivalence relation that identifies

$$[\tau_1 \cdots \tau_n] \equiv [\tau_{\sigma(1)} \cdots \tau_{\sigma(n)}] \tag{2.25}$$

for every $\sigma \in S_n$ (the group of permutations on n indices), and the same for any subtree of a given tree.

Moreover, it will be useful to consider the following types of subsets of $\mathcal{T}$: For $n \in \mathbf{N}$ the set of **sub-n-ary trees**

$$\mathcal{T}_{\leqslant n} := \{\tau \in \mathcal{T} \ : \text{ any inner node in } \tau \text{ has at most } n \text{ descendants}\} \tag{2.26}$$

and the set of n**-ary trees**

$$\mathcal{T}_n := \{\tau \in \mathcal{T} \setminus \{\mathbf{1}\} \,:\, \text{any inner node in } \tau \text{ has exactly } n \text{ descendants}\} \tag{2.27}$$

For example, all trees displayed in (2.22) lie in $\mathcal{T}_{\leqslant 3}$, but only the second, third and last tree lie in $\mathcal{T}_3$. Note that we work with the conventions that the empty tree belongs to $\mathcal{T}, \mathcal{T}_{\leqslant n}$, but **not** $\mathcal{T}_n$ and that the "single node" tree $\tau = \bullet$ belongs to $\mathcal{T}_n$ (because by convention the root counts as a leaf and not as an inner node).

Let us also introduce some algebraic quantities for general rooted trees, which will come in handy, when dealing with related combinatorial factors:

- **Symmetry factor**. For any $\tau \in \mathcal{T}$ we write $s(\tau)$ for the *symmetry factor*, counting the number symmetries that leave the **unordered** tree structure unchanged. It is given by a recursive formula as follows: First, $s(\bullet) = s(\mathbf{1}) = 1$. If $\tau = [(\tau_1)^{k_1} \cdots (\tau_n)^{k_n}]$ for distinct $\tau_i \neq \mathbf{1}$, $i = 1, ..., n$, each one appearing k_i times, then

$$s(\tau) = \prod_{i=1}^{n} k_i!\, s(\tau_i)^{k_i}\,. \tag{2.28}$$

  This recursive formula is based on the fundamental fact that

$$\begin{aligned} \frac{|\mathbf{k}|!}{s(\tau)} \prod_{i=1}^{n} s(\tau_i)^{k_i} = \frac{|\mathbf{k}|!}{k_1! \cdots k_n!} &= \frac{|\mathbf{k}|!}{\mathbf{k}!} \\ &= \binom{|\mathbf{k}|}{k_1} \binom{|\mathbf{k}| - k_1}{k_2} \cdots \binom{|\mathbf{k}| - (k_1 + \cdots + k_{n-2})}{k_{n-1}} \end{aligned} \tag{2.29}$$

  counts the number of possibilities to arrange the trees τ_i to form τ, or more generally, the number of arrangements of $|\mathbf{k}|$ elements of n families, each of size k_i. Here $\mathbf{k}$ denotes the the multi–index $(k_1, \ldots, k_n)$.

  Furthermore, the symmetry factor $s(\tau)$ equals the size of the symmetry group associated with τ, hence the name. The symmetry group is given by the family of automorphisms $\phi : \mathcal{V}(\tau) \to \mathcal{V}(\tau)$ such that

$$\phi(\mathfrak{o}) = \mathfrak{o} \quad \text{and} \qquad \{v, u\} \in \mathcal{E}(\tau) \quad \Leftrightarrow \quad \{\phi(v), \phi(u)\} \in \mathcal{E}(\tau)\,, \tag{2.30}$$

  see also [But08, p. 154].

- **Tree factorial.** For the empty tree $\mathbf{1}$ we define $\mathbf{1}! := 1$ and, inductively, for

a tree $\tau = [\tau_1 \cdots \tau_n]$ we define $\tau!$ as

$$\tau! := |\tau|\, \tau_1! \cdots \tau_n! \,. \tag{2.31}$$

where $|\tau| = 1 + |\tau_1| + \cdots |\tau_n|$ is the total number of vertices of the tree τ. It is easy to see that the tree factorial of a linear tree of n vertices, i.e. there is no branching and n vertices lie on a line, is equal to $n!$. Thus, the notion of tree factorial generalises the usual notion of a factorial.

τ	$h^{(\tau)}(x)$	$s(\tau)$	$\tau!$	$\lvert\tau\rvert$	$\alpha(\tau)$
$\mathbf{1}$	x	1	1	0	1
•	$h(x)$	1	1	1	1
	$h'(x)h(x)$	1	2	2	1
	$h''(x)h(x)^2$	2	3	3	1
	$h'(x)^2h(x)$	1	6	3	1
	$h^{(3)}(x)h(x)^3$	6	4	4	1
	$h''(x)h'(x)h(x)^2$	1	8	4	3
	$h''(x)h'(x)h(x)^2$	2	12	4	1
	$h'(x)^3h(x)$	1	24	4	1

Table 2.1: Examples of rooted trees and their corresponding quantities, up to $|\tau| \leqslant 4$. For the definition of the symmetry factor and the tree factorial, recall (2.28) and (2.31). On the other hand, the elementary differential $h^{(\tau)}(x)$ will be introduced below in (2.35) and the coefficient α at the end of this section.

Let us conclude the section by giving an example of how fruitful the introduction of the tree factorial and symmetry factor are, when expressing combinatorics associated to rooted trees. Let $\tau \in \mathcal{T}$, we say $\mathfrak{e} : \mathcal{V}(\tau) \to \{1, \ldots, |\tau|\}$ is an **ascending labelling of** τ if

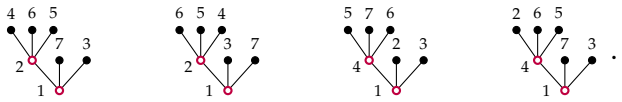
- $\mathfrak{e}$ is a bijection,

- and $\mathfrak{e}(v) > \mathfrak{e}(\mathfrak{p}(v))$ for every $v \in \mathcal{V}(\tau) \setminus \mathfrak{o}$. In other words, any path from the root to a leaf is labelled in an ascending order.

In particular, given an ascending labelling $\mathfrak{e}$, the root is always labelled by $\mathfrak{e}(\mathfrak{o}) = 1$. We say that two ascending labellings $\mathfrak{e}, \mathfrak{f}$ are equivalent $\mathfrak{e} \sim \mathfrak{f}$, if there exists an automorphism $\phi : \mathcal{V}(\tau) \to \mathcal{V}(\tau)$ satisfying (2.30) such that

$$\mathfrak{e}(v) = \mathfrak{f}(\phi(v)), \quad \forall v \in \mathcal{V}(\tau).$$

For example, consider the following labellings:



The first three labellings are ascending, with the first two being in fact equivalent. We denote the set of all ascending labellings (up to action of the symmtry group by the equivalence relation $\sim$) of a rooted tree $\tau \in \mathcal{T}$ by $\mathfrak{E}(\tau)$. It's cardinality can be expressed conveniently in terms of its tree factorial and symmetry factor.

Lemma 2.2.1. *For any $\tau \in \mathcal{T}$, we have*

$$\alpha(\tau) := |\mathfrak{E}(\tau)| = \frac{|\tau|!}{\tau!\, s(\tau)}.$$

The proof of the lemma can be found in [But08, Theorem 305A].

2.2.2 Associating trees to nested derivatives

In this section, we embed rooted trees into the analysis of ordinary differential equations (ODEs). We are mainly interested in series representations of the solutions to ODEs of the form

$$\dot{x} = h(x), \quad x_0 = a \in \mathbf{R}, \tag{2.32}$$

for some function $h : \mathbf{R} \to \mathbf{R}$. At this point let us assume that both h and the solution $(x_t)_{t \geqslant 0}$ admit locally a unique series expansion, given by their Taylor

series. Hence, the integral representation of (2.32) reads

$$\begin{aligned} x_t &= a + \int_0^t \sum_{n=0}^{\infty} \frac{(x_s - a)^n}{n!} h^{(n)}(a)\, \mathrm{d}s \\ &= a + \int_0^t \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=1}^{\infty} \frac{s^k}{k!} x_0^{(k)} \right)^n h^{(n)}(a)\, \mathrm{d}s\,, \end{aligned} \tag{2.33}$$

where in the last step we replaced x_s by its Taylor series. Here and throughout the rest of the thesis, we write

$$x_t^{(k)} = \frac{\mathrm{d}^k}{\mathrm{d}^k r} x_r \Big|_{r=t} \qquad \text{or} \qquad h^{(n)}(a) = \frac{\mathrm{d}^n}{\mathrm{d}^n x} h(x) \Big|_{x=a},$$

for the k^{th} (or n^{th}) derivative of a function. From (2.33), we deduce that it suffices to control derivatives of x at time zero. Now we make the useful observation, thanks to the information encoded in the differential equation (2.32) and application of the chain rule, that

$$\begin{aligned} \dot{x}_r &= x_r^{(1)} = h(x_r)\,, \\ \ddot{x}_r &= x_r^{(2)} = h'(x_r)\dot{x}_r = h'(x_r)h(x_r)\,, \\ & \quad x_r^{(3)} = h''(x_r)h(x_r)^2 + h'(x_r)^2 h(x_r)\,, \\ & \quad x_r^{(4)} = h^{(3)}(x_r)h(x_r)^3 + 4h''(x_r)h'(x_r)h(x_r)^2 + h'(x_r)^3 h(x_r)\,. \end{aligned}$$

Hence, derivatives of arbitrary order of the solution x can be expressed in terms of higher-order derivatives of h. However, it will be difficult to keep track of the iterative application of the chain rule for higher order terms. At this point, we use rooted trees as a guiding aid, that will even turn out more fruitful than expected, see Lemma 2.2.2 below.

Let us begin by considering the tree • containing a single vertex, which represents the function $h(x)$. Now, differentiating $h(x)$, there is only a single function we can apply the derivative to (a single vertex), which yields

$$h(x) \mapsto \frac{\mathrm{d}}{\mathrm{d}t} h(x) = h'(x)\dot{x} = h'(x)h(x) = x^{(2)} \quad \text{represented as} \quad \bullet \mapsto \overset{\bullet}{\circ},$$

where we chose the only available vertex (representing an occurrence of $h(x)$) and attached a new vertex to it (taking the derivative). This increases the degree (number of derivatives) of the original vertex, whereas the new leaf represents an occurrence of $\dot{x} = h(x)$ by the chain rule. Note that at this stage

the colour of vertices in the tree may be ignored, as both leaves and inner vertices play the same role.

Let us also formulate the remaining terms, up to order four, in terms of rooted trees. The first non-trivial case is $x^{(3)}$ for which

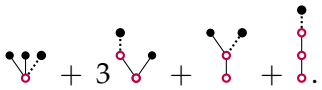
$$h'(x)h(x) \mapsto h''(x)h(x)^2 + h'(x)^2 h(x) = x^{(3)}$$

is represented in terms of

In words: Using the product rule we have to take the derivative of either term in the product $h'(x)h(x)$. Thus, we either attach a new vertex to $\bullet$ or to $\circ$ in  , corresponding to deriving $h'(x)$ or $h(x)$, respectively. Lastly, for $x^{(4)}$ we represent

$$h^{(3)}(x)h(x)^3 + 4h''(x)h'(x)h(x)^2 + h'(x)^3 h(x)$$

in terms of



(2.34)

Here we split the second term on the left-hand side into two seperate trees on the right-hand side. Both trees represent the same product of derivatives, because the number of vertices having a specific degree agrees. Similarly, it is possible to follow a procedure using multi-indices instead of trees, in which case the two terms are combined into a single one, see Faà di Bruno's formula [Fra78].

That the successive application of differential operators can be expressed in terms of rooted trees, as presented above, was first observed by Cayley [Cay57] and a century later refined by Merson [Mer57], with a focus on numerical analysis. In fact, the relation between higher-order derivatives and trees can be made more accurate by also expressing the occurring coefficients, cf. (2.34), in terms of their corresponding trees [But63]. For this, we have to introduce a last quantity:

- **Tree (or elementary) differential.** For a smooth function h, we define, inductively, its *tree differential* (also called *elementary differential*) $h^{(\tau)} : \mathbf{R} \to \mathbf{R}$ as follows: For all $a \in \mathbf{R}$, for the empty tree $\mathbf{1}$ we define $h^{(\mathbf{1})}(a) := a$ and

for the tree $\bullet$, $h^{(\bullet)}(a) := h(a)$. Now, for $\tau = [\tau_1 \cdots \tau_n]$ we define inductively

$$h^{(\tau)}(a) := h^{(n)}(a) \prod_{j=1}^{n} h^{(\tau_j)}(a) \,, \quad \text{for all} \quad a \in \mathbf{R} \,. \tag{2.35}$$

Unravelling the recursive definition of the elementary differential using an inductive argument, we can write

$$h^{(\tau)}(x_t) = \prod_{v \in \mathcal{V}(\tau)} h^{(\deg_\uparrow(v))}(x_t) \,, \tag{2.36}$$

where $\deg_\uparrow(v)$ counts the number of edges connected to v, that lead deeper into the tree, away from the root (which is equivalent to the number of descendants). More precisely,

$$\deg_\uparrow(v) := \big|\{\{v, u\} \in \mathcal{E}(\tau) \,:\, \mathfrak{p}(v) \neq u\}\big| \,,$$

with the convention that $\deg_\uparrow(\mathfrak{o}_\tau) = \deg(\mathfrak{o}_\tau)$.

Now, we have all the tools at hand to express higher-order derivatives of the solution x of the ODE (2.32) in terms of rooted trees.

Lemma 2.2.2. *Let $h \in C^n(\mathbf{R})$, $n \in \mathbf{N}$, and let $(x_t)_{t\in[0,t^*)}$ solve the differential equation (2.32) on $[0, t^*)$, for some $t^* > 0$. Then*

$$x_t^{(k)} = \sum_{\substack{\tau\in\mathcal{T} \\ |\tau|=k}} \alpha(\tau) h^{(\tau)}(x_t) = \sum_{\substack{\tau\in\mathcal{T} \\ |\tau|=k}} \frac{k!\, h^{(\tau)}(x_t)}{\tau!\, s(\tau)} \,, \quad \forall k \leqslant n+1 \,,\ t \in [0, t^*) \,,$$

where $\alpha(\tau)$ was defined in Lemma 2.2.1.

The identity was made precise in [But63], see also [But08, Theorem 311C]. We give a proof for completeness.

Proof. Let $t^* > 0$ such that $(x_t)_{t\in[0,t^*)}$ solves the differential equation (2.32) on $[0, t^*)$. First, consider $\tau = \bullet$ for which the claimed identity holds, since $\dot{x}_t = h(x_t)$ and $s(\bullet) = \bullet! = 1$. Now, assume the identity holds for any $k \leqslant n$, $n \in \mathbf{N}$. Then, by the induction hypothesis

$$x_t^{(n+1)} = \sum_{\substack{\tau\in\mathcal{T} \\ |\tau|=n}} \frac{n!}{\tau!\, s(\tau)} \frac{\mathrm{d}}{\mathrm{d}t} h^{(\tau)}(x_t) \,. \tag{2.37}$$

Note that by definition of the elementary differential, $h^{(\tau)}(x_t)$ is a product of $|\tau|$ derivatives (including 0^{th}-order), cf. (2.36), and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}h^{(\tau)}(x_t) = h(x_t)\sum_{v\in\mathcal{V}(\tau)} h^{(\deg_\uparrow(v)+1)}(x_t)\prod_{u\in\mathcal{V}(\tau)\setminus\{v\}} h^{(\deg_\uparrow(u))}(x_t)\,.$$

Hence, together with (2.37) and Lemma 2.2.1, we can write

$$x_t^{(n+1)} = \sum_{\substack{\tau\in\mathcal{T}\\|\tau|=n}}\sum_{\mathfrak{e}\in\mathfrak{E}(\tau)}\sum_{v\in\mathcal{V}(\tau)} h(x_t)h^{(\deg_\uparrow(v)+1)}(x_t)\prod_{u\in\mathcal{V}(\tau)\setminus\{v\}} h^{(\deg_\uparrow(u))}(x_t)\,, \tag{2.38}$$

where $\mathfrak{E}(\tau)$ denotes the set of ascending labellings, which was introduced above Lemma 2.2.1. In fact, each summand is an elementary differential for a tree $\tau' \in \mathcal{T}$ with $n+1$ nodes, where we added a new leaf neighbouring the node v, cf. (2.36).

It is only left to rewrite the above sum in terms of trees of size $n+1$. However, a tree of size $n+1$ can arise from various trees of size n, by grafting a new vertex onto the tree, for example

To keep track of such multiplicites, let us define the index set over the sums on the right-hand side of (2.38) as the set of labelled rooted trees (of size n) with a marked vertex

$$\mathcal{LT}^{(n)}_{\mathrm{mrk}} := \{(\tau, v, \mathfrak{e}) \,:\, v\in\mathcal{V}(\tau)\,,\ \mathfrak{e}\in\mathfrak{E}(\tau) \text{ for some } \tau\in\mathcal{T} \text{ with } |\tau| = n\}\,,$$

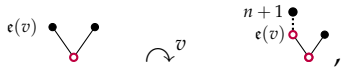
furthermore, we write the set of labelled trees of size n in terms of

$$\mathcal{LT}^{(n)} := \{(\tau, \mathfrak{e}) \,:\, \mathfrak{e}\in\mathfrak{E}(\tau) \text{ for some } \tau\in\mathcal{T} \text{ with } |\tau| = n\}\,.$$

Then the (simplified) grafting operator is the mapping from $\mathcal{LT}^{(n)}_{\mathrm{mrk}}$ into $\mathcal{LT}^{(n+1)}$ via

$$(\tau, v, \mathfrak{e}) \mapsto {}^{\mathrm{n+1}}\bullet \curvearrowright^{v} \tau_{\mathfrak{e}} =: (\tau', \mathfrak{f})\,. \tag{2.39}$$

Here, τ' is the tree arising from τ by attaching a new leaf u to v, and $\mathfrak{f}$ is the labelling that extends $\mathfrak{e}$ with the additional component $u \mapsto n+1$. For example,



with $n = 3$.

In order to conclude the proof, we have to show that the grafting operator $\curvearrowright^{\cdot}$, (2.39), defines a bijection between marked, labelled trees $\mathcal{LT}^{(n)}_{\text{mrk}}$ and labelled trees $\mathcal{LT}^{(n+1)}$. Consider $(\tau', \mathfrak{f}) \in \mathcal{LT}^{(n+1)}$ and write the labelled tree where we removed the node with label $n+1$ and the corresponding edge (note that the node is necessarily a leaf) by

$$\left(\tau, \mathfrak{p}\left(\mathfrak{f}^{-1}(n+1)\right), \mathfrak{f}\big|_{\mathcal{V}(\tau)}\right) \in \mathcal{LT}^{(n)}_{\text{mrk}}, \tag{2.40}$$

with $\mathfrak{p}\left(\mathfrak{f}^{-1}(n+1)\right)$ denoting the parent of the node with label $n+1$. We can see that the generated tree (2.40) maps to $(\tau', \mathfrak{f})$ under the grafting $\curvearrowright^{\cdot}$. This also yields injectivity, because the grafting $\curvearrowright^{v}$ only acts locally on a tree at the marked node v.

Overall, we can perform a change of variable in the summation in (2.38) from $\mathcal{LT}^{(n)}_{\text{mrk}}$ to $\mathcal{LT}^{(n+1)}$, which yields

$$x_t^{(n+1)} = \sum_{\substack{\tau \in \mathcal{T} \\ |\tau| = n+1}} \sum_{\mathfrak{f} \in \mathfrak{E}(\tau)} h^{(\tau)}(x_t) = \sum_{\substack{\tau \in \mathcal{T} \\ |\tau| = n+1}} \frac{(n+1)!}{\tau!\, s(\tau)} h^{(\tau)}(x_t)\,.$$

In the last step, we applied once more Lemma 2.2.1 since the elementary differential does not depend on the chosen labelling. This concludes the proof. □

2.2.3 The Butcher series

Using Lemma 2.2.2, the Taylor series of x_t can be written (formally) in terms of trees and their quantities introduced in (2.28), (2.31) and (2.35):

$$B_h(t, a) := \sum_{\tau \in \mathcal{T}} \frac{h^{(\tau)}(a)}{\tau!\, s(\tau)} t^{|\tau|}\,. \tag{2.41}$$

This formal power series over rooted trees (2.41) was introduced by Butcher [But63, Display (11)] and named **B-series** or **Butcher series** in subsequent work by Hairer and Wanner [HW74]. As for the Taylor series, one expects that the B-series agrees with the solution x_t, for t small enough. The content of the following lemma is that the new order of summation does not change the value of the series and that the Butcher series indeed represents the solution of the ODE.

Lemma 2.2.3. *Assume the solution $(x_t)_{t\geqslant 0}$ of the ODE (2.32) exists and is locally analytic around zero with radius t^*. Then, for all $t \in [0, t^*)$*

$$x_t = B_h(t, a) \in \mathbf{R}\,,$$

i.e. the solution x_t is represented by its B-series (2.41) whenever it admits a power series in t centred at zero.

Proof. Let $t^\star \in (0, \infty]$ such that $(x_t)_{t\in[0,t^\star)}$ has an analytic extension on the complex plane $\mathbf{C}$, i.e. $t^\star$ is the radius of convergence. Then, for $t \in [0, t^\star)$

$$x_t = \sum_{k=0}^{\infty} \frac{t^k}{k!} x_0^{(k)}\,, \tag{2.42}$$

with the series converging absolutely [MH99, Theorem 3.2.7]. Now, due to the absolute convergence of the series and Lemma 2.2.2, we can switch to a summation over rooted trees which yields

$$x_t = \sum_{k=0}^{\infty} \sum_{\substack{\tau\in\mathcal{T}\\ |\tau|=k}} \frac{h^{(\tau)}(x_0)}{\tau!\, s(\tau)} t^{|\tau|} = B_h(a, t)\,.$$

This concludes the proof. □

For t sufficently small, analyticity of h implies analyticity of the solution x_t, which is then represented by the B-series (2.41).

Proposition 2.2.4. *Let $h : \mathbf{R} \to \mathbf{R}$ be a function that is analytic on $B(a, R)$, $R \in (0, \infty]$, then the unique solution $(x_t)_{t\geqslant 0}$ to the ODE (2.32) is locally described by the B-series (2.41). More precisely,*

$$x_t = B_h(t, a)\,, \quad \textit{for all } \; t \in \left[0, \tfrac{R}{2\,\overline{M}}\right) \,, \quad \textit{with } \overline{M} := \sup_{|x-a|\leqslant R} |h(x)|\,.$$

A proof of the proposition can be found in [HNW93, pp. 46], see also [Gub10, Theorem 5.1]. We repeat the argument here for completeness.

Proof. Uniqueness of the equation follows for example from [HNW93, Theorem I.7.4]. The second part of the statement is a consequence of the method of majorants: Assuming that h is analytic on $B(a, R)$, Cauchy's inequality [MH99, Theorem 2.4.7] yields for all $n \in \mathbf{N}$

$$|h^{(n)}(a)| \leqslant \frac{n!\,\overline{M}}{R^n} = \overline{M} R \frac{\mathrm{d}^n}{\mathrm{d}r^n} (R - r)^{-1}\Big|_{r=0} = g^{(n)}(0)\,, \tag{2.43}$$

where $g(r) := \overline{M}R(R-r)^{-1}$ and $\overline{M} := \sup_{|x-a|\leqslant R} |h(x)|$.

First, the differential equation

$$\dot{r} = \frac{\overline{M}\,R}{R-r} = g(r)\,, \quad r_0 = 0\,,$$

has the explicit solution

$$r_t = R\left(1 - \sqrt{1 - \tfrac{2\overline{M}t}{R}}\right),$$

which is analytic on $B(0, \frac{R}{2\overline{M}})$, thus, admits a power series representation around 0. On the other hand, by Lemma 2.2.2,

$$\frac{1}{k!}\left|x_0^{(k)}\right| = \sum_{\substack{\tau\in\mathcal{T}\\|\tau|=k}} \frac{|h^{(\tau)}(a)|}{\tau!\,s(\tau)} \leqslant \sum_{\substack{\tau\in\mathcal{T}\\|\tau|=k}} \frac{g^{(\tau)}(0)}{\tau!\,s(\tau)} = \frac{1}{k!}r_0^{(k)}\,, \quad \text{for all } k\in\mathbf{N}\,,$$

because every product term in the elementary differential (2.35) can be estimated using (2.43). This yields absolute convergence of the Taylor series for x_t, since

$$\sum_{k=1}^{\infty} \frac{t^k}{k!}\left|x_0^{(k)}\right| \leqslant r_t = \sum_{k=1}^{\infty} \frac{t^k}{k!} r_0^{(k)} < \infty\,,$$

for every $t \in [0, \frac{R}{2\overline{M}})$. The statement of the proposition follows from Lemma 2.2.3. □

Remark 2.2.5. *Note that Proposition 2.2.4 is far from being optimal and may only yield a sub-interval on which the solution $(x_t)_{t\geqslant 0}$ is analytic. Consider for example the differential equation*

$$\dot{x} = x^3\,, \quad x_0 = 1\,. \tag{2.44}$$

The function $h(x) = x^3$ is entire, thus, Proposition 2.2.4 yields that the solution is analytic on the interval $[0, \hat{t})$ with

$$\hat{t} := \sup_{R>0} \frac{R}{2(1+R)^3} = \frac{2}{27}\,.$$

However, (2.44) has the explicit solution

$$x_t = \frac{1}{\sqrt{1-2t}}\,, \quad \textit{for all } \; t\in[0,\tfrac{1}{2})\,, \tag{2.45}$$

which is analytic on the interval $[0, \frac{1}{2})$, that is strictly larger than the one derived

from Proposition 2.2.4. Hence, x_t agrees with its B-series $B_h(t,1)$ for any $t \in [0, \frac{1}{2})$, following Lemma 2.2.3.

2.2.4 A brief note on rough paths

It is worth mentioning that Butcher's original work [But63, But72] influenced fields and works far beyond the original scope. Butcher started in the field of numerical analysis, namely building a general theory of Runge-Kutta methods. However, the thorough study of the trees appearing unveiled a finer structure in terms of a Hopf algebra formalism, which reached much farther. For example, the same rooted trees emerged in non-commutative geometry and also quantum field theory, where they were used for the removal of subdivergences in Feynman diagrams. We refer to the survey articles [Bro04, But18] for a detailed overview. Finally, rooted trees made their debut in stochastic analysis in the seminal work [Gub10], in the context of rough path theory.

From the point of view of probability theory, the theory of rough paths, introduced in [Lyo98], allows for a deterministic analysis of stochastic differential equation. Starting point is the replacement of the driving signal in (2.32) with a (rough) stochastic process $W = (W_t)_{t \geqslant 0} \in \mathcal{C}^\alpha([0,\infty))$, $\alpha \in (0,1)$, e.g. (fractional) Brownian motion. This yields the SDE

$$\mathrm{d}X_t = h(X_t)\,\mathrm{d}W_t\,, \qquad X_0 = a\,, \tag{2.46}$$

with h a smooth function. Because the solution X should look like W on small scales, one expects that

$$X_t = \sum_{\substack{\tau \in \mathcal{T} \\ |\tau| \leqslant N}} \frac{h^{(\tau)}(X_s)}{s(\tau)} \mathbb{W}_{s,t}^{(\tau)} + O\big(|t-s|^{(N+1)\alpha}\big)\,, \tag{2.47}$$

following the same steps as in the case of the ODE (2.32). Here, $\mathbb{W}^{(\tau)}$ denotes iterated stochastic integrals with respect to W, which replace the role of $\frac{t^{|\tau|}}{\tau!}$ in the previous section. For example

$$\mathbb{W}_{s,t}^{(\bullet)} = \int_s^t \mathrm{d}W_r = W_t - W_s \qquad \text{and} \qquad \mathbb{W}_{s,t}^{(\bullet\!\bullet)} = \int_s^t \int_s^r \mathrm{d}W_{r'}\,\mathrm{d}W_r\,, \tag{2.48}$$

see [Gub10, Display (10)] for their explicit definition.

In order to proceed in a pathwise fashion, we have to interpret the stochastic integrals in (2.48) for a fixed realisation $(W_t(\omega))_{t\geqslant 0}$, which is apriori not possible. For simplicity, assume that W is given by a standard Brownian motion, then the integrals (2.48) can be interpreted both in the Itô or Stratonovich sense (and many more), giving rise to different solutions of the SDE (2.46). This really is a manifestation of the fact that the "integral-map" is continuous, in fact well-defined, only if $W \in \mathcal{C}^\alpha([0,\infty))$, $\alpha > \frac{1}{2}$ [You36]. Clearly, Brownian motion just about fails this property, allowing for different interpretations of the stochastic integrals.

The idea of rough path theory is to lift the driving signal W to an element $\mathbb{W}$ in an extended space (the so called rough path space equipped with a suitable topology), by postulating the values of iterated integrals (2.48). In other words, the sole probabilistic component is the choice of integration in (2.48). Together with (2.47), this allows to construct the integral

$$\int_0^t h(X_s)\,\mathrm{d}W_s$$

as a limit of higher-order Riemann sums [Lyo98][2]. Here N has to be chosen large enough such that $(N+1)\alpha > 1$. Hence, in the case of Brownian motion we have $N = 2$ and giving meaning to a single (non-trivial) iterated integral suffices, cf. (2.48). Moreover, by setting up a fixed point problem on the rough path space, uniqueness follows and the solution map $\mathbb{W} \mapsto X$ can be proven to be continuous, see for instance [Gub10, Theorem 8.8].

2.3 Wild expansion of SPDEs with additive noise

After having seen how tree expansions are useful to encode solutions of ODEs and SDEs, we shall now proceed to the infinite dimensional setting. Throughout this section we will focus on SPDEs with additive noise of the form

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \mathfrak{m}\, u_\varepsilon + h(u_\varepsilon) + \Xi^\varepsilon\,, \quad u_\varepsilon(0,\cdot) = 0\,, \tag{2.49}$$

where h is a real analytic functions, satisfying $h(0) = h'(0) = 0$, and $\mathfrak{m} \in \mathbf{R}$. At this point, consider a noise Ξ of either the form

$$\Xi^\varepsilon(t,x) = \lambda_\varepsilon\, p_{\varepsilon^2} \star \xi(t,\cdot)(x) \quad \text{or} \quad \Xi^\varepsilon(t,x) = \lambda_\varepsilon\, p_{\varepsilon^2} \star \eta(x)\delta_0(t)\,, \tag{2.50}$$

[2] This was also explained beautifully in a course by Hendrik Weber on Singular SPDEs, see `https://drive.google.com/file/d/1s4XX--zGP7JfKgBxn-eoSZupqiM6QIFH/view`.

with $\lambda_\varepsilon \geqslant 0$ a real parameter.[3] Here, $\star$ denotes convolution in the space-variable. Thus, either Ξ is given by a mollified space-time white noise or a mollified white noise forcing at $t = 0$. We will write $\Xi = \xi$ or $\Xi = \delta_0\, \eta$, when distinguishing between the two cases. Let us also recall that $P_t = \exp\left(\frac{t}{2}\Delta\right)$ is the heat semigroup on $\mathbf{R}^d$:

$$\begin{aligned} P_t\varphi(x) &= p_t \star \varphi(x)\,, \\ p_t(x) &= \frac{1}{(2\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{2t}\right) \mathbb{1}_{[0,\infty)}(t)\,, \quad \forall (t,x) \in \mathbf{R} \times \mathbf{R}^d\,. \end{aligned} \tag{2.51}$$

Similarly, the heat semigroup with non-trivial potential $\mathfrak{m}$ is given by $P_t^{(\mathfrak{m})} := e^{\mathfrak{m}\, t} P_t$ with associated kernel $p_t^{(\mathfrak{m})}(x) := e^{\mathfrak{m}\, t} p_t(x)$.

Because h is a real analytic function satisfying $h(0) = h'(0) = 0$, the mild formulation of (2.49) reads

$$\begin{aligned} u_\varepsilon(t,x) &= \tree{I}_\varepsilon(t,x) + \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x) \sum_{n=2}^{\infty} \frac{u_\varepsilon(s,y)^n}{n!} h^{(n)}(0)\, \mathrm{d}y\, \mathrm{d}s \\ &= \tree{I}_\varepsilon(t,x) + \sum_{n=2}^{\infty} \frac{h^{(n)}(0)}{n!} P^{(\mathfrak{m})} * u_\varepsilon^n(t,x)\,, \end{aligned} \tag{2.52}$$

where we introduced the tree notation

$$\tree{I}_\varepsilon(t,x) := \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x) \Xi^\varepsilon(s,y)\, \mathrm{d}y\, \mathrm{d}s\,, \tag{2.53}$$

which is the solution of the linearisation of the SPDE (2.49), with $h = 0$. Expression (2.53) should be interpreted as follows: We assign the time-space variable (t,x) to the root, the time-space variable (s,y) as well as the weight $\Xi^\varepsilon(s,y)$ to the leaf and the kernel $p_{t-s}^{(\mathfrak{m})}(x-y)$ to the connecting edge. Finally, we integrate over the variables associated to all nodes except the root. In other words, the edge represents a time-space convolution between the heat kernel and the weight of the leaf, evaluated at the time-space variables assigned to the root.

The integral $\tree{I}_\varepsilon$ is the first step in an approximation of the solution of u_ε. Viewing the right-hand side of (2.52) as a function of u_ε, we can perform an iterative expansion, such as a Picard iteration. However, due to the non-linear nature of the equation, the stochastic integrals appearing become

[3] In some cases the scaling parameter λ_ε is only well-defined on an interval $(0,\delta)$, e.g. $\lambda_\varepsilon \sim \left(\log \frac{1}{\varepsilon}\right)^{-1/2}$. In the following, we will assume $\varepsilon \in (0,\delta)$, with the (arbitrary) choice $\delta = \frac{1}{2}$.

overwhelming quickly, both in number and complexity.

Example 2.3.1. Consider the equation

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \mathfrak{m}\, u_\varepsilon - (u_\varepsilon)^3 + \Xi^\varepsilon\,, \quad u_\varepsilon(0,\cdot) = 0\,, \tag{2.54}$$

i.e. $h(x) = -x^3$. Note that this is the non-linearity of the Allen-Cahn equation (Chapter 3) or the dynamic φ^4 model (1.20). Its mild formulation reads

$$u_\varepsilon(t,x) = \; \bullet\!\!-\!\!\circ_\varepsilon(t,x) - \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x)(u_\varepsilon(s,y))^3 \,\mathrm{d}y\,\mathrm{d}s\,. \tag{2.55}$$

Performing the next step in the Picard iteration, we insert (2.53) in place of u_ε into the right-hand side of (2.55), which yields

$$\bullet\!\!-\!\!\circ_\varepsilon(t,x) - P^{(\mathfrak{m})} * \left(\bullet\!\!-\!\!\circ_\varepsilon\right)^3 (t,x)\,. \tag{2.56}$$

Building on the identity (2.53), we represent the cubic power in terms of a branched tree as

$$\Yright_\varepsilon(s,y) := \left(\bullet\!\!-\!\!\circ_\varepsilon(s,y)\right)^3\,, \tag{2.57}$$

where we have glued together three copies of the tree $\bullet\!\!-\!\!\circ$ at a common root. At this point, we notice that we represent occurrences of Ξ^ε by $\bullet$ and kernels $p^{(\mathfrak{m})}$ by straight edges, which are linked by vertices that correspond to space-time convolutions (excluding the root). This motivates us to finally write

$$(2.56) =: \; \bullet\!\!-\!\!\circ_\varepsilon(t,x) - \Yright_\varepsilon(t,x)\,. \tag{2.58}$$

Both terms lie in a finite inhomogeneous Wiener chaos and we are able to control them with tools from stochastic analysis. Of course, we can continue this expansion at will, for example the next iteration yields the expression

$$\bullet\!\!-\!\!\circ_\varepsilon(t,x) - P^{(\mathfrak{m})} * \left(\bullet\!\!-\!\!\circ_\varepsilon - \Yright_\varepsilon\right)^3 (t,x)\,,$$

and by iterating the procedure we obtain a formal expression which is hoped to represent the solution of (2.54) arbitrarily well.

Any such iterative procedure leads to a formal power expansion in the noise variable Ξ^ε. In the following, we will use a recursive tree expansion, often referred to as Wild expansion, instead of performing an iterative Picard

expansion. This type of expansion was popularised in the context of stochastic PDEs by the work [Hai13], and is attributed to the work of Wild [Wil51].

Definition 2.3.2 (Wild expansion)**.** *The Wild expansion of* (2.52) *is the formal series over rooted trees*

$$\sum_{\tau\in\mathcal{T}\setminus\{\mathbf{1}\}} X_\varepsilon^\tau\,, \tag{2.59}$$

with X_ε^τ, $\tau\in\mathcal{T}\setminus\{\mathbf{1}\}$, defined recursively as follows:

- *First, define $X_\varepsilon^\bullet(t,x)$ to be the solution of the linearised equation*

$$\partial_t X_\varepsilon^\bullet = \frac{1}{2}\Delta X_\varepsilon^\bullet + \mathfrak{m}\, X_\varepsilon^\bullet + \Xi^\varepsilon\,, \quad X_\varepsilon^\bullet(0,\cdot) = 0\,. \tag{2.60}$$

- *For $\tau\in\mathcal{T}\setminus\{\mathbf{1},\bullet\}$ of the form $\tau = [(\tau_1)^{k_1}\cdots(\tau_n)^{k_n}]$ for distinct $\tau_i\in\mathcal{T}$'s, $i=1,\dots,n$, each one appearing k_i times, we define X_ε^τ to be the solution of*

$$\partial_t X_\varepsilon^\tau = \frac{1}{2}\Delta X_\varepsilon^\tau + \mathfrak{m}\, X_\varepsilon^\tau + \frac{h^{(|\mathbf{k}|)}(0)}{\mathbf{k}!}\prod_{i=1}^n (X_\varepsilon^{\tau_i})^{k_i}\,, \quad X_\varepsilon^\tau(0,\cdot) = 0\,. \tag{2.61}$$

where we used the multi-index notation $|\mathbf{k}| = \sum_{i=1}^n k_i$ and $\mathbf{k}! = k_1!\cdots k_n!$.

In other words, every term X_ε^τ can be expressed as an iterated stochastic integral over "lower order" terms:

$$X_\varepsilon^\tau(t,x) = \int_0^t\int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x)\frac{h^{(|\mathbf{k}|)}(0)}{\mathbf{k}!}\prod_{i=1}^n (X_\varepsilon^{\tau_i}(s,y))^{k_i}\,\mathrm{d}y\,\mathrm{d}s\,, \tag{2.62}$$

where the integral is to be interpreted in the sense of (2.16). We will make this connection more precise in the next section. Note the discrepancy between $\frac{h^{(|\mathbf{k}|)}(0)}{|\mathbf{k}|!}$ in (2.62) and the Taylor coefficient in (2.52), which is due to considering unordered trees: The difference of the two coefficients

$$\frac{|\mathbf{k}|!}{\mathbf{k}!} = \frac{|\mathbf{k}|!}{k_1!\cdots k_n!}$$

counts the number of possibilities to graft the tree τ from the given τ_i's, see (2.29).

As we are interested in approximating the solution u_ε using this expansion, it will be useful to introduce a truncated version which only considers finitely many terms.

Definition 2.3.3 (Truncated Wild expansion)**.** *We define the* ***truncated Wild expansion of order*** *N to be*

$$u_\varepsilon^N(t,x) := \sum_{\tau\in\mathcal{T}^N} X_\varepsilon^\tau(t,x)\,, \qquad \forall (t,x)\in(0,\infty)\times\mathbf{R}^d\,,$$

with $\mathcal{T}^N := \{\tau\in\mathcal{T}\setminus\{\mathbf{1}\} : |\tau|\leqslant N\}$.

The truncated Wild expansion agrees in the limit with a formal Picard iteration. Notably, it has the advantage of still being closely associated to the original equation, allowing for (some) control on the imposed error. More precisely, since every term X_ε^τ solves a PDE (2.61), their sum u_ε^N solves the differential equation

$$\partial_t u_\varepsilon^N = \frac{1}{2}\Delta u_\varepsilon^N + \mathfrak{m}\, u_\varepsilon^N + \mathrm{P}_{h,0}^{(N)}(u_\varepsilon^N) + R_\varepsilon^N + \Xi_\varepsilon\,, \quad u_\varepsilon^N(0,\cdot) = 0\,, \tag{2.63}$$

where $\mathrm{P}_{h,0}^{(N)}$ is the N^{th}-order Taylor polynomial of h around zero

$$\mathrm{P}_{h,0}^{(N)}(u_\varepsilon^N(t,x)) = \sum_{m=1}^{N}\frac{h^{(m)}(0)}{m!}\sum_{\tau_1,\dots,\tau_m\in\mathcal{T}^N}\prod_{i=1}^{m} X_\varepsilon^{\tau_i}(t,x)\,,$$

and R_ε^N an error term. The error term R_ε^N is explicit and only depends on trees in the "boundary" of $\mathcal{T}^N$

$$\begin{aligned} R_\varepsilon^N &:= -\mathrm{P}_{h,0}^{(N)}(u_\varepsilon^N) + \sum_{\substack{\tau\in\mathcal{T}^N\\ \tau=[(\tau_1)^{k_1}\cdots(\tau_n)^{k_n}]}}\frac{h^{(|\mathbf{k}|)}(0)}{\mathbf{k}!}\prod_{i=1}^{n}\left(X_\varepsilon^{\tau_i}\right)^{k_i} \\ &= -\sum_{m=1}^{N}\frac{h^{(m)}(0)}{m!}\sum_{\substack{\tau_1,\dots,\tau_m\in\mathcal{T}^N\\ [\tau_1\cdots\tau_m]\notin\mathcal{T}^N}}\prod_{i=1}^{m} X_\varepsilon^{\tau_i}\,. \end{aligned}$$

In the second equality, we used that

$$\begin{aligned} \sum_{\substack{\tau\in\mathcal{T}^N\\ \tau=[(\tau_1)^{k_1}\cdots(\tau_n)^{k_n}]}}\frac{h^{(|\mathbf{k}|)}(0)}{\mathbf{k}!}\prod_{i=1}^{n}\left(X_\varepsilon^{\tau_i}\right)^{k_i} &= \sum_{m=1}^{N}\frac{h^{(m)}(0)}{m!}\sum_{\substack{\tau\in\mathcal{T}^N\\ \tau=[(\tau_1)^{k_1}\cdots(\tau_n)^{k_n}]\\ |\mathbf{k}|=m}}\frac{m!}{\mathbf{k}!}\prod_{i=1}^{n}\left(X_\varepsilon^{\tau_i}\right)^{k_i} \\ &= \sum_{m=1}^{N}\frac{h^{(m)}(0)}{m!}\sum_{\substack{\tau_1,\dots,\tau_m\in\mathcal{T}^N\\ [\tau_1\cdots\tau_m]\in\mathcal{T}^N}}\prod_{i=1}^{m} X_\varepsilon^{\tau_i}\,, \end{aligned}$$

which is a consequence of (2.29). At this stage, we note that the choice of in-

dex set $\mathcal{T}^N$ is rather arbitrary and in general one may consider any sequence of (increasing) subsets of trees that exhaust all of $\mathcal{T}$, as $N \to \infty$. Depending on the structure of the equation, other choices might be more appropriate. For example, when h is a monomial of degree $n \in \mathbf{N}$, it suffices to consider trees in $\mathcal{T}_n$ (2.27), see also Chapter 3.

Remark 2.3.4. *The concept of Wild expansions can be implemented more generally to include multiplicative SPDEs or SPDEs with a non-linearity acting on ∇u, instead of u. In the latter case, the principal example is the KPZ equation (1.19), for which (2.61) would instead read*

$$\partial_t X_\varepsilon^\tau = \frac{1}{2}\Delta X_\varepsilon^\tau + a(\tau)\langle \nabla X_\varepsilon^{\tau_1}, \nabla X_\varepsilon^{\tau_2}\rangle\,, \quad X_\varepsilon^\tau(0,\cdot) = 0\,,$$

where $\tau = [\tau_1\ \tau_2] \in \mathcal{T}_2$, with $a(\tau) = \mathbb{1}_{\tau_1=\tau_2} + 2\mathbb{1}_{\tau_1\neq\tau_2} \in \{1,2\}$. See also [Hai13], where a truncated Wild expansion (in terms of ordered trees) was used to solve the one-dimensional KPZ equation.

2.3.1 Iterated stochastic integrals, trees and Wiener chaoses

Throughout this section, $\mathcal{H}$ will denote the Gaussian Hilbert space generated by Ξ, see Corollary 2.1.5 and Remark 2.1.19. Whenever we write stochastic integrals with respect to Ξ, such as (2.9), they are to be interpreted in the sense of L^2-isometries, cf. Lemma 2.1.8.

Our goal is to represent terms in the Wild expansion by iterated Wiener integrals, indexed by trees, as for example in (2.53) and (2.57). With this in mind, let us first define for any tree $\tau \in \mathcal{T} \setminus \{\mathbf{1}, \bullet\}$ the Stratonovich integral

$$\begin{aligned}\tau_\varepsilon(t,x) &:= \mathring{\mathbb{I}}_{\ell(\tau)}\big(K_{\tau,\varepsilon}^{t,x}\big) \\ &= \int_{(\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau)}} K_{\tau,\varepsilon}^{t,x}(s_{\mathcal{L}(\tau)}, y_{\mathcal{L}(\tau)}) \prod_{v\in\mathcal{L}(\tau)} \Xi(\,\mathrm{d}s_v, \,\mathrm{d}y_v)\,,\end{aligned} \tag{2.64}$$

cf. Definition 2.1.14, with

$$K_{\tau,\varepsilon}^{t,x}(s_{\mathcal{L}(\tau)}, y_{\mathcal{L}(\tau)}) := \int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \widehat{K}_{\tau,\varepsilon}^{t,x}(s_{\mathcal{V}(\tau)\setminus\mathfrak{o}}, y_{\mathcal{V}(\tau)\setminus\mathfrak{o}})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,, \tag{2.65}$$

where $D_t := [0,t]\times\mathbf{R}^d$ and

$$\widehat{K}_{\tau,\varepsilon}^{t,x}(s_{\mathcal{V}(\tau)\setminus\mathfrak{o}}, y_{\mathcal{V}(\tau)\setminus\mathfrak{o}}) := \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)}) \tag{2.66}$$

$$\prod_{u\in\mathcal{L}(\tau)} \lambda_\varepsilon\, e^{\mathfrak{m}\,(s_{\mathfrak{p}(u)}-s_u)}\, p_{s_{\mathfrak{p}(u)}-s_u+\varepsilon^2}(y_u - y_{\mathfrak{p}(u)})\,.$$

We recall that $\mathfrak{p}(u)$ denotes the unique parent of u. Note that we are **not** integrating over the root variable of the tree τ, which is assigned to the point (t,x), but only integrate over all other inner vertices of τ. In other words, every edge in τ corresponds to an occurrence of $p^{(\mathfrak{m})}$ and every leaf corresponds to a noise Ξ^ε, recall Example 2.3.1. Note that the condition $s_{\mathfrak{p}(u)} \geqslant s_u$ is implicit in the definition of the heat kernel (2.51). At this stage, it is not clear that the function $K^{t,x}_{\tau,\varepsilon}$ is contractable, cf. Definition 2.1.13, and thus if the Stratonovich integral is well defined. We will verify this in Lemma 2.3.8 below.

Remark 2.3.5. *In* (2.49) *we assume zero initial data. However, non-trivial intial data u_0 may be included in the term Ξ by replacing it in* (2.64) *with*

$$\delta_0(\,\mathrm{d}s_v)\,u_0(y)\;\mathrm{d}y_v + \Xi(\,\mathrm{d}s_v,\,\mathrm{d}y_v)\,,$$

assuming u_0 is regular enough.

The terms of the Wild expansion (2.62) are related to the Wiener–Stratonovich integrals (2.64) via a combinatorial factor. For example, the formulation (2.64) allows us to write

$$X_\varepsilon \;=\; \frac{h^{(3)}(0)}{3!}\;\; \varepsilon\,, \qquad X_\varepsilon \;=\; \frac{h^{(3)}(0)^2}{3!\,2!}\;\; \varepsilon\,.$$

The coefficients in the display above, and in the formulation of the Wild expansion (2.61), are a consequence of considering unordered trees: In classical expansions, such as the Picard iteration, integrals are multiplied in different orders but give rise to the same terms, cf. (2.23) or (2.25). It will be important to obtain a precise expression for these coefficients, which we shall do now. To this end, let us define recursively for $\tau = [\tau_1 \cdots \tau_n]$

$$c_\tau := h^{(n)}(0) \prod_{i=1}^n c_{\tau_i}\,, \tag{2.67}$$

with $c_\bullet := 1$. Then the following identity holds:

Lemma 2.3.6. *For any $\tau \in \mathcal{T}\setminus\{\mathbf{1}\}$, we have for every $\varepsilon \in (0,\frac{1}{2})$ and $(t,x) \in (0,\infty)\times\mathbf{R}^d$*

$$X^\tau_\varepsilon(t,x) = \frac{c_\tau}{s(\tau)}\,[\tau]_\varepsilon(t,x)\,,$$

with the Wild expansion term (2.62) *on the left and the Wiener–Stratonovich integral* (2.64) *on the right, represented by the planted tree* $[\tau]$.

Note that c_τ (2.67) defines an elementary differential (not necessarily associated to a function as in (2.35)). We note that an identity of similar nature has been derived in [BCCH20], to determine coefficients of renormalisation constants in the context of (subcritical) singular SPDEs.

At this point, we adapt the concept of contractions from Section 2.1.2 to the kernels considered in (2.65). Unlike the stochastic integrals indexed by the initial tree, the stochastic integral indexed by a contracted tree will lie in the homogeneous Wiener chaos, whose order is given by the number of uncontracted leaves. We will see below that one can recover the integral associated to the original tree by summing over all integrals indexed by possible contractions, in analogy to Definition 2.1.14. Let us now be more precise and start with the definition of a **contracted tree**.

Definition 2.3.7. *For any $\tau \in \mathcal{T}$ we define a **contraction** to be a subset of (unordered) pairs of leaves $\kappa \subset \binom{\mathcal{L}(\tau)}{2}$ of the tree τ, such that every $v \in \mathcal{L}(\tau)$ lies in at most one element of κ. We define the corresponding set of contractions by*

$$\mathcal{K}(\tau) := \left\{ \kappa \subset \binom{\mathcal{L}(\tau)}{2} \,:\, \kappa \text{ is a contraction of } \tau \right\} . \tag{2.68}$$

Moreover, we will denote a tree $\tau = (\mathcal{V}, \mathcal{E})$ that is being contracted according to a contraction κ by τ_κ:

$$\tau_\kappa := (\mathcal{V}, \mathcal{E} \cup \kappa) ,$$

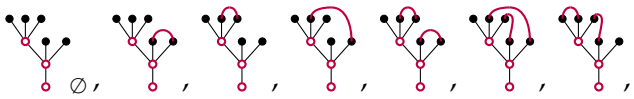
*and call this a κ-contracted tree or simply a **contracted tree**. We will also denote by $\mathcal{L}(\tau_\kappa)$ the set of leaves of τ_κ, namely, the subset of leaves $\mathcal{L}(\tau)$ which are not included in the contraction κ.*

If all leaves of τ are contracted via κ, we will call κ a *complete contraction* and τ_κ a *completely contracted* tree. If this is not the case, we will often (but not necessarily always) talk of *partial contractions* and *partially contracted* trees.

At this point, we see a clear advantage of using rooted trees to represent stochastic integrals, as we can visualise contractions $\kappa \in \mathcal{K}(\tau)$ pictorially by pairing leaves in the tree τ. For example, the possible contractions of the tree

(2.69)

are (up to symmetries)



(2.70)

where we denoted the tree without any contraction with a subscript $\emptyset$ to emphasize the empty contraction, and distinguish it from the Stratonovich integral. Then, given a contraction $\kappa \in \mathcal{K}(\tau)$, we write the kernel $K^{t,x}_{\tau,\varepsilon}$ (2.65) contracted with respect to κ as

$$K^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{L}(\tau_\kappa)}, y_{\mathcal{L}(\tau_\kappa)}) := \int_{D_t^{\mathcal{V}(\kappa)}} K^{t,x}_{\tau,\varepsilon}(s_{\mathcal{L}(\tau)}, y_{\mathcal{L}(\tau)}) \prod_{\{u,u'\}\in\kappa} \mathbb{E}\left[\Xi(s_u, y_u)\,\Xi(s_{u'}, y_{u'})\right] \mathrm{d}y_{\mathcal{V}(\kappa)}\,\mathrm{d}s_{\mathcal{V}(\kappa)}\,, \tag{2.71}$$

where $\mathcal{V}(\kappa)$ denotes the vertices included in the contraction κ and

$$\mathbb{E}\left[\Xi(s_u, y_u)\,\Xi(s_{u'}, y_{u'})\right] := \begin{cases} \delta_{s_u}(s_{u'})\,\delta_{y_u}(y_{u'}) & , \quad \text{if } \Xi = \xi\,, \\ \delta_0(s_u)\delta_0(s_{u'})\,\delta_{y_u}(y_{u'}) & , \quad \text{if } \Xi = \delta_0\,\eta\,. \end{cases}$$

Now, to a contracted tree τ_κ we associate the Wiener–Itô integral

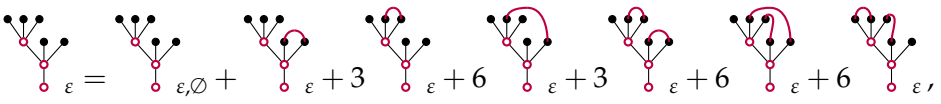
$$\tau_{\kappa,\varepsilon}(t,x) := \mathrm{I}_{\ell(\tau_\kappa)}\left(K^{t,x}_{\tau_\kappa,\varepsilon}\right)\,. \tag{2.72}$$

To lighten notation, we will sometimes drop the coordinates t, x that indicate the time-space coordinates of the root, if the explicit indication is not necessary. With this definition, $\tau_{\kappa,\varepsilon}$ lies in the homogeneous Wiener chaos of order $\ell(\tau_k) = |\mathcal{L}(\tau_\kappa)|$.

By Definition 2.1.14 of the Stratonovich integral, each stochastic integral associated to a tree τ_ε has the decomposition

$$\tau_\varepsilon = \sum_{\kappa\in\mathcal{K}(\tau)} \tau_{\kappa,\varepsilon}\,, \tag{2.73}$$

with $\tau_{\kappa,\varepsilon}$ defined in (2.72). Thus, in the example of the trees in (2.69) and (2.70), we have



where the right-hand side corresponds to the homogeneous stochastic inte-

grals indexed by the contracted trees in (2.70). Here we have taken into account multiplicities of homogeneous components due to equivalent contractions, e.g. the contractions

are all different, however, correspond to the same stochastic integrals.

Up to this point, we have not given a justification that the stochastic integrals (2.64) or (2.72) are indeed well defined, which we shall do now.

Lemma 2.3.8. *Let $\tau \in \mathcal{T}$ and $T > 0$. Then uniformly over all $(t,x) \in (0,T] \times \mathbf{R}^d$ and $\varepsilon \in (0, \frac{1}{2})$*

$$\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \leqslant C \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}} \right)^{\ell(\tau)}, \qquad \text{for all } \kappa \in \mathcal{K}(\tau)\,,$$

where $C = C(\mathfrak{m}, T, i(\tau), \ell(\tau))$ is a finite positive constant, independent of ε. In particular, the kernels $K^{t,x}_{\tau_\kappa,\varepsilon}$ are contractable, cf. Definition 2.1.13, and the Stratonovich integral (2.64) is well defined as a sum of Itô integrals.

The proof of Lemma 2.3.8 can be found at the end of the section. We stress that the bound obtained is far from being optimal: The divergence of the ε^{-2}-terms, as $\varepsilon \to 0$, is in general not compensated (even by the usual choices of $\lambda_\varepsilon \to 0$ in the weak coupling limit, e.g. (1.31)). That the kernels are indeed uniformly bounded in ε is left as a case by case analysis, which strongly depends on the model at hand and the choice λ_ε being made.

In view of Lemma 2.3.8, we are ready to prove Lemma 2.3.6.

Proof of Lemma 2.3.6. First, the identity holds for $\tau = \bullet$ since

$$X^\bullet_\varepsilon(t,x) = \lambda_\varepsilon \int_{D_t} e^{\mathfrak{m}(t-s)} p_{t-s+\varepsilon^2}(y-x)\, \Xi(\mathrm{d}s, \mathrm{d}y) = [\bullet]_\varepsilon(t,x)\,.$$

Now, assume the claimed identity holds for any tree τ satisfying $|\tau| \leqslant N$, $N \in \mathbf{N}$. Let $\tau = [(\tau_1)^{k_1} \cdots (\tau_m)^{k_m}]$ with distinct $\tau_i \in \mathcal{T}$'s, $i = 1, ..., m$, each one appearing k_i times, such that $|\tau| = N + 1$. Then, by (2.62) and the induction

hypothesis

$$X_\varepsilon^\tau(t,x) = \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x) \frac{h^{(|\mathbf{k}|)}(0)}{\mathbf{k}!} \prod_{i=1}^m \left(\frac{c_{\tau_i}}{s(\tau_i)} [\tau_i]_\varepsilon(s,y) \right)^{k_i} \mathrm{d}y \, \mathrm{d}s$$
$$= \frac{c_\tau}{s(\tau)} \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x) \prod_{i=1}^m \big([\tau_i]_\varepsilon(s,y) \big)^{k_i} \mathrm{d}y \, \mathrm{d}s \,, \tag{2.74}$$

where we used (2.67) and (2.28) in the second equality. Because every kernel $K^{s,y}_{[\tau_i],\varepsilon}$ is contractable by Lemma 2.3.8, we introduce

$$K^{s,y}_{\tau,\varepsilon} := \bigotimes_{i=1}^m \left(K^{s,y}_{[\tau_i],\varepsilon} \right)^{\otimes k_i} \in L^2\big((\mathbf{R} \times \mathbf{R}^d)^{\mathcal{L}(\tau)} \big) \,, \tag{2.75}$$

where we recall the $\otimes$-notation from (2.13). Thus, we can use the product identity for the Stratonovich integral, Lemma 2.1.15, to write

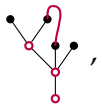
$$\prod_{i=1}^m \big([\tau_i]_\varepsilon(s,y) \big)^{k_i} = \mathring{\mathrm{I}}_{\ell(\tau)}\big(K^{s,y}_{\tau,\varepsilon} \big) = \tau_\varepsilon(s,y) \,,$$

where used that $\ell(\tau) = k_1\ell(\tau_1) + \cdots + k_m\ell(\tau_m)$. Together with Lemma 2.1.18, this finally allows to rewrite (2.74) in terms of

$$X_\varepsilon^\tau(t,x) = \frac{c_\tau}{s(\tau)} \mathring{\mathrm{I}}_{\ell(\tau)} \left(\int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x) K^{s,y}_{\tau,\varepsilon}(\cdot) \, \mathrm{d}y \, \mathrm{d}s \right)$$
$$= \frac{c_\tau}{s(\tau)} \, [\tau]_\varepsilon(t,x) \,,$$

where we used that $K^{t,x}_{[\tau],\varepsilon}$ is contractable, by another application of Lemma 2.3.8. □

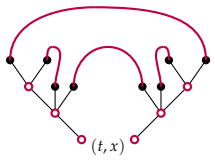
Finally, we pay our debt and state a proof of Lemma 2.3.8. While being heavy notationally, the proof is elementary and best explained in a picture. Consider for example the kernel $K^{t,x}_{\tau_\kappa,\varepsilon}$ associated to the (contracted) tree

$\tau_\kappa =$  ,

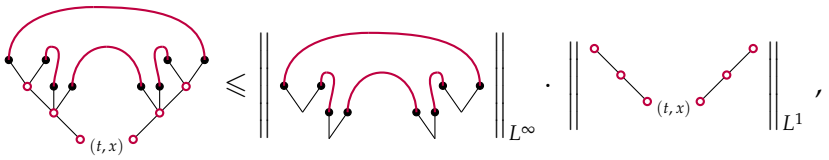
then we can represent the L^2-norm of the kernel diagrammatically in terms

of

$$\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} = \quad , \tag{2.76}$$



where both time-space points associated to the respective roots equal (t,x). See Section 3.2 for a detailed exposition of this fact. Again, identifications of time-space points are represented by pairwise contractions (purple edges). Black edges are interpreted as the convolution kernels consisting of $p^{(\mathfrak{m})}$'s, cf. (2.71). To give a brief (diagrammatic) summary of the proof: We will extract the kernels associated to black edges which are connected to purple edges using a crude "L^1–L^∞-estimate":



The first term on the right-hand side is bounded since we considered a mollified noise Ξ^ε. Whereas the second term can be explicitly estimated by "integrating out" all remaining kernels using the time-space points associated to inner nodes.

Proof of Lemma 2.3.8. Let us first consider the case $\Xi = \xi$, and assume $\mathfrak{m} > 0$. Explicitly writing out (2.71), we have

$$\begin{aligned}
&K^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{L}(\tau_\kappa)}, y_{\mathcal{L}(\tau_\kappa)}) \\
&= \int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \int_{D_t^{\mathcal{V}(\kappa)}} \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)}) \\
&\qquad \prod_{\{u,u'\}\in\kappa} \delta_{s_u}(s_{u'})\,\delta_{y_u}(y_{u'}) \\
&\qquad \prod_{u\in\mathcal{L}(\tau)} \lambda_\varepsilon\, e^{\mathfrak{m}\,(s_{\mathfrak{p}(u)}-s_u)} p_{s_{\mathfrak{p}(u)}-s_u+\varepsilon^2}(y_u - y_{\mathfrak{p}(u)})\,\mathrm{d}y_{\mathcal{V}(\kappa)}\,\mathrm{d}s_{\mathcal{V}(\kappa)}\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}} \\
&\leqslant \int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)}) \\
&\qquad \int_{[0,t]^{\mathcal{V}(\kappa)}} \prod_{\{u,u'\}\in\kappa} \lambda_\varepsilon^2\, e^{2\mathfrak{m}\,(t-s_u)} p_{s_{\mathfrak{p}(u)}+s_{\mathfrak{p}(u')}-2s_u+2\varepsilon^2}(y_{\mathfrak{p}(u)} - y_{\mathfrak{p}(u')})\delta_{s_u}(s_{u'})\,\mathrm{d}s_{\mathcal{V}(\kappa)} \\
&\qquad \prod_{u\in\mathcal{L}(\tau_\kappa)} \lambda_\varepsilon\, e^{\mathfrak{m}\,(t-s_u)} p_{s_{\mathfrak{p}(u)}-s_u+\varepsilon^2}(y_u - y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,,
\end{aligned} \tag{2.77}$$

where we separated out the space-time points assosciated to leaves contained in κ and applied the semigroup property of the heat-kernel when integrating out the spatial variables $y_{\mathcal{V}(\kappa)}$, and used the fact that $s_{\mathfrak{p}(u)} \leqslant t$.

Now, the second line of the right-hand side of (2.77) can be upper bounded as follows

$$\begin{aligned}
&\int_{[0,t]^{\mathcal{V}(\kappa)}} \prod_{\{u,u'\}\in\kappa} \lambda_\varepsilon^2\, e^{2\mathfrak{m}\,(t-s_u)} p_{s_{\mathfrak{p}(u)}+s_{\mathfrak{p}(u')}-2s_u+2\varepsilon^2}(y_{\mathfrak{p}(u)} - y_{\mathfrak{p}(u')})\delta_{s_u}(s_{u'})\,\mathrm{d}s_{\mathcal{V}(\kappa)} \\
&\leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\int_0^t e^{2\mathfrak{m}\,(t-s)}\,\mathrm{d}s\right)^{|\kappa|} \leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{|\kappa|}, \qquad (2.78)
\end{aligned}$$

where we used that $p_{s+2\varepsilon^2}(x) \leqslant p_{s+2\varepsilon^2}(0) \leqslant p_{2\varepsilon^2}(0)$ in the first inequality. Hence,

$$\begin{aligned}
&K^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{L}(\tau_\kappa)}, y_{\mathcal{L}(\tau_\kappa)}) \\
&\leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{|\kappa|} \int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)}) \qquad (2.79) \\
&\qquad \prod_{u\in\mathcal{L}(\tau_\kappa)} \lambda_\varepsilon\, e^{\mathfrak{m}(t-s_u)} p_{s_{\mathfrak{p}(u)}-s_u+\varepsilon^2}(y_u - y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,.
\end{aligned}$$

This yields an upper bound of the L^2-norm of the kernel, since

$$\begin{aligned}
&\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \\
&= \int_{D_t^{\mathcal{L}(\tau_\kappa)}} K^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{L}(\tau_\kappa)}, y_{\mathcal{L}(\tau_\kappa)})^2\,\mathrm{d}s_{\mathcal{L}(\tau_\kappa)}\,\mathrm{d}y_{\mathcal{L}(\tau_\kappa)} \\
&\leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{2|\kappa|} \int_{D_t^{\mathcal{L}(\tau_\kappa)}} \Bigg(\int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)}) \\
&\qquad \prod_{u\in\mathcal{L}(\tau_\kappa)} \lambda_\varepsilon\, e^{\mathfrak{m}(t-s_u)} p_{s_{\mathfrak{p}(u)}-s_u+\varepsilon^2}(y_u - y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\Bigg)^2 \mathrm{d}s_{\mathcal{L}(\tau_\kappa)}\,\mathrm{d}y_{\mathcal{L}(\tau_\kappa)} \\
&\leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{2|\kappa|} \int_{[0,t]^{\mathcal{L}(\tau_\kappa)}} \prod_{u\in\mathcal{L}(\tau_\kappa)} \frac{\lambda_\varepsilon^2\, e^{2\mathfrak{m}(t-s_u)}}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\,\mathrm{d}s_{\mathcal{L}(\tau_\kappa)} \\
&\qquad \left(\int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}} \prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u - y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\right)^2,
\end{aligned}$$

where we applied (2.79) in the first inequality. In the second inequality, we once more integrated over spatial variables associated to $\mathcal{L}(\tau_\kappa)$, while keeping in mind the semigroup property of the heat kernel, and again used

$p_{s+2\varepsilon^2}(x) \leqslant p_{2\varepsilon^2}(0)$. Overall, repeating the same steps as in (2.78), we have

$$\begin{aligned}
&\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \\
&\leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{\ell(\tau)} \\
&\quad\times\left(\int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}}\prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u-y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\right)^2,
\end{aligned} \tag{2.80}$$

since $2|\kappa| + |\mathcal{L}(\tau_\kappa)| = \ell(\tau)$. The integral inside the square-bracket can now be evaluated by integrating over spatial variables from leaves towards the root (note that every inner node has exactly one edge (heat kernel) connecting it to its parent):

$$\begin{aligned}
&\int_{D_t^{\mathcal{I}(\tau)\setminus\mathfrak{o}}}\prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} p^{(\mathfrak{m})}_{s_{\mathfrak{p}(u)}-s_u}(y_u-y_{\mathfrak{p}(u)})\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}}\,\mathrm{d}y_{\mathcal{I}(\tau)\setminus\mathfrak{o}} \\
&\leqslant \int_{[0,t]^{\mathcal{I}(\tau)\setminus\mathfrak{o}}}\prod_{u\in\mathcal{I}(\tau)\setminus\mathfrak{o}} e^{\mathfrak{m}(t-s_u)}\,\mathrm{d}s_{\mathcal{I}(\tau)\setminus\mathfrak{o}} \leqslant \left(\frac{1}{\mathfrak{m}}e^{\mathfrak{m}\,t}\right)^{i(\tau)-1}.
\end{aligned} \tag{2.81}$$

Lastly, combining (2.80) and (2.81) yields

$$\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\frac{1}{2\mathfrak{m}}e^{2\mathfrak{m}\,t}\right)^{\ell(\tau)}\left(\frac{1}{\mathfrak{m}}e^{\mathfrak{m}\,t}\right)^{2(i(\tau)-1)}.$$

When considering $\mathfrak{m} \leqslant 0$, we use the estimate $\int_0^t \mathbb{1}_{s_{\mathfrak{p}(u)}\geqslant s_u} e^{\mathfrak{m}(s_{\mathfrak{p}(u)}-s_u)}\,\mathrm{d}s_u \leqslant t$, which yields

$$\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\right)^{\ell(\tau)} t^{2(i(\tau)-1)+\ell(\tau)},$$

following the same lines of the proof above.

Lastly, let us comment on the case when $\Xi = \delta_0\eta$, in which all time-variables associated to leaves are evaluated at 0. Then the equivalent estimate to (2.78) reads

$$\begin{aligned}
&\int_{[0,t]^{\mathcal{V}(\kappa)}}\prod_{\{u,u'\}\in\kappa}\lambda_\varepsilon^2\,e^{2\mathfrak{m}\,(t-s_u)}\,p_{s_{\mathfrak{p}(u)}+s_{\mathfrak{p}(u')}-2s_u+2\varepsilon^2}(y_{\mathfrak{p}(u)}-y_{\mathfrak{p}(u')})\delta_0(s_{u'})\delta_0(s_u)\,\mathrm{d}s_{\mathcal{V}(\kappa)} \\
&\quad\leqslant\left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}\int_0^t e^{2\mathfrak{m}\,(t-s)}\delta_0(s)\,\mathrm{d}s\right)^{|\kappa|} \leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}}e^{2\mathfrak{m}\,t}\right)^{|\kappa|},
\end{aligned}$$

which eventually yields, for $\mathfrak{m} > 0$,

$$\|K^{t,x}_{\tau_\kappa,\varepsilon}\|^2_{L^2((\mathbf{R}\times\mathbf{R}^d)^{\mathcal{L}(\tau_\kappa)})} \leqslant \left(\frac{\lambda_\varepsilon^2}{(4\pi\varepsilon^2)^{\frac{d}{2}}} e^{2\mathfrak{m}t}\right)^{\ell(\tau)} \left(e^{\mathfrak{m}t}\right)^{2(i(\tau)-1)} .$$

In the case $\mathfrak{m} \leqslant 0$, the exponential factor can be neglected. This concludes the proof. □

2.3.2 From the Wild expansion to a Butcher series

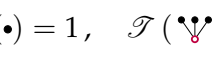
This sections main result is to establish a link between the Wild expansion, cf. Definition 2.3.2, and a corresponding Butcher series (2.41). First we need to introduce an operator on trees.

Definition 2.3.9. *Let $n \in \mathbf{N}$. We call the map*

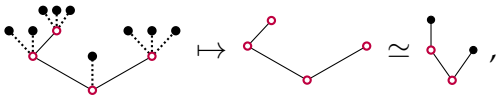
$$\mathscr{T} : \mathcal{T} \setminus \{\mathbf{1}\} \to \mathcal{T} , \qquad \mathscr{T}(\tau) = \tau , \tag{2.82}$$

*the **trimming** operator, where $\mathscr{T}(\tau)$ is the tree that is spanned by the inner nodes of τ, i.e. $\mathscr{T}$ "cuts off" all the leaves and their attached edges.*

To lighten later notation, we have also used the chromatic notation by which $\tau = \mathscr{T}(\tau)$, for example

$\bullet = \mathscr{T}(\bullet) = 1$, $\mathscr{T}($  $) = \bullet$, and $\mathscr{T}\Big($  $\Big) =$  .

The reader should have the following pictorial description of $\mathscr{T}$ in mind



(2.83)

where we again coloured leaves in the tree on the right-hand side in black, according to our convention.

Having the definition of trimmed trees at hand, we can recover the coefficients of the Butcher series, with non-linearity h, by summing over all suitable coefficients of Wild expansion terms.

Lemma 2.3.10. *Let $\sigma \in \mathcal{T}$, then*

$$\frac{h^{(\sigma)}(1)}{s(\sigma)} = \sum_{\substack{\tau\in\mathcal{T} \\ \tau=\mathscr{T}(\tau)=\sigma}} \frac{c_\tau}{s(\tau)} ,$$

with c_τ defined in (2.67). Recall (2.28) and (2.35) for the definition of the symmetry factor $s(\tau)$ and the elementary differential $h^{(\tau)}$, respectively.

To the author's best knowledge the identity in Lemma 2.3.10 was not observed before. However, it is not entirely surprising: The fundamental building block in the construction of terms in the Wild expansion is the lollipop  . Hence, the expansion can be viewed as a power series expansion in terms of  , and not the noise $\Xi^\varepsilon = \bullet$. The corresponding combinatorial factor, associated to a Wild term X_ε^τ, should thus be related to the trimmed tree $\mathscr{T}(\tau) = \tau$.

Proof of Lemma 2.3.10. The statement is true for $\sigma = \mathbf{1} \in \mathcal{T}$, since $\mathscr{T}^{-1}(\mathbf{1}) = \{\bullet\}$ with $h^{(\bullet)}(1) = h^{(\mathbf{1})}(1) = 1 = c_\bullet$ and $s(\bullet) = s(\mathbf{1})$. Moreover, for $\sigma = \bullet$ the pre-image $\mathscr{T}^{-1}(\bullet)$ consists of trees of the form

$$\tau = \quad = [\bullet^n], \qquad n \in \mathbf{N}.$$

Hence, because h is analytic and $h(0) = 0$

$$\sum_{\substack{\tau \in \mathcal{T} \\ \tau = \mathscr{T}(\tau) = \bullet}} \frac{c_\tau}{s(\tau)} = \sum_{n=1}^{\infty} \frac{c_{[\bullet^n]}}{s([\bullet^n])} = \sum_{n=1}^{\infty} \frac{h^{(n)}(0)}{n!} \cdot c_\bullet^n = h(1) = \frac{h^{(\bullet)}(1)}{s(\bullet)}.$$

Now we proceed by induction. Assume that the statement is true for all trees $\hat{\sigma} \in \mathcal{T} \setminus \{\mathbf{1}\}$ with $|\hat{\sigma}| \leqslant n$, for some given $n \in \mathbf{N}$, and let $\sigma \in \mathcal{T}$ be a tree with $|\sigma| = n + 1$, of the form $\sigma = [(\sigma_1)^{k_1} \cdots (\sigma_n)^{k_n}]$ for distinct $\sigma_i \in \mathcal{T} \setminus \{\mathbf{1}\}$, $i = 1, ..., n$, each one appearing k_i times. Just as above, trees in the pre-image $\mathscr{T}^{-1}(\sigma)$ must be of the form

$$\tau = \quad = [\tau_1 \cdots \tau_{|\mathbf{k}|} \bullet^{m-|\mathbf{k}|}], \qquad m = |\mathbf{k}|, |\mathbf{k}| + 1, \dots,$$

for some $\tau_j \in \mathcal{T} \setminus \{\mathbf{1}, \bullet\}$, $j = 1, \dots, |\mathbf{k}|$, where we recall that $|\mathbf{k}| = \sum_{i=1}^n k_i$. Because we consider unordered trees, we may assume that the trees $\tau_1, \dots, \tau_{k_1}$ satisfy $\mathscr{T}(\tau_j) = \sigma_1$, the trees $\tau_{k_1+1}, \dots, \tau_{k_1+k_2}$ satisfy $\mathscr{T}(\tau_j) = \sigma_2$ and so on. At this point it will be useful to write the grafting operator in terms of

$$\begin{aligned} \tau = [\tau_1 \cdots \tau_{|\mathbf{k}|} \bullet^{m-|\mathbf{k}|}] &= [\tau_1 \cdots \tau_{k_1}] \cdot [\tau_{k_1+1} \cdots \tau_{k_1+k_2}] \cdots [\bullet^{m-|\mathbf{k}|}] \\ &=: \overline{\tau}_1 \cdots \overline{\tau}_n \cdot \quad , \end{aligned} \tag{2.84}$$

where the right-hand side is to be interpreted as gluing together all grafted trees at **one** common root (without introducing new edges). Note that

$\mathscr{T}(\overline{\tau}_i) = [\sigma_i^{k_i}]$. In particular, we have

$$s(\tau) = (m - |\mathbf{k}|)!\, s(\overline{\tau}_1) \cdots s(\overline{\tau}_n)\,, \tag{2.85}$$

because $\tau_j \neq \tau_\ell$ whenever $\mathscr{T}(\tau_j) \neq \mathscr{T}(\tau_\ell)$. As a consequence of (2.84) and (2.85), we can write

$$\begin{aligned}\sum_{\substack{\tau\in\mathcal{T}\\ \tau=\mathscr{T}(\tau)=\sigma}} \frac{c_\tau}{s(\tau)} &= \sum_{m=|\mathbf{k}|}^{\infty} h^{(m)}(0)\, c_\bullet^{m-|\mathbf{k}|} \sum_{\substack{\tau=[\tau_1\cdots\tau_{|\mathbf{k}|}\bullet^{m-|\mathbf{k}|}]\in\mathcal{T}\\ \tau=\mathscr{T}(\tau)=\sigma}} \frac{1}{s(\tau)} \prod_{i=1}^{|\mathbf{k}|} c_{\tau_i} \\ &= \sum_{m=|\mathbf{k}|}^{\infty} \frac{h^{(m)}(0)\, c_\bullet^{m-|\mathbf{k}|}}{(m-|\mathbf{k}|)!} \prod_{i=1}^{n} \sum_{\substack{\overline{\tau}_i=[\hat{\tau}_1\cdots\hat{\tau}_{k_i}]\in\mathcal{T}\\ \mathscr{T}(\overline{\tau}_i)=[\sigma_i^{k_i}]}} \frac{1}{s(\overline{\tau}_i)} \prod_{j=1}^{k_i} c_{\hat{\tau}_j}\,,\end{aligned} \tag{2.86}$$

where $\hat{\tau}_1, \dots, \hat{\tau}_{k_i}$ play the role of $\tau_{k_1+\cdots+k_i+1}, \dots, \tau_{k_1+\cdots+k_{i+1}}$. Now, we have

$$\begin{aligned}k_i! \sum_{\substack{\overline{\tau}_i=[\hat{\tau}_1\cdots\hat{\tau}_{k_i}]\in\mathcal{T}\\ \mathscr{T}(\overline{\tau}_i)=[\sigma_i^{k_i}]}} \frac{1}{s(\overline{\tau}_i)} \prod_{j=1}^{k_i} c_{\hat{\tau}_j} &= \sum_{\substack{\overline{\tau}_i=[\hat{\tau}_1\cdots\hat{\tau}_{k_i}]\in\mathcal{T}\\ \mathscr{T}(\overline{\tau}_i)=[\sigma_i^{k_i}]}} k_i! \frac{s(\hat{\tau}_1)\cdots s(\hat{\tau}_{k_i})}{s(\overline{\tau}_i)} \prod_{j=1}^{k_i} \frac{c_{\hat{\tau}_j}}{s(\hat{\tau}_j)} \\ = \sum_{\substack{\hat{\tau}_1,\dots,\hat{\tau}_{k_i}\in\mathcal{T}\\ \mathscr{T}(\hat{\tau}_i)=\sigma_i}} \prod_{j=1}^{k_i} \frac{c_{\hat{\tau}_j}}{s(\hat{\tau}_j)} &= \Bigg(\sum_{\substack{\hat{\tau}\in\mathcal{T}\\ \mathscr{T}(\hat{\tau})=\sigma_i}} \frac{c_{\hat{\tau}}}{s(\hat{\tau})} \Bigg)^{k_i} = \left(\frac{h^{(\sigma_i)}(1)}{s(\sigma_i)} \right)^{k_i}\,,\end{aligned} \tag{2.87}$$

where the second equality is a consequence of the size of the pre-image $[\cdots]^{-1}(\overline{\tau}_i) \subset (\mathcal{T})^{k_i}$ being equal to

$$k_i! \frac{s(\hat{\tau}_1)\cdots s(\hat{\tau}_{k_i})}{s(\overline{\tau}_i)} \in \mathbf{N}\,,$$

cf. (2.29). The last equality holds by application of the induction hypothesis.

Then, (2.86) together with the chain of identities in (2.87) yields

$$\sum_{\substack{\tau\in\mathcal{T}\\ \tau=\mathscr{T}(\tau)=\sigma}} \frac{c_\tau}{s(\tau)} = \sum_{m=|\mathbf{k}|}^{\infty} \frac{h^{(m)}(0)\, c_\bullet^{m-|\mathbf{k}|}}{(m-|\mathbf{k}|)!} \prod_{i=1}^{n} \frac{1}{k_i!} \left(\frac{h^{(\sigma_i)}(1)}{s(\sigma_i)} \right)^{k_i}. \tag{2.88}$$

Thus, together with

$$\sum_{m=|\mathbf{k}|}^{\infty} \frac{h^{(m)}(0)}{(m-|\mathbf{k}|)!} c_\bullet^{m-|\mathbf{k}|} = h^{(|\mathbf{k}|)}(c_\bullet) = h^{(|\mathbf{k}|)}(1)\,,$$

(2.88) implies

$$\sum_{\substack{\tau\in\mathcal{T}\\ \tau=\mathscr{T}(\tau)=\sigma}} \frac{c_\tau}{s(\tau)} = \frac{1}{s(\sigma)} h^{(|\mathbf{k}|)}(1) \prod_{i=1}^{n} \left(h^{(\sigma_i)}(1)\right)^{k_i} = \frac{h^{(\sigma)}(1)}{s(\sigma)},$$

where we used the definition of the symmetry factor (2.28) and the elementary differential (2.35). This concludes the proof. □

Let us state a useful property of the trimming operation, namely it is a bijection between finite families of n-ary and sub-n-ary trees: For arbitrary $N \in \mathbf{N}$, let us define

$$\begin{aligned} \mathcal{T}_n^{N,\circ} &:= \{\tau \in \mathcal{T}_n : \ i(\tau) \leqslant N\} \subseteq \mathcal{T}_n\,, \\ \mathcal{T}_{\leqslant n}^{N} &:= \{\tau \in \mathcal{T}_{\leqslant n} : \ |\tau| \leqslant N\} \subseteq \mathcal{T}_{\leqslant n}\,. \end{aligned} \tag{2.89}$$

The next lemma summarises this and other properties of $\mathscr{T}$.

Lemma 2.3.11. *The following holds for any $n \in \mathbf{N}$.*

(i) *The map $\mathscr{T}$ is a bijection from $\mathcal{T}_n^{N,\circ}$ to $\mathcal{T}_{\leqslant n}^{N}$, for every $N \in \mathbf{N}$.*

(ii) *Let $\tau \in \mathcal{T}_n \setminus \{\bullet\}$ and $\tau_1, \dots, \tau_n \in \mathcal{T}_n$ such that $\tau = [\tau_1 \ \cdots \ \tau_n]$, then*

$$\tau = [\tau_1 \ \cdots \ \tau_n]\,.$$

In other words, trimming via $\mathscr{T}$ and grafting via $[\dots]$ commute.

Proof. By definition of $\mathcal{T}_n^{N,\circ}$, its image under $\mathscr{T}$ is a subset of $\mathcal{T}_{\leqslant n}^{N}$. On the other hand, for any $\sigma \in \mathcal{T}_{\leqslant n}^{N}$ we can construct a $\tau \in \mathcal{T}_n^{N,\circ}$ such that $\mathscr{T}(\tau) = \tau = \sigma$ as follows: To each node that has $n-k$ descendants, for $k \in \{1, \dots, n\}$, we append exactly k lollipops $\bullet\!\!\circ$, so that in the new tree that node has exactly n outgoing edges. The constructed tree τ lies in $\mathcal{T}_n^{N,\circ}$, since $i(\tau) = |\sigma| \leqslant N$ and every inner node of τ has exactly n descendants. Moreover, it satisfies $\mathscr{T}(\tau) = \sigma$. This concludes the proof of the first part of the statement.

In order to prove (ii), let $\tau \in \mathcal{T}_n \setminus \{\bullet\}$ and $\tau_1, \dots, \tau_n \in \mathcal{T}_n$ such that $\tau = [\tau_1 \ \cdots \ \tau_n]$. Now, because $\mathscr{T}$ does not act on the root of τ, see Definition 2.3.9, we necessarily have

$$\tau = \mathscr{T}(\tau) = [\mathscr{T}(\tau_1) \ \cdots \ \mathscr{T}(\tau_n)] = [\tau_1 \ \cdots \ \tau_n]\,.$$

In order to avoid confusion, let us discuss explicitly the case where $\tau_i = \bullet$ and thus $\tau_i = \mathbf{1}$, for some $i \in \{1, \dots, n\}$. Without loss of generality, let us assume

that $\tau_2 = \bullet$. Then, by convention of $[\dots]$, we have

$$\tau = [\tau_1 \ \cdots \ \tau_n] = [\tau_1 \ \tau_3 \ \cdots \ \tau_n] .$$

This identity propagates to $\tau = [\tau_1]$ if additionally $\tau_i = \bullet$ for all $i \in \{3, \dots, n\}$. Moreover, in the most extreme case $\tau =$  , this reduces further to $\tau = \bullet = [\mathbf{1}] = [\mathbf{1}^n] = [\bullet^n]$. □

The bijection property of $\mathscr{T}$ for n-ary trees allows us to link each term in the Wild expansion to a corresponding term in the Butcher expansion of the ODE $\dot{x} = h(x)$, whenever h is a monomial.

Corollary 2.3.12. *Let h be a monomial of degree $n \in \mathbf{N}$, then the terms of the Wild expansion for the SPDE (2.49) satisfy the following identity: For any $\tau \in \mathcal{T}_n$*

$$X_\varepsilon^\tau(t,x) = \frac{h^{(\tau)}(1)}{s(\tau)} [\tau]_\varepsilon(t,x) , \quad \forall \varepsilon \in (0, \tfrac{1}{2}) , \ (t,x) \in (0,\infty) \times \mathbf{R}^d ,$$

with the symmetry factor $s(\tau)$ and elementary differential $h^{(\tau)}$ defined in (2.28) and (2.35), respectively.

Proof. Let $\sigma \in \mathcal{T}_{\leqslant n}$, then there exists a unique $\tau \in \mathcal{T}_n$ such that $\mathscr{T}(\tau) = \tau = \sigma$, by Lemma 2.3.11. Therefore, the statement of Lemma 2.3.10 simplifies to

$$\frac{h^{(\sigma)}(1)}{s(\sigma)} = \sum_{\substack{\tau' \in \mathcal{T} \\ \mathscr{T}(\tau') = \sigma}} \frac{c_{\tau'}}{s(\tau')} = \frac{c_\tau}{s(\tau)} ,$$

because $c_{\tau'} = 0$ for every $\tau' \notin \mathcal{T}_n$, since $h^{(m)}(0) = 0$ whenever $m \neq n$, cf. (2.67). The statement of the corollary is now a consequence of Lemma 2.3.6. □

2.3.3 A brief note on regularity structures

In Section 2.2.4, we saw that finitely many iterated integrals suffice to express the solution of a SDE in a pathwise sense. This idea was pushed further to a multi-parameter setting in [Hai13] and lead to Martin Hairer's theory of regularity structures [Hai14], which has been expanded and further developed in [CH16, BCCH20, BHZ19]. In this section, we aim to give a condensed summary of the theory and its limitation to (scaling) subcritical SPDEs. For more details and a general overview, we point the reader to

[FH20, CW17, CZ21, Zam21].

Consider a SPDE of the form (2.49) that is subcritical, recall Table 1.4. The basic idea of the theory of regularity structures is an expansion as we performed in Definition 2.3.2. Subcriticality of the equation suggests that $u_\varepsilon \simeq$  on small scales. In particular, all other terms in the expansion are of higher regularity. This is both a curse and blessing for subcritical equations. On the one hand, the gain in regularity suggests that finitely many terms in the expansion suffice to give sense to (2.49). On the other hand, we will encounter objects of regularity $\alpha > 1$, which require "recentering" to make sense in this context. In particular, the latter is accountable for a lot of the complexity in regularity structures [CW17].

The expansion in regularity structures is expressed in terms of abstract symbols $\tau = (\tau)_{\tau\in\overline{\mathcal{T}}}$, where $\overline{\mathcal{T}}$ denotes an index set such as rooted trees. Usually $\overline{\mathcal{T}}$ holds a larger class of objects, including abstract monomials X^k for the "recentering".[4] To every point (t,x) one then associates a local expansion $\Pi^{(\varepsilon)}_{t,x}$, a so called **model**, which maps abstract symbols in $\overline{\mathcal{T}}$ to actual functions/distributions. For example,

$$\Pi^{(\varepsilon)}_{t,x} \, \bullet \!\!-\!\!\circ(\cdot) = P^{(\mathfrak{m})} * \xi^\varepsilon(\cdot) - P^{(\mathfrak{m})} * \xi^\varepsilon(t,x) \qquad \text{and} \qquad \Pi^{(\varepsilon)}_{t,x} X^k(\cdot) = ((\cdot)_2 - x)^k .$$

In the context of rough paths, the model $\Pi^{(\varepsilon)}_{t,x}$ corresponds to the iterated integrals $\mathbb{W}^{(\cdot)}_{t,\cdot}$. Now, in order to describe the solution of the given SPDE up to regularity $\gamma > 0$, it suffices to look at all terms in the expansion below this regularity threshold (**finitely many**), i.e.

$$u_\varepsilon(\varphi^{(\lambda)}_{t,x}) = \sum_{\substack{\tau\in\overline{\mathcal{T}} \\ |\tau|_h<\gamma}} F_\tau(t,x)\left(\Pi^{(\varepsilon)}_{t,x}\tau\right)(\varphi^{(\lambda)}_{t,x}) + O(\lambda^\gamma) ,$$

where $|\tau|_h$ measures how the associated distribution vanishes or blows up when evaluated on small scales, with $\varphi^{(\lambda)}$ a scaled test function as in (1.21). Here, F_τ are the coefficients in the expansion, some of which depend non-trivially on (t,x), to correct the "recentering". Note that in the case of the Wild expansion with vanishing initial data, Definition 2.3.2, no such recentering was implemented and the associated coefficients are constant, cf. Lemma 2.3.6.

[4] In the literature, these abstract symbols are often coloured in blue. We will do the same in order to distinguish them from iterated integrals introduced in (2.64).

Finally, taking the small ε-limit in the space of models leads to a (random) model $\hat{\Pi}_{t,x}$ and the family of (modelled) distributions

$$\hat{F}_{t,x} = \sum_{\substack{\tau \in \overline{\mathcal{T}} \\ |\tau|_h < \gamma}} F_\tau(t,x) \left(\hat{\Pi}_{t,x} \tau \right),$$

which locally describes the solution of the SPDE at hand. Notably, $\hat{\Pi}_{t,x}$ encodes, how to interpret ill-defined products in the formulation of (2.49). We stress that in most cases the small ε-limit of $\hat{\Pi}_{t,x}$ only exists after suitable renormalisation of the model.

The remainder of the procedure resembles rough path theory, in the sense that it is performed deterministically, for a fixed realisation of the model $\hat{\Pi}_{t,x}(\omega)$. Namely, the celebrated reconstruction theorem states that there exists a unique distribution $\mathcal{R}F$ which is locally described by $\hat{F}_{t,x}$, i.e.

$$\left| \mathcal{R}F(\varphi^{(\lambda)}_{t,x}) - \hat{F}_{t,x}(\varphi^{(\lambda)}_{t,x}) \right| \lesssim \lambda^\gamma,$$

uniformly over $\varphi, \lambda \in (0,1]$ and locally uniformly in (t,x).

In summary, the theory of regularity structures provides a small-scale analysis of subcritical SPDEs, locally in time. For this, only finitely many terms in the expansion are necessary. It is a general theory that makes sense of ill-defined SPDEs and allows to systematically renormalise such equations. On the other hand, considering a critical equation, infinitely many expansion-terms will have the same regularity as $\mathcal{I}_\varepsilon(\bullet)$, which does not allow for pathwise techniques as sketched above.

Chapter 3

The Allen–Cahn equation with critical initial datum

We consider the 2D Allen–Cahn equation

$$\partial_t u = \frac{1}{2}\Delta u + \mathfrak{m}\, u - u^3 \,, \quad u(0,\cdot) = \eta \,, \tag{3.1}$$

$\hat{\lambda} \geqslant 0$ and $\mathfrak{m} \in \mathbf{R}$, with initial condition given by space white noise η on $\mathbf{R}^2$. The SPDE is scaling critical and there exists no solution theory, due to the irregularity of the noise.
In this chapter, based on the work [GRZ23], we approach the SPDE (3.1) by studying its weak coupling limit and establish non-trivial Gaussian fluctuations. The result builds on treating an infinite Wild expansion of the equation, cf. Section 2.3, and on the introduction of certain combinatorial "cycle removal" estimates, which allow to control the inflicted error in the expansion. To the author's best knowledge, this is the first time an infinite expansion has been treated successfully for a non-linear critical SPDE.

The weakly coupled version of (3.1) is given in terms of

$$\partial_t u_\varepsilon = \frac{1}{2}\Delta u_\varepsilon + \mathfrak{m}\, u_\varepsilon - u_\varepsilon^3 \,, \qquad u_\varepsilon(0,\cdot) = \eta_\varepsilon \,, \tag{3.2}$$

with $\varepsilon \in (0, \frac{1}{2})$ and initial condition

$$\eta_\varepsilon := \lambda_\varepsilon\, p_{\varepsilon^2} \star \eta \,, \quad \text{with} \quad \lambda_\varepsilon := \frac{\hat{\lambda}}{\sqrt{\log \frac{1}{\varepsilon}}} \,,$$

which approximates space white noise, as $\varepsilon \to 0$. Here $\hat{\lambda} > 0$ is a coupling constant that we will later on require to be sufficiently small. Also recall that p denotes the heat kernel

$$p_t(x) = \frac{1}{2\pi t} \exp\left(-\frac{|x|^2}{2t}\right) \mathbb{1}_{[0,\infty)}(t)\,, \qquad \text{for all} \quad (t,x) \in \mathbf{R} \times \mathbf{R}^2\,.$$

Note that $p_{\varepsilon^2} \star \eta$ corresponds to a mollified white noise on scales ε, since $p_{\varepsilon^2}(\cdot) = \varepsilon^{-2} p_1(\cdot/\varepsilon)$.

Because the initial condition vanishes under the considered scaling, one expects u_ε to converge to the trivial zero-solution. Hence, we should multiply the solution by the same factor the initial data vanishes, to study non-trivial fluctuations around the zero-limit. Therefore, we define

$$\mathcal{U}_\varepsilon(t,x) := \sqrt{\log \tfrac{1}{\varepsilon}} \cdot u_\varepsilon(t,x)\,, \tag{3.3}$$

which solves the equation

$$\partial_t \mathcal{U}_\varepsilon = \frac{1}{2}\Delta \mathcal{U}_\varepsilon + \mathfrak{m}\,\mathcal{U}_\varepsilon - \big(\log \tfrac{1}{\varepsilon}\big)^{-1} \cdot \mathcal{U}_\varepsilon^3\,, \qquad \mathcal{U}_\varepsilon(0,\cdot) = \hat{\lambda}\, p_{\varepsilon^2} \star \eta\,.$$

Our main result is the following.

Theorem 3.0.1 ([GRZ23])**.** *There exists a $\hat{\lambda}_{\text{fin}} > 0$ such that for $\hat{\lambda} \in (0, \hat{\lambda}_{\text{fin}})$, $\mathfrak{m} \in \mathbf{R}$ and $T \in (0,\infty)$ satisfying*

$$\overline{\mathfrak{m}}\, T \leqslant \log \frac{\hat{\lambda}_{\text{fin}}}{\hat{\lambda}}\,, \tag{3.4}$$

where $\overline{\mathfrak{m}} := \mathfrak{m} \vee 0$, we have

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left|\mathcal{U}_\varepsilon(t,x) - \sigma_{\hat{\lambda}} \cdot v(t,x)\right|^2\right] = 0 \qquad \textit{with} \qquad \sigma_{\hat{\lambda}} := \frac{\hat{\lambda}}{\sqrt{1 + \frac{3}{\pi}\hat{\lambda}^2}}\,,$$

for all $(t,x) \in (0,T] \times \mathbf{R}^2$. Here, v denotes the solution of the linear equation

$$\partial_t v = \frac{1}{2}\Delta v + \mathfrak{m}\, v\,, \qquad v(0,\cdot) = \eta\,. \tag{3.5}$$

As a consequence of our main result, we see that the limiting statistics of the mollified SPDE (3.2) are given in terms of the linearised equation (3.5), where the non-linearity was dropped. However, the non-linearity is relevant and affects the size of the limiting fluctuations $\sigma_{\hat{\lambda}}$, which are strictly

weaker than the fluctuations of the limit when dropping the non-linearity in (3.2), see also (3.10) and the discussion below. Furthermore, it is interesting to note that, despite the restriction $\hat{\lambda} < \hat{\lambda}_{\text{fin}}$ due to technical reasons, the limit in Theorem 3.0.1 is well-defined for arbitrary $\hat{\lambda} > 0$. In particular, for $\hat{\lambda} = \sqrt{\log \frac{1}{\varepsilon}}$, we formally recover the strong coupling regime, which suggests fluctuations of order $O(1)$.

The 2D Allen–Cahn equation (3.1) falls into the regime of scaling criticality. Recalling (1.25), we have that $u^{(\lambda)}(t,x) := \lambda^{-\alpha} u(\lambda^2 t, \lambda x)$ solves

$$\partial_t u^{(\lambda)} = \frac{1}{2}\Delta u^{(\lambda)} + \lambda^2 \mathfrak{m}\, u^{(\lambda)} - \lambda^{2(1+\alpha)} (u^{(\lambda)})^3 , \quad u^{(\lambda)}(0,\cdot) \overset{d}{=} \lambda^{-\alpha-\frac{d}{2}} \eta . \tag{3.6}$$

For $\alpha \in (-1, 0)$ the factor in front of the non-linearity vanishes as $\lambda \to 0$, which corresponds to the (scaling) subcritical regime. For $\alpha = -1$ we are in the critical regime. Together with the choice $d = 2$, $\alpha = -1$ also keeps the initial condition scale invariant. In fact, we may reformulate (3.1) as a SPDE with additive forcing

$$\partial_t u = \frac{1}{2}\Delta u + \mathfrak{m}\, u - u^3 + \delta_0 \eta , \qquad u(0,\cdot) = 0 . \tag{3.7}$$

The noise $\delta_0 \eta$ lies in $\mathcal{C}^{-3-}(\mathbf{R}_+ \times \mathbf{R}^d)$ when $d = 2$, cf. Table 1.3. Notably, this agrees with the regularity of space-time white noise in $d = 4$, for which the additive equation (the dynamic φ^4_4-model (1.20)) is known to be scaling critical, cf. (1.27).

We should mention that that the Allen–Cahn equation is only locally well-posed for initial data in $C^{\alpha-}(\mathbf{R}^d)$, for $\alpha > -\frac{2}{3}$, and ill-posed otherwise [COW22]. In particular, it is not possible to make sense of (3.1) pathwise, and only stochastic cancellations make the regime $\alpha \leqslant -\frac{2}{3}$ applicable. Up to this point in time, the study of (scaling) critical SPDEs remains difficult as there is no general toolbox available for their treatment, in contrast to subcritical equations, cf. Section 2.3.3. Instead, critical SPDEs have only been treated individually, by exploiting certain characteristics of the equation at hand, as we presented in Section 1.3. We believe that our approach in analysing the **full expansion** for the critical Allen–Cahn equation, may be a first step towards analysing infinite expansions of critical singular SPDEs more generally.

Besides regarding (3.1) as a toy model for a critical singular SPDE, our interest in the dynamics is motivated by the paper [HLR23], where (3.1) was considered in $d \geqslant 2$ with a weak rescaling of the initial datum (in the sub-critical regime), namely

$$u_\varepsilon(0,\cdot) = \hat{\lambda}\, \varepsilon^{\frac{d}{2}+\alpha} p_{\varepsilon^2} \star \eta \quad \text{with} \quad \alpha \in (-1,0)\,. \tag{3.8}$$

This yields initial conditions $u_\varepsilon(0,\cdot) \in \mathcal{C}^{\alpha-}(\mathbf{R}^d)$, uniformly in $\varepsilon \in (0,\frac{1}{2})$, and as a result, one observes[1]

$$\varepsilon^{-\frac{d}{2}-\alpha} u_\varepsilon(t,x) \to \hat{\lambda} \cdot v(t,x)\,, \quad \text{in } L^2(\mathbb{P})\,, \tag{3.9}$$

where the right-hand side is simply described by the solution of linearised equation of (3.1), without seeing any influence of the non-linearity. However, we stress that the work [HLR23] focuses instead on a non-trivial limit arising at **large times**, the so called Bargmann–Fock field. The fronts formed by the dynamics of the Allen–Cahn equation are known to evolve according to mean-curvature flow, see for example [BK91, ESS92]. Thus, [HLR23] yields a natural construction of initial data for mean curvature flows, constructed from "absolute chaos" (white noise). In contrast, Theorem 3.0.1 in the present paper studies (3.1) at the critical value $\alpha = -1$ (in $d = 2$) with a weak coupling constant and non-trivial influence of the non-linearity can already be observed on scales of order one. In particular, when comparing our result with (3.9), we observe a **dampening effect of the non-linearity**, because

$$\sigma_{\hat{\lambda}} = \frac{\hat{\lambda}}{\sqrt{1+\frac{3}{\pi}\hat{\lambda}^2}} < \hat{\lambda}\,, \qquad \text{for all } \hat{\lambda} > 0\,. \tag{3.10}$$

However, in the case $\mathfrak{m} = 1$, we are only able to treat the problem at **small times**, and the large time behaviour in the flavour of [HLR23] seems not to be covered by our approach, due to technical restrictions.

Remark 3.0.2. *Note that for $\mathfrak{m} \leqslant 0$, Theorem 3.0.1 covers arbitrary large time-horizons T, for all $\hat{\lambda} \in (0, \hat{\lambda}_{\text{fin}})$, and our methods should allow for an extension to slowly growing time horizon T_ε, as $\varepsilon \to 0$.*

The explicit value $\sigma_{\hat{\lambda}}$ appears from the study of a Wild expansion of u_ε, see Definition 2.3.2, and arises in terms of a Butcher series (2.41), which in

[1] In [HLR23] the initial condition is stated for $\hat{\lambda} = 1$, we added the additional factor to allow for comparison to our result. They also consider a wider class of mollifiers.

turn allows us to link it to the solution of an ordinary differential equation. More precisely, after performing an expansion of $\mathcal{U}_\varepsilon(t,x)$ in terms of iterated integrals (represented by a certain class of rooted trees), we can identify a power series in $\hat{\lambda}$ such that

$$\mathcal{U}_\varepsilon(t,x) \sim \hat{\lambda} \left(\sum_{n=0}^{\log \frac{1}{\varepsilon}} a_n \hat{\lambda}^{2n} \right) v(t,x)\,,$$

where the approximation is to be understood in a $L^2(\mathbb{P})$-sense. The sum inside the parenthesis agrees miraculously with the Taylor expansion of the solution of the ordinary differential equation $\dot{y} = -y^3$ with $y(0) = 1$, where the variable $\frac{3\hat{\lambda}^2}{2\pi}$ plays the role of time. Thus, for $\hat{\lambda}$ small enough,

$$\lim_{\varepsilon \to 0} \sum_{n=0}^{\log \frac{1}{\varepsilon}} a_n \hat{\lambda}^{2n} = \frac{1}{\sqrt{1 + \frac{3\hat{\lambda}^2}{\pi}}}\,, \tag{3.11}$$

with the right-hand side being the solution of the aforementioned ODE. As we could see in Chapter 2, the concept of Wild expansions can be implemented more generally for SPDEs. However, to the author's best knowledge, this is the first time that the (full) Butcher series is used to study and identify the limit of a (non-linear) singular SPDE.

Let us close with some general remarks on Theorem 3.0.1.

Remark 3.0.3. *The fluctuation coefficient $\sigma_{\hat{\lambda}}$ can also be identified as the solution of the ODE*

$$\frac{\mathrm{d}}{\mathrm{d}\hat{\lambda}}\sigma = \frac{1}{\hat{\lambda}} \left(\sigma - \frac{3}{\pi}\sigma^3 \right)\,, \qquad \sigma_0 = 0\,,$$

as a function of $\hat{\lambda}$. However, this ODE does not appear naturally in our analysis. Generally, it would be interesting to link the original SPDE directly to the ODE that determines the fluctuation coefficient.

Remark 3.0.4. *In Theorem 3.0.1 we impose $\hat{\lambda}$ to be chosen small enough, since we require the power series in (3.11) to converge absolutely. However, we expect the result to hold for arbitrary choices of $\hat{\lambda}$ since $\sigma_{\hat{\lambda}}$ is well defined on all of $\mathbf{R}_+$ and does not explode.*

Remark 3.0.5. *For $d \geqslant 3$, we could have considered critical initial data of the form (3.8) with $\alpha = -1$. In this case, the equivalent statement of Theorem 3.0.1 should*

read

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big[\big|\varepsilon^{-\frac{d}{2}+1}u_\varepsilon(t,x)-\sigma_{\hat\lambda}(d)\cdot v(t,x)\big|^2\Big]=0\,,$$

with

$$\sigma_{\hat\lambda}(d):=\frac{\hat\lambda}{\sqrt{1+\left(\frac{d}{2}-1\right)^{-1}\frac{6\hat\lambda^2}{(4\pi)^{d/2}}}}\,.$$

The statement should follow almost verbatim from our arguments.

Structure of the chapter

The remainder of this chapter is structured as follows. In Section 3.1, we present the main steps of the proof of Theorem 3.0.1, while assuming the chapter's key ingredient (Proposition 3.1.2). Section 3.2 introduces the notion of v-cycles and related estimates, which are required for the the control of second moment estimates for terms in the Wild expansion. Lastly, we provide a proof of Proposition 3.1.2 in Section 3.3.

3.1 Proof steps for the main result

In Chapter 2, we introduced the Wild expansion and set the stage for iterative expansions of additive singular SPDEs. There, we learned that, on a formal level, the solution of (3.2) can be represented in terms of the Wild expansion (Definition 2.3.2) as

$$u_\varepsilon=\sum_{\tau\in\mathcal{T}_3}X_\varepsilon^\tau\,,\tag{3.12}$$

where it suffices to consider ternary trees $\mathcal{T}_3$ (2.27) only, due to the cubic nonlinearity, cf. Example 2.3.1. The linchpin of our argument, to conclude Theorem 3.0.1, is a precise control of the Wild expansion (3.12). Our asymptotic analysis builds on a precise understanding of the decomposition of the components of the Wild expansion into its homogeneous chaos terms, cf. (2.73). The goal of this detailed study will be to show that that only terms in the first chaos (and not all of them) contribute to the Gaussian limit in Theorem 3.0.1. The key result in this direction is presented in the upcoming Section 3.1.1. The proof of Theorem 3.0.1 is then carried out in Section 3.1.2.

Chapter 2 provides all the elements which allow us to discuss the proof of Theorem 3.0.1, without entering into technical details. These are deferred to later sections and require the introduction of additional tools.

3.1.1 From the Wild expansion to single-tree estimates

Let us write the truncated Wild expansion, cut off at a level $N \in \mathbf{N}$, as

$$u_\varepsilon^N := \sum_{\tau \in \mathcal{T}_3^{N,\circ}} X_\varepsilon^\tau ,$$

with $\mathcal{T}_3^{N,\circ}$ introduced in (2.89). Here we slighlty deviated from the formulation in Definition 2.3.3, however, the idea remains the same and unlike the full Wild expansion, u_ε^N is well defined as it is a finite sum. The structure of the Allen–Cahn equation, and in particular the fact that the non-linearity $-u^3$ is monotone in u, allows to circumvent direct treatment of the infinite part of the series. More precisely, in analogy to (2.63) and the discussion below, we have that u_ε^N solves

$$\partial_t u_\varepsilon^N = \frac{1}{2}\Delta u_\varepsilon^N + \mathfrak{m}\, u_\varepsilon^N - (u_\varepsilon^N)^3 + R_\varepsilon^N , \quad u_\varepsilon^N(0,\cdot) = \eta_\varepsilon , \tag{3.13}$$

where the error term R_ε^N depends only on trees at the "boundary" of $\mathcal{T}_3^{N,\circ}$:

$$R_\varepsilon^N = \sum_{\substack{\tau_1,\tau_2,\tau_3 \in \mathcal{T}_3^{N,\circ} \\ [\tau_1\, \tau_2\, \tau_3] \notin \mathcal{T}_3^{N,\circ}}} X_\varepsilon^{\tau_1} X_\varepsilon^{\tau_2} X_\varepsilon^{\tau_3} . \tag{3.14}$$

Utilising a maximum principle in combination with more structural estimates, which we will describe in detail below, we are able to control the error of the approximation as follows.

Proposition 3.1.1. *Let $\hat{\lambda} > 0$ and $T > 0$ satisfy*

$$\hat{\lambda} e^{\overline{\mathfrak{m}}\, T} < \frac{1}{10\sqrt{C}} , \tag{3.15}$$

with $\overline{\mathfrak{m}} = \mathfrak{m} \vee 0$ and $C \in (0,\infty)$ the constant defined in (3.82). Then uniformly over all $(t,x) \in (0,T] \times \mathbf{R}^2$, $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and $N \leqslant \log \frac{1}{\varepsilon}$

$$\sqrt{\log \tfrac{1}{\varepsilon}} \left\| u_\varepsilon^N(t,x) - u_\varepsilon(t,x) \right\|_{L^2(\mathbb{P})} \leq \frac{C_0}{\log \frac{1}{\varepsilon}} \frac{\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, t}\right)^N}{\varepsilon} , \tag{3.16}$$

with $C_0 = C_0(T, \mathfrak{m}, \hat{\lambda}) \in (0,\infty)$ the constant defined in (3.31).

The proof of this proposition is deferred to Section 3.1.2. In order for this estimate to help us prove Theorem 3.0.1, we would like the right-hand side of

(3.16) to vanish as $\varepsilon \to 0$. This forces us to choose a cut-off level $N = N_\varepsilon$ that explodes as $\varepsilon \to 0$, and a suitably small coupling constant $\hat{\lambda}$. In particular, we fix the cut-off level N_ε given by

$$N_\varepsilon = \lfloor \log \tfrac{1}{\varepsilon} \rfloor \,, \tag{3.17}$$

and choose $\hat{\lambda} \in (0, \hat{\lambda}_{\text{fin}})$ for some $\hat{\lambda}_{\text{fin}}$ small enough, which we fix later on. Given the error estimate above, the next task is to identify the convergence of the truncated sequence $u_\varepsilon^{N_\varepsilon} = \sum_{\tau \in \mathcal{T}_3^{N_\varepsilon,\circ}} X_\varepsilon^\tau$. This convergence is very delicate, in particular because the number of terms in the sum now grows with ε. The next proposition contains the key estimate that allows us to overcome this issue. For this, let us fix now and throughout this chapter $h(y) := -y^3$.

Proposition 3.1.2. *Let $T > 0$ and $\hat{\lambda} > 0$, then uniformly over any $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and $\tau \in \mathcal{T}_3^{N_\varepsilon,\circ}$, with $N_\varepsilon = \lfloor \log \frac{1}{\varepsilon} \rfloor$, and uniformly over all $(t,x) \in [0,T] \times \mathbf{R}^2$, we have*

$$\begin{aligned}
&\Big\| \sqrt{\log \tfrac{1}{\varepsilon}} \cdot X_\varepsilon^\tau(t,x) - \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \left(\frac{3\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \hat{\lambda} e^{\mathfrak{m}\, t} P_{t+\varepsilon^2} \eta(x) \Big\|_{L^2(\mathbb{P})} \\
&\qquad \leqslant \frac{|h^{(\tau)}(1)|}{\tau!\, s(\tau)} \left(C\, \hat{\lambda}^2 e^{2\overline{\mathfrak{m}}\, t} \right)^{|\tau|} \frac{e^{2|\mathfrak{m}|\, t} + |\log(t+\varepsilon^2)| + \sqrt{\log \frac{1}{\varepsilon}}}{2 \log \frac{1}{\varepsilon}} \frac{\hat{\lambda} e^{\mathfrak{m}\, t}}{\sqrt{4(t+\varepsilon^2)}} \,,
\end{aligned}$$

where τ is the trimmed tree $\mathscr{T}(\tau)$ defined in (2.82) and C is the finite constant defined in (3.82).

The above result both identifies the limit of $\sqrt{\log \frac{1}{\varepsilon}} \cdot X_\varepsilon^\tau$ and gives a quantitative result on its rate of convergence. The proof of Proposition 3.1.2 is at the heart of this chapter and can be found at the end of Section 3.3. It builds on all the results that are derived on the way. We highlight that the bound we obtain is uniform over all trees $\tau \in \mathcal{T}_3$ with $|\tau| = O(\log \frac{1}{\varepsilon})$. This is rather remarkable: As $|\tau|$ grows, every tree consists of a growing number of components living in distinct homogeneous chaoses and it requires precise estimates to bound all of them at once. In particular, the right-hand side is summable over τ and decays for $\varepsilon \to 0$, a fact that is necessary to make rigorously sense of the following approximations:

$$\sqrt{\log \tfrac{1}{\varepsilon}} \sum_{\tau \in \mathcal{T}_3^{N_\varepsilon,\circ}} X_\varepsilon^\tau(t,x) \sim \hat{\lambda} \sum_{\tau \in \mathcal{T}_3^{N_\varepsilon,\circ}} \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \Big(\frac{3\hat{\lambda}^2}{2\pi} \Big)^{|\tau|} e^{\mathfrak{m}\, t} P_{t+\varepsilon^2} \eta(x)$$

$$\sim \hat{\lambda} \sum_{\tau \in \mathcal{T}_3} \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \Big(\frac{3\hat{\lambda}^2}{2\pi} \Big)^{|\tau|} P_t^{(\mathfrak{m})} \eta(x), \qquad \text{as} \quad \varepsilon \to 0.$$

The series in the expression equals

$$\hat{\lambda} \sum_{\tau \in \mathcal{T}_3} \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \Big(\frac{3\hat{\lambda}^2}{2\pi} \Big)^{|\tau|} = \sigma_{\hat{\lambda}} = \hat{\lambda}\, y\Big(\frac{3\hat{\lambda}^2}{2\pi} \Big)\,, \tag{3.18}$$

by Lemma 2.2.3, with

$$y(\zeta) := \frac{1}{\sqrt{1+2\zeta}} \tag{3.19}$$

being the solution of the differential equation $\dot{y} = -y^3$, $y(0) = 1$. The above limiting behaviour is the content of the following proposition, which is the final step towards the proof of Theorem 3.0.1.

Proposition 3.1.3. *Let $\hat{\lambda} > 0$ and $T > 0$ satisfy*

$$\hat{\lambda} e^{\overline{\mathfrak{m}}\, T} < \frac{1}{\sqrt{2C}}\,, \tag{3.20}$$

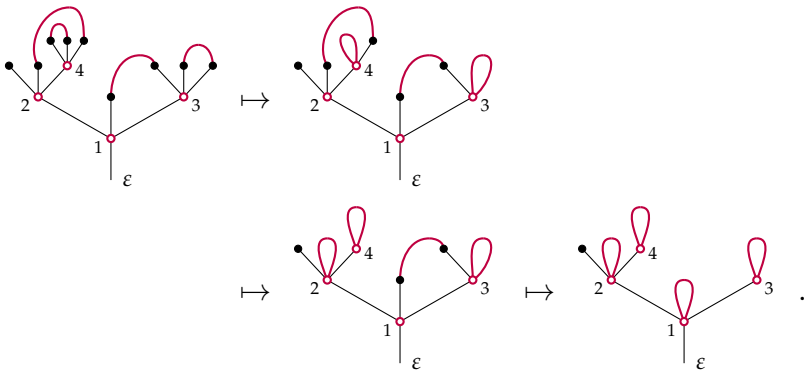
with $\overline{\mathfrak{m}} = \mathfrak{m} \vee 0$ and C be the positive constant defined in (3.82). Then for all $(t,x) \in (0,T] \times \mathbf{R}^2$

$$\lim_{\varepsilon \to 0} \Big\| \sqrt{\log \tfrac{1}{\varepsilon}}\, u_\varepsilon^{N_\varepsilon}(t,x) - \sigma_{\hat{\lambda}} P_t^{(\mathfrak{m})} \eta(x) \Big\|_{L^2(\mathbb{P})} = 0\,,$$

where $\sigma_{\hat{\lambda}}$ as in (3.18).

We will provide the proof of Proposition 3.1.3 at the end of the section. Before we pass to the proof of Theorem 3.0.1, let us explain the structure that underlies our main estimate, contained in Proposition 3.1.2. In Chapter 2, we saw that each Wiener integral X_ε^τ lies in an *inhomogeneous Wiener chaos* and can be decomposed into its homogeneous chaos components. These projections are realised in terms of all possible *pairwise contractions* of noises, cf. Definition 2.1.14 or the discussion below (2.73). An example of a possible contraction, appearing in the study of (3.2), is shown in the left-most element

of the following display:



(3.21)

In this display, **each** inner node can be associated to a **unique** contraction. A crucial observation will be that only homogeneous chaos configurations consisting of precisely such contractions (i.e. internal contractions can be extracted from all inner nodes, possibly in a nested manner) contribute in the limit $\varepsilon \to 0$. This will be the content of Section 3.2. Once the configurations with the dominant contribution have been identified, the next step is to determine their limit. Let us describe the relevant structure based on the example diagram (3.21). In the left-most tree the contraction among the two leaves neighbouring 4 (which creates the loop $4 \to 4$) will produce the weight, cf. (2.66),

$$\lambda_\varepsilon^2 \int_{\mathbf{R}^2} \left(e^{\mathrm{m} s_4} p_{s_4+\varepsilon^2}(z-y_4)\right)^2 \mathrm{d}z = \lambda_\varepsilon^2 e^{2\mathrm{m} s_4} p_{2(s_4+\varepsilon^2)}(0)\,,$$

by using the Chapman–Kolmogorov equations. Similarly, the loop $3 \to 3$ will have a weight $\lambda_\varepsilon^2 e^{2\mathrm{m} s_3} p_{2(s_3+\varepsilon^2)}(0)$. Next, moving to the contraction that creates the path $2 \to 4 \to 2$, we see it also produces the weight

$$\lambda_\varepsilon^2 \int_{(\mathbf{R}^2)^2} e^{\mathrm{m} s_2} p_{s_2+\varepsilon^2}(z-y_2) p^{(\mathrm{m})}_{s_2-s_4}(y_4-y_2) e^{\mathrm{m} s_4} p_{s_4+\varepsilon^2}(z-y_4)\,\mathrm{d}z\,\mathrm{d}y_4 = \lambda_\varepsilon^2 e^{2\mathrm{m} s_2} p_{2(s_2+\varepsilon^2)}(0),$$

which we represent by the loop $2 \to 2$. For the same reason, the contraction $1 \to 3 \to 1$ produces the weight $\lambda_\varepsilon^2 e^{2\mathrm{m} s_1} p_{2(s_1+\varepsilon^2)}(0)$, represented by the loop $1 \to 1$ in the right-most tree. Some crucial observations, that can be made in light of this example are the following.

- The condition that every inner node is associated to a contraction, leads to a nested contraction structure with only one leaf (one noise) that is not

contracted.

- Successively separating the weights from contracted trees and integrating out the spatial dependency, leads to a diagram indexed by the *trimmed tree* $\mathscr{T}(\tau) = \tau$ (2.82). To each of the trimmed trees' vertices v we assign a weight $\lambda_\varepsilon^2 e^{2\mathfrak{m}\, s_v} p_{2(s_v+\varepsilon^2)}(0) = \frac{\lambda_\varepsilon^2}{4\pi(s_v+\varepsilon^2)} e^{2\mathfrak{m}\, s_v}$.

- The temporal variables $\{s_v \colon v \in \mathcal{V}(\tau)\}$ will inherit the ordering imposed by the trimmed tree τ. In the above example: $s_1 > s_2, s_3$ and $s_2 > s_4$. The integration of the total weight $\prod_{v\in\mathcal{V}(\tau)} \lambda_\varepsilon^2 e^{2\mathfrak{m}\, s_v} p_{2(s_v+\varepsilon^2)}(0)$ under the constraints of τ leads to the factor $(\tau!)^{-1} \frac{3\hat{\lambda}^2}{2\pi}$ in the Butcher series (3.18). Here the numerical factor 3 is due to the fact that there are $\binom{3}{2} = 3$ ways to contract the noises of a trident.

Verifying the above observations rigorously for arbitrary trees will be the bulk of this chapter. To do so, we will formally introduce pairings and the notion of v-cycles in Section 3.2. In Section 3.3 we will then finally prove Proposition 3.1.2.

3.1.2 Proof of Theorem 3.0.1

We are now ready to prove our main result, given the estimates in Proposition 3.1.1 and Proposition 3.1.3, the proofs of which are postponed to further below in the section.

Proof of Theorem 3.0.1. We define $\hat{\lambda}_{\text{fin}} := \frac{1}{10\sqrt{C}}$, where C is the positive constant defined in (3.82). Let $\hat{\lambda} \in (0, \hat{\lambda}_{\text{fin}})$ and $T > 0$ such that (3.4) is satisfied, which equivalently reads

$$\hat{\lambda} e^{\overline{\mathfrak{m}}\, T} < \frac{1}{10\sqrt{C}}\,. \tag{3.22}$$

By the triangle inequality, for $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and $(t,x) \in (0,T] \times \mathbf{R}^2$

$$\begin{aligned}\big\|\mathcal{U}_\varepsilon(t,x) - \sigma_{\hat{\lambda}}\, P_t^{(\mathfrak{m})}\eta(x)\big\|_{L^2(\mathbb{P})} \leqslant \sqrt{\log\tfrac{1}{\varepsilon}}\, \Big\|u_\varepsilon(t,x) - u_\varepsilon^{N_\varepsilon}(t,x)\Big\|_{L^2(\mathbb{P})} \\ + \Big\|\sqrt{\log\tfrac{1}{\varepsilon}}\; u_\varepsilon^{N_\varepsilon}(t,x) - \sigma_{\hat{\lambda}}\, P_t^{(\mathfrak{m})}\eta(x)\Big\|_{L^2(\mathbb{P})}\,.\end{aligned}$$

The second term on the right-hand side vanishes as $\varepsilon \to 0$ by Proposition 3.1.3, since (3.22) implies (3.20). On the other hand, by Proposition 3.1.1,

the first term is upper bounded by

$$\sqrt{\log \tfrac{1}{\varepsilon}} \left\| u_\varepsilon(t,x) - u_\varepsilon^{N_\varepsilon}(t,x) \right\|_{L^2(\mathbb{P})} \leqslant \frac{C_0}{\log \frac{1}{\varepsilon}} \frac{\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, T}\right)^{N_\varepsilon}}{\varepsilon} .$$

The polynomial blow-up on the right-hand side must be compensated, and here we will crucially use that $N_\varepsilon \sim \log \frac{1}{\varepsilon}$, so that we have a compensating effect coming from the term $\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, T}\right)^{N_\varepsilon}$. More precisely, by choice of T in (3.22), we have

$$-\log\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, T}\right) \geqslant \log 10 > 1 ,$$

thus, for all $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$

$$\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, T}\right)^{N_\varepsilon} \leqslant \exp\left(\log\left(\sqrt{C}\hat{\lambda} e^{\overline{\mathfrak{m}}\, T}\right)\left(\log \tfrac{1}{\varepsilon} - 1\right)\right) \leqslant 10\, \varepsilon^{\log 10} .$$

Hence, we obtain that

$$\sqrt{\log \tfrac{1}{\varepsilon}} \left\| u_\varepsilon(t,x) - u_\varepsilon^{N_\varepsilon}(t,x) \right\|_{L^2(\mathbb{P})} \leqslant \frac{10\, C_0}{\log \frac{1}{\varepsilon}} \varepsilon^{(\log 10)-1} ,$$

which vanishes in the limit $\varepsilon \to 0$. This concludes the proof. □

In the remainder of this section we prove that the truncated Wild expansion $u_\varepsilon^{N_\varepsilon}$ indeed approximates the solution u_ε (Proposition 3.1.1) and that $u_\varepsilon^{N_\varepsilon}(t,x)$ is close to $\sigma_{\hat{\lambda}} P_t \eta(x)$ in $L^2(\mathbb{P})$ (Proposition 3.1.3). The proof of Proposition 3.1.2 will be given at the end of Section 3.3.

Proof of Proposition 3.1.1. Let $T > 0$, $(t,x) \in (0,T] \times \mathbf{R}^2$ and $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$. From (3.2) and (3.13), we obtain that the difference $w_\varepsilon^N = u_\varepsilon^N - u_\varepsilon$ solves the equation

$$\partial_t w_\varepsilon^N = \frac{1}{2}\Delta w_\varepsilon^N + \mathfrak{m} w_\varepsilon^N - (u_\varepsilon^N)^3 + u_\varepsilon^3 + R_\varepsilon^N , \quad w_\varepsilon^N(0,\cdot) = 0 , \tag{3.23}$$

with R_ε^N defined in (3.14). Defining $V_\varepsilon^N(t,x) := \frac{(u_\varepsilon^N)^3 - u_\varepsilon^3}{u_\varepsilon^N - u_\varepsilon}$, we can write (3.23) as

$$\partial_t w_\varepsilon^N = \frac{1}{2}\Delta w_\varepsilon^N + \mathfrak{m} w_\varepsilon^N - V_\varepsilon^N \cdot w_\varepsilon^N + R_\varepsilon^N , \quad w_\varepsilon^N(0,\cdot) = 0 .$$

The Feynman–Kac formula [KS91, Theorem 5.7.6] then allows to represent

w_ε^N as

$$w_\varepsilon^N(t,x) = \mathbf{E}_x\left[\int_0^t R_\varepsilon^N(t-s,W_s)\exp\left(\mathfrak{m}\, s - \int_0^s V_\varepsilon^N(s-r,W_r)\,\mathrm{d}r\right)\mathrm{d}s\right],$$

where W is a two dimensional Brownian path and $\mathbf{E}_x$ is the expectation with respect to it when the path starts from $x \in \mathbf{R}^2$. Using the fact that $V_\varepsilon^N \geqslant 0$, which is due to the monotonicity of the mapping $u \mapsto u^3$, we obtain that

$$\begin{aligned}
&\left|\, u_\varepsilon^N(t,x) - u_\varepsilon(t,x)\,\right| \\
&\qquad \leqslant \mathbf{E}_x\left[\int_0^t \left|R_\varepsilon^N(t-s,W_s)\right|\exp\left(\mathfrak{m}\, s - \int_0^s V_\varepsilon^N(s-r,W_r)\,\mathrm{d}r\right)\mathrm{d}s\right] \\
&\qquad \leqslant \mathbf{E}_x\left[\int_0^t e^{\mathfrak{m}\, s}\left|R_\varepsilon^N(t-s,W_s)\right|\mathrm{d}s\right].
\end{aligned}$$

Writing the latter in terms of the heat kernel we conclude that

$$\left|\, u_\varepsilon^N(t,x) - u_\varepsilon(t,x)\,\right| \leqslant \int_0^t \int_{\mathbf{R}^d} p_{t-s}^{(\mathfrak{m})}(y-x)\,|R_\varepsilon^N(s,y)|\,\mathrm{d}y\,\mathrm{d}s\,.$$

Taking the $L^2(\mathbb{P})$-norm, we arrive at the bound that we will be working with:

$$\|u_\varepsilon^N(t,x) - u_\varepsilon(t,x)\|_{L^2(\mathbb{P})} \leqslant e^{\overline{\mathfrak{m}}\, t}\int_0^t \|R_\varepsilon^N(s,0)\|_{L^2(\mathbb{P})}\,\mathrm{d}s\,, \tag{3.24}$$

where we have used that R_ε^N is spatially homogeneous. To continue, we use the definition of R_ε^N from (3.14), the triangle inequality and Hölder's inequality, to obtain

$$\begin{aligned}
\|R_\varepsilon^N(s,0)\|_{L^2(\mathbb{P})} &\leqslant \sum_{\substack{\tau_1,\dots,\tau_3 \in \mathcal{T}_3^{N,\circ} \\ [\tau_1\,\tau_2\,\tau_3] \notin \mathcal{T}_3^{N,\circ}}} \|(X_\varepsilon^{\tau_1} X_\varepsilon^{\tau_2} X_\varepsilon^{\tau_3})(s,0)\|_{L^2(\mathbb{P})} \qquad (3.25)\\
&\leqslant \sum_{\substack{\tau_1,\dots,\tau_3 \in \mathcal{T}_3^{N,\circ} \\ [\tau_1\,\tau_2\,\tau_3] \notin \mathcal{T}_3^{N,\circ}}} \|X_\varepsilon^{\tau_1}(s,0)\|_{L^6(\mathbb{P})}\|X_\varepsilon^{\tau_2}(s,0)\|_{L^6(\mathbb{P})}\|X_\varepsilon^{\tau_3}(s,0)\|_{L^6(\mathbb{P})}\,.
\end{aligned}$$

At this point we use a hypercontractivity estimate, which we recall from Proposition 2.1.7:

$$\|X_\varepsilon^{\tau_i}(s,y)\|_{L^6(\mathbb{P})} \leqslant 5^{\frac{\ell(\tau_i)}{2}}\,\|X_\varepsilon^{\tau_i}(s,y)\|_{L^2(\mathbb{P})}\,.$$

To bound the $L^2(\mathbb{P})$-norm, we will make use of Proposition 3.1.2 to obtain

$$\|X_\varepsilon^\tau(s,y)\|_{L^2(\mathbb{P})} \leqslant \frac{\tilde{c}(T,\mathfrak{m})}{\sqrt{\log\frac{1}{\varepsilon}}} \frac{|h^{(\tau)}(1)|}{\tau!\, s(\tau)} \left(C\,\hat{\lambda}^2 e^{2\overline{\mathfrak{m}}\, s}\right)^{|\tau|} \frac{\hat{\lambda} e^{\mathfrak{m}\, s}}{2\sqrt{s+\varepsilon^2}}\,, \tag{3.26}$$

with $\tilde{c}(T,\mathfrak{m}) := e^{2|\mathfrak{m}|\,T} + 4$. The verification of this bound is deferred to the bottom of this proof. Now, assuming (3.26) is true, from the identity $\ell(\tau_i) = 2|\tau_i| + 1$ (which holds since $\tau \in \mathcal{T}_3$) we obtain

$$\|X_\varepsilon^{\tau_i}(s,y)\|_{L^6(\mathbb{P})} \leqslant 5^{|\tau_i|+\frac{1}{2}} \frac{\tilde{c}(T,\mathfrak{m})}{\sqrt{\log\frac{1}{\varepsilon}}} \frac{|h^{(\tau_i)}(1)|}{\tau_i!\, s(\tau_i)} \left(C\,\hat{\lambda}^2 e^{2\overline{\mathfrak{m}}\, s}\right)^{|\tau_i|} \frac{\hat{\lambda} e^{\mathfrak{m}\, s}}{2\sqrt{s+\varepsilon^2}}\,. \tag{3.27}$$

Combining (3.24) with (3.25) and (3.27), we conclude that

$$\begin{aligned}
&\sqrt{\log\tfrac{1}{\varepsilon}} \cdot \|u_\varepsilon^N(t,x) - u_\varepsilon(t,x)\|_{L^2(\mathbb{P})} \\
&\qquad \leqslant \sqrt{\log\tfrac{1}{\varepsilon}} \cdot e^{\overline{\mathfrak{m}} t} \int_0^\infty \|R^\varepsilon(s,0)\|_{L^2(\mathbb{P})}\, \mathrm{d}s \\
&\qquad \leqslant \frac{\left(\sqrt{5}\hat{\lambda}\,\tilde{c}(T,\mathfrak{m})\right)^3 e^{4\overline{\mathfrak{m}} t}}{8\,\log\frac{1}{\varepsilon}} \left\{ \int_0^\infty \frac{1}{(s+\varepsilon^2)^{\frac{3}{2}}}\, \mathrm{d}s \right\} \cdot \\
&\qquad\qquad \times \sum_{\substack{\tau_1,\dots,\tau_3 \in \mathcal{T}_3^{N,\circ} \\ [\tau_1\,\tau_2\,\tau_3] \notin \mathcal{T}_3^{N,\circ}}} \prod_{i=1}^3 \left\{ \frac{|h^{(\tau_i)}(1)|}{\tau_i!\, s(\tau_i)} \left(5C\hat{\lambda}^2 e^{2\overline{\mathfrak{m}} t}\right)^{|\tau_i|} \right\}.
\end{aligned} \tag{3.28}$$

At this point, we notice that the time integral appearing in the last estimate blows up polynomially in ε, since

$$\int_0^\infty \frac{1}{(s+\varepsilon^2)^{\frac{3}{2}}}\, \mathrm{d}s = \left[-2(s+\varepsilon^2)^{-\frac{1}{2}}\right]_{s=0}^{s=\infty} = \frac{2}{\varepsilon}\,. \tag{3.29}$$

On the other hand, for any $\{\tau_i\}_{i=1}^3$ such that $[\tau_1\ \tau_2\ \tau_3] \notin \mathcal{T}_3^{N,\circ}$ we have that

$$N < i([\tau_1\ \tau_2\ \tau_3]) = i(\tau_1) + i(\tau_2) + i(\tau_3) + 1 = |\tau_1| + |\tau_2| + |\tau_3| + 1\,.$$

Moreover, by assumption $\hat{\lambda}$ is sufficiently small to satisfy $5\sqrt{C}\,\hat{\lambda} e^{\overline{\mathfrak{m}} t} < \frac{1}{2}$.

Therefore, we can estimate the sum in the last line of (3.28) as follows:

$$\begin{aligned}
&\sum_{\substack{\tau_1,\dots,\tau_3\in\mathcal{T}_3^{N,\circ}\\ [\tau_1\,\tau_2\,\tau_3]\notin\mathcal{T}_3^{N,\circ}}}\prod_{i=1}^{3}\left\{\frac{|h^{(\tau_i)}(1)|}{\tau_i!\,s(\tau_i)}\left(5C\hat{\lambda}^2e^{2\overline{\mathfrak{m}}\,t}\right)^{|\tau_i|}\right\}\\
&\qquad\leqslant\left(\sqrt{C}\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^N\sum_{\substack{\tau_1,\dots,\tau_3\in\mathcal{T}_3^{N,\circ}\\ [\tau_1\,\tau_2\,\tau_3]\notin\mathcal{T}_3^{N,\circ}}}\prod_{i=1}^{3}\left\{\frac{|h^{(\tau_i)}(1)|}{\tau_i!\,s(\tau_i)}\left(5\sqrt{C}\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^{|\tau_i|}\right\} \qquad (3.30)\\
&\qquad\leqslant\left(\sqrt{C}\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^N\frac{1}{(1-10\sqrt{C}\hat{\lambda}e^{\overline{\mathfrak{m}}\,t})^{\frac{3}{2}}}\,.
\end{aligned}$$

In the last step of (3.30), we completed the remaining sum to a Butcher series:

$$\begin{aligned}
&\sum_{\substack{\tau_1,\dots,\tau_3\in\mathcal{T}_3^{N,\circ}\\ [\tau_1\,\tau_2\,\tau_3]\notin\mathcal{T}_3^{N,\circ}}}\prod_{i=1}^{3}\left\{\frac{|h^{(\tau_i)}(1)|}{\tau_i!\,s(\tau_i)}\left(5\sqrt{C}\,\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^{|\tau_i|}\right\}\\
&\qquad\leqslant\left(\sum_{\tau\in\mathcal{T}_3^{N,\circ}}\frac{|h^{(\tau_i)}(1)|}{\tau!\,s(\tau)}\left(5\sqrt{C}\,\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^{|\tau|}\right)^3\\
&\qquad\leqslant\left(\overline{y}\left(5\sqrt{C}\,\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)\right)^3=\frac{1}{(1-2\cdot5\sqrt{C}\,\hat{\lambda}e^{\overline{\mathfrak{m}}\,t})^{\frac{3}{2}}}\,,
\end{aligned}$$

where $\overline{y}$ is the solution to the ODE $\dot{\overline{y}}=\overline{y}^3$ with **positive** initial condition $\overline{y}(0)=1$. The solution of this ODE is $\overline{y}(\zeta)=(1-2\zeta)^{-1/2}$ and so the associated Butcher series converges for $\zeta<1/2$, i.e. if (3.15) holds in the present case. See also Lemma 2.2.3 and Remark 2.2.5.

Hence, by putting together (3.28), (3.29) and (3.30), we obtain

$$\sqrt{\log\tfrac{1}{\varepsilon}}\cdot\|u_\varepsilon^N(t,\cdot)-u_\varepsilon(t,\cdot)\|_{L^2(\mathbb{P})}\leqslant\frac{C_0(T,\mathfrak{m},\hat{\lambda})}{\log\frac{1}{\varepsilon}}\frac{\left(\sqrt{C}\hat{\lambda}e^{\overline{\mathfrak{m}}\,t}\right)^N}{\varepsilon}$$

with

$$C_0(T,\mathfrak{m},\hat{\lambda}):=\frac{e^{4\overline{\mathfrak{m}}\,T}}{4}\left(\frac{\sqrt{5}\hat{\lambda}\,\tilde{c}(T,\mathfrak{m})}{\sqrt{1-10\sqrt{C}\,\hat{\lambda}e^{\overline{\mathfrak{m}}\,T}}}\right)^3. \qquad (3.31)$$

This concludes the proof of the proposition, up to the proof of (3.26). The lat-

ter bound follows simply from the triangle inequality and Proposition 3.1.2:

$$\begin{aligned}
&\|X_\varepsilon^\tau(t,x)\|_{L^2(\mathbb{P})} \\
&\leqslant \frac{|h^{(\tau)}(1)|}{\tau!\, s(\tau)} \left(\frac{3\hat\lambda^2}{2\pi}\right)^{|\tau|} \left\| \Big|_{\circ}^{\bullet} \mathcal{I}_\varepsilon(t,x) \right\|_{L^2(\mathbb{P})} \\
&\quad + \frac{|h^{(\tau)}(1)|}{\tau!\, s(\tau)} \left(C\,\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}\right)^{|\tau|} \frac{2e^{2|\mathfrak{m}|\,T}+4}{\sqrt{\log\frac{1}{\varepsilon}}} \frac{\hat\lambda e^{\mathfrak{m}\, t}}{\sqrt{4(t+\varepsilon^2)}}\,,
\end{aligned} \tag{3.32}$$

where we made use of the fact that $|\tau| \leqslant \log\frac{1}{\varepsilon}$ and the crude estimate

$$\frac{e^{2|\mathfrak{m}|\,t} + |\log(t+\varepsilon^2)| + \sqrt{\log\frac{1}{\varepsilon}}}{2\log\frac{1}{\varepsilon}} \leqslant \frac{e^{2|\mathfrak{m}|\,t}}{2\log 2} + 2 + \frac{1}{2\sqrt{\log 2}} \leqslant e^{2|\mathfrak{m}|\,T} + 3\,,$$

which is a consequence of $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and the uniform estimate

$$\begin{aligned}
\sup_{0\leqslant t\leqslant T} \frac{|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}} &= \sup_{0\leqslant t\leqslant 1-\varepsilon^2} \frac{|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}} \vee \sup_{1-\varepsilon^2<t<T} \frac{\log(t+\varepsilon^2)}{2\log\frac{1}{\varepsilon}} \\
&\leqslant 1 + \frac{\log(\frac{1}{\varepsilon}+1)}{2\log\frac{1}{\varepsilon}} \leqslant 2\,.
\end{aligned} \tag{3.33}$$

Now the statement follows, since

$$\begin{aligned}
\left(\frac{3\hat\lambda^2}{2\pi}\right)^{|\tau|} \left\| \Big|_{\circ}^{\bullet} \mathcal{I}_\varepsilon(t,x) \right\|_{L^2(\mathbb{P})} &= \left(\frac{3\hat\lambda^2}{2\pi}\right)^{|\tau|} \left(\hat\lambda_\varepsilon^2 e^{2\mathfrak{m}\, t} \int_{\mathbf{R}^2} p_{t+\varepsilon^2}(y)^2\, \mathrm{d}y\right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{\log\frac{1}{\varepsilon}}} \left(\frac{3\hat\lambda^2}{2\pi}\right)^{|\tau|} \frac{\hat\lambda e^{\mathfrak{m}\, t}}{\sqrt{4\pi(t+\varepsilon^2)}} \\
&\leqslant \frac{\left(C\,\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}\right)^{|\tau|}}{\sqrt{\log\frac{1}{\varepsilon}}} \frac{\hat\lambda e^{\mathfrak{m}\, t}}{2\sqrt{t+\varepsilon^2}}\,,
\end{aligned}$$

with C being the constant from (3.82). Thus, together with (3.32), we obtain

$$\|X_\varepsilon^\tau(t,x)\|_{L^2(\mathbb{P})} \leqslant \frac{\tilde c(T,\mathfrak{m})}{\sqrt{\log\frac{1}{\varepsilon}}} \frac{|h^{(\tau)}(1)|}{\tau!\, s(\tau)} \left(C\,\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}\right)^{|\tau|} \frac{\hat\lambda e^{\mathfrak{m}\, t}}{2\sqrt{t+\varepsilon^2}}\,.$$

with $\tilde c(T,\mathfrak{m}) = e^{2|\mathfrak{m}|\,T} + 4$. This completes the proof. □

Proof of Proposition 3.1.3. For $h(y) = -y^3$, we introduce the truncated Butcher

series

$$B_h^\varepsilon(\zeta,1) = \sum_{\tau\in\mathcal{T}_{\leqslant 3}^{N_\varepsilon}} \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \zeta^{|\tau|} = \sum_{\tau\in\mathcal{T}_3^{N_\varepsilon,\circ}} \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \zeta^{|\tau|}\,, \tag{3.34}$$

where the second equality is a consequence of Lemma 2.3.11. Then, Proposition 3.1.3 will follow if we can show that the following two limits hold true with $\zeta_{\hat\lambda} = \frac{3\hat\lambda^2}{2\pi}$:

$$\lim_{\varepsilon\to 0} \left\| \sqrt{\log\tfrac{1}{\varepsilon}}\, u_\varepsilon^{N_\varepsilon}(t,x) - \hat\lambda B_h^\varepsilon(\zeta_{\hat\lambda},1)\, e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) \right\|_{L^2(\mathbb{P})} = 0\,, \tag{3.35}$$

$$\lim_{\varepsilon\to 0} \left\| \hat\lambda B_h^\varepsilon(\zeta_{\hat\lambda},1)\, e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) - \sigma_{\hat\lambda} P_t^{(\mathfrak{m})}\eta(x) \right\|_{L^2(\mathbb{P})} = 0\,. \tag{3.36}$$

The limit (3.35) follows from Proposition 3.1.2, provided (3.20) holds: We can bound for $\varepsilon \in (0, \frac{1}{T}\wedge\frac{1}{2})$

$$\begin{aligned}
&\left\| \sqrt{\log\tfrac{1}{\varepsilon}}\, u_\varepsilon^{N_\varepsilon}(t,x) - \hat\lambda B_h^\varepsilon(\zeta_{\hat\lambda},1)\, e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) \right\|_{L^2(\mathbb{P})} \\
&\leqslant \sum_{\tau\in\mathcal{T}_3^{N_\varepsilon,\circ}} \left\| \sqrt{\log\tfrac{1}{\varepsilon}}\, X_\varepsilon^\tau(t,x) - \hat\lambda \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \zeta_{\hat\lambda}^{|\tau|} e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) \right\|_{L^2(\mathbb{P})} \\
&\leqslant \sum_{\tau\in\mathcal{T}_{\leqslant 3}^{N_\varepsilon}} \frac{|h^{(\tau)}(1)|}{s(\tau)\,\tau!} \left(C\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}\right)^{|\tau|} \frac{e^{2|\mathfrak{m}|\, t} + |\log(t+\varepsilon^2)| + \sqrt{\log\frac{1}{\varepsilon}}}{2\log\frac{1}{\varepsilon}} \frac{\hat\lambda e^{\mathfrak{m}\, t}}{\sqrt{4(t+\varepsilon^2)}}\,,
\end{aligned} \tag{3.37}$$

where we used (3.34). Now, (3.35) will follow, if we can show that the series on the right-hand side is summable, that is if

$$\sum_{\tau\in\mathcal{T}_{\leqslant 3}} \frac{|h^{(\tau)}(1)|}{s(\tau)\,\tau!} \left(C\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}\right)^{|\tau|} < \infty\,. \tag{3.38}$$

This is the Butcher series associated to the ODE $\dot{\overline{y}} = \overline{y}^3$ with initial condition $\overline{y}(0) = 1$, which converges as long as (3.20) holds. See also the discussion in the proof of Proposition 3.1.1. Hence, (3.37), and thus (3.35), vanishes for arbitrary fixed $t \in (0,T]$, because the second-to-last ratio on the right-hand side in (3.37) vanishes in the limit $\varepsilon \to 0$.

To complete the proof of the proposition we must now check (3.36). Here

we observe that

$$\begin{aligned}
&\left\| \hat{\lambda} B_h^\varepsilon(\zeta_{\hat{\lambda}}, 1)\, e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) - \sigma_{\hat{\lambda}}\, P_t^{(\mathfrak{m})}\eta(x) \right\|_{L^2(\mathbb{P})} \\
&\leqslant |\hat{\lambda} B_h^\varepsilon(\zeta_{\hat{\lambda}}, 1) - \sigma_{\hat{\lambda}}| \cdot \left\| e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) \right\|_{L^2(\mathbb{P})} \\
&\qquad + \sigma_{\hat{\lambda}} \left\| e^{\mathfrak{m}\, t} P_{t+\varepsilon^2}\eta(x) - P_t^{(\mathfrak{m})}\eta(x) \right\|_{L^2(\mathbb{P})} .
\end{aligned}$$

The second term is converging to 0 by the continuity properties of the heat semigroup. Instead, for the first term we observe that $B_h^\varepsilon(\zeta, 1)$ is an approximation to the Butcher series associated to the solution $y(\zeta)$ of the ODE $\dot{y} = -y^3, y(0) = 1$, which is given by $y(\zeta) = (1 + 2\zeta)^{-1/2}$. This solution is analytic for $|\zeta| < 1/2$ and the associated Butcher series converges, see the discussion in Section 2.2.3. Thus, recalling Lemma 2.2.3 and the explicit form of $\sigma_{\hat{\lambda}}$ (3.18), we have that $\lim_{\varepsilon\to 0} |\hat{\lambda} B_h^\varepsilon(\zeta_{\hat{\lambda}}, 1) - \sigma_{\hat{\lambda}}| = 0$, as long as $\frac{3\hat{\lambda}_{\text{fin}}^2}{2\pi} < \frac{1}{2}$, which is implied by (3.20). This concludes the proof. □

3.2 Paired trees and their structure

In Section 3.1.1, in particular around (3.21) and the discussion thereafter, we outlined the structure underlying the main estimate contained in Proposition 3.1.2. Namely, only certain terms in the first chaos contribute to the Gaussian limit in Theorem 3.0.1. In this section, we introduce the notion of pairings and present estimates on the integration kernels associated to the Wild expansion terms X_ε^τ. This will allow us to analyse X_ε^τ rigorously and identify its limiting contribution in Section 3.3.

As a first step towards the analysis of components in the Wild expansion (3.12), we can identify each term X_ε^τ with a term in a Butcher series, which is a direct consequence of Corollary 2.3.12.

Lemma 3.2.1. *For every term in the Wild expansion* (3.12) *indexed by* $\tau \in \mathcal{T}_3$*, the following identity holds:*

$$X_\varepsilon^\tau(t, x) = \frac{h^{(\tau)}(1)}{s(\tau)} [\tau]_\varepsilon(t, x) , \quad \forall \varepsilon \in (0, \tfrac{1}{2}) , \ (t, x) \in (0, \infty) \times \mathbf{R}^2 ,$$

with $h(y) = -y^3$*. Here* $[\tau]_\varepsilon$ *denotes the stochastic integral introduced in* (2.64) *(with respect to* $\Xi = \delta_0\eta$*) and* $\tau = \mathscr{T}(\tau)$ *is the trimmed version of* τ (2.82).

Recall from (2.73) that an iterated stochastic Stratonovich integral $[\tau]_\varepsilon$ can be represented in terms of Itô integrals $[\tau]_{\kappa,\varepsilon}$, by summing over all contrac-

tions of the associated tree's leaves (2.68):

$$[\tau]_\varepsilon = \sum_{\kappa \in \mathcal{K}(\tau)} [\tau]_{\kappa,\varepsilon} \,. \tag{3.39}$$

Here, each contracted tree $[\tau]_\kappa$, $\kappa \in \mathcal{K}(\tau)$, is associated to a stochastic integral lying in a *homogeneous* Wiener chaos. With $[\tau]_{\kappa,\varepsilon}$ as in (2.72), for $\Xi = \delta_0 \eta$, such that

$$\begin{aligned} &[\tau]_{\kappa,\varepsilon}(t,x) \\ &= \int_{D_t^{\mathcal{V}(\tau)}} \widehat{K}^{t,x}_{[\tau]_\kappa,\varepsilon}(s_{\mathcal{V}(\tau)}, y_{\mathcal{V}(\tau)}) \, \mathrm{d}y_{\mathcal{V}(\tau)\setminus\mathcal{L}(\tau_\kappa)} \, \mathrm{d}s_{\mathcal{V}(\tau)} \, \eta^{\otimes \ell(\tau_\kappa)}(\, \mathrm{d}y_{\mathcal{L}(\tau_\kappa)}) \,, \end{aligned}$$

with $D_t = [0,t] \times \mathbf{R}^2$ and

$$\begin{aligned} &\widehat{K}^{t,x}_{[\tau]_\kappa,\varepsilon}(s_{\mathcal{V}(\tau)}, y_{\mathcal{V}(\tau)}) \\ &:= \widehat{K}^{t,x}_{[\tau],\varepsilon}(s_{\mathcal{V}(\tau)}, y_{\mathcal{V}(\tau)}) \prod_{v \in \mathcal{L}(\tau)} \delta_0(s_v) \prod_{\{u,u'\} \in \kappa} \delta_0(y_u - y_{u'}) \,, \end{aligned} \tag{3.40}$$

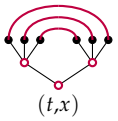
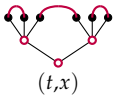
where $\widehat{K}^{t,x}_{[\tau],\varepsilon}$ is the kernel defined in (2.66) with respect to $[\tau]$. For convenience, we also included all arising Dirac-δ's in the kernel $\widehat{K}^{t,x}_{[\tau]_\kappa,\varepsilon}$. We will write both $[\tau]_{\kappa,\varepsilon}$ and $[\tau_\kappa]_\varepsilon$ for the Itô stochastic integral, when considering the planted tree $[\tau]$.

3.2.1 Pairings between trees and $L^2(\mathbb{P})$ estimates

We want to extend the notion of contractions from a single tree to a pair of trees. This will be necessary in order to encode second moments of stochastic iterated integrals. Recalling Lemma 2.1.15 and (2.73), we see that second moments of $[\tau]_\varepsilon$ can be expressed as the sum over all possible *pairwise contractions* over the (uncontracted) leaves of two copies of the tree, connected to the same root with time-space variables (t,x). In other words, we look at the stochastic integral $[\tau,\tau]_\varepsilon(t,x)$ corresponding to the tree $[\tau,\tau]$ with root variable (t,x).

Example 3.2.2. Let us compute the second moment of the stochastic integral $[\tau]_\varepsilon(t,x) =$  $_\varepsilon(t,x)$, represented by

$$\mathbb{E}\big[[\tau]_\varepsilon(t,x)^2\big] = \mathbb{E}\big[[\tau,\tau]_\varepsilon(t,x)\big] = \mathbb{E}\big[\big[$$  $$\big]_\varepsilon(t,x)\big]$$

$$= 6$$  $$+ 9$$  , (3.41)

similar to (2.76). In other words, the computation of the second moment of an integral $[\tau]_\varepsilon$ gives rise to *completely* contracted trees $[\tau, \tau]_\kappa$ in accordance with the definition of (2.68). We also note that this computation is consistent if we decomposed $[\tau]_\varepsilon(t,x)$ into its homogeneous Wiener chaos components first:

$$\mathbb{E}[[\tau](t,x)^2] = \mathbb{E}\Big|$$  $$\Big|^2 + \mathbb{E}\Big| 3$$  $$\Big|^2 ,$$

where we also used the orthogonality between different homogeneous components.

Let us introduce a notation that will allow us to encode contractions between glued together trees, in a way that distinguishes them from the contractions of Definition 2.3.7. This will be useful to encode covariances between $[\tau]_\varepsilon(t,x)$ and $[\tau']_\varepsilon(t,x)$.

Definition 3.2.3. *For two rooted trees $\tau, \tau' \in \mathcal{T}$ define the set of **pairings** among the union of leaves as*

$$\mathcal{Y}(\tau, \tau') := \big\{ \gamma \in \mathcal{K}([\tau, \tau']) \,:\, \gamma \text{ is a complete contraction } \big\} .$$

We also define the subsets of pairings which complete a given pair of contractions $(\kappa, \kappa') \in \mathcal{K}(\tau) \times \mathcal{K}(\tau')$ by

$$\mathcal{Y}(\tau_\kappa, \tau'_{\kappa'}) := \big\{ \gamma \in \mathcal{Y}(\tau, \tau') \,:\, \kappa \cup \kappa' \subset \gamma \text{ and all remaining pairs in } \gamma \text{ connect } \tau \text{ to } \tau' \big\} .$$

We will write $[\tau, \tau']_\gamma$ to denote the tree $[\tau, \tau']$ where all leaves are contracted according to $\gamma \in \mathcal{Y}(\tau, \tau')$.

The pictorial representations in (3.41) shows all possible elements (up to symmetries) of $\mathcal{Y}(\tau, \tau)$ for that example. Note that $\mathcal{Y}(\tau_\kappa, \tau'_{\kappa'})$ gives rise to a *completely contracted* tree and, therefore, $\mathcal{Y}(\tau_\kappa, \tau'_{\kappa'}) = \emptyset$, if the number of uncontracted leaves in τ_κ and $\tau'_{\kappa'}$ differ. This is in harmony with the fact that homogeneous chaoses are orthogonal with respect to one another. We can

now express covariances between contracted trees as follows:

$$\mathbb{E}\Big[\,[\tau]_{\kappa,\varepsilon}\,[\tau']_{\kappa',\varepsilon}\,\Big] = \sum_{\gamma\in\mathcal{Y}(\tau_\kappa,\tau'_{\kappa'})} [\,\tau,\tau']_{\gamma,\varepsilon}\,. \tag{3.42}$$

Moreover, it is clear that $\mathcal{Y}(\tau_\kappa,\tau'_{\kappa'})$ allows to partition $\mathcal{Y}(\tau,\tau')$ as

$$\mathcal{Y}(\tau,\tau') = \bigsqcup_{(\kappa,\kappa')\in\mathcal{K}(\tau)\times\mathcal{K}(\tau')} \mathcal{Y}(\tau_\kappa,\tau'_{\kappa'})\,, \tag{3.43}$$

where $\bigsqcup$ denotes a disjoint union. This is clear since, if we want to find all pairwise contractions of $[\tau,\tau']$, we can first identify the contractions that are internal to each τ,τ' and then identify the contractions that connect the leaves of one tree to those of the other. This partitioning then allows us to express covariances in terms of

$$\begin{aligned}\mathbb{E}\Big[\,[\tau]_{\varepsilon}\,[\tau']_{\varepsilon}\,\Big] &= \sum_{\kappa\in\mathcal{K}(\tau)}\sum_{\kappa'\in\mathcal{K}(\tau')} \mathbb{E}\Big[\,[\tau]_{\kappa,\varepsilon}\,[\tau']_{\kappa',\varepsilon}\,\Big]\\ &= \sum_{\kappa\in\mathcal{K}(\tau)}\ \sum_{\kappa'\in\mathcal{K}(\tau')}\ \sum_{\gamma\in\mathcal{Y}(\tau_\kappa,\tau'_{\kappa'})} [\,\tau,\tau']_{\gamma,\varepsilon}\\ &= \sum_{\gamma\in\mathcal{Y}(\tau,\tau')} [\,\tau,\tau']_{\gamma,\varepsilon}\,,\end{aligned}$$

where we used (3.42) in the second step, and (3.43) in the last. Finally, the partitioning (3.43) allows us to recover the internal contractions associated to a given pairing. This motivates the following definition.

Definition 3.2.4. *For any $\tau,\tau'\in\mathcal{T}_3$, let $\mathfrak{s}_{[\tau,\tau']}:\mathcal{Y}(\tau,\tau')\to\mathcal{K}(\tau)\times\mathcal{K}(\tau')$ be the map that for any $\gamma\in\mathcal{Y}(\tau,\tau')$ identifies the unique pair $\mathfrak{s}_{[\tau,\tau']}(\gamma) := (\kappa_1(\gamma),\kappa_2(\gamma))$ such that*

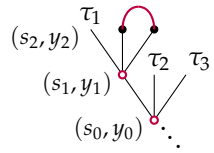
$$\gamma\in\mathcal{Y}(\tau_{\kappa_1(\gamma)},\tau'_{\kappa_2(\gamma)})\,.$$

In other words, the map $\mathfrak{s}$ identifies the subset of edges in γ that only connect within τ and τ', respectively.

3.2.2 1-cycles and their removal

Let us now define the notion of a 1-*cycle*. To this end, we consider a tree τ with a contraction κ and denote the corresponding contracted tree by τ_κ. Suppose

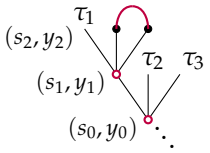
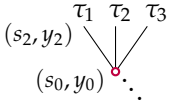
that τ_κ contains a component of the form

$\tau_\kappa =$  , (3.44)

where we observe a cycle consisting of an inner vertex (s_1, y_1) connected to two leaves that are themselves connected to one another by a purple edge (part of κ). We call such a cycle, that is a cycle in τ_κ which contains only *one* inner vertex of τ, a 1-*cycle*. Let us remark that in the above picture, (s_2, y_2) denotes the coordinates of the basis of the sub-tree τ_1 and (s_0, y_0) denotes the coordinates of the parent of the inner vertex with coordinates (s_1, y_1). A more formal definition is the following.

Definition 3.2.5. *Given a tree τ and a contraction $\kappa \in \mathcal{K}(\tau)$, we call a* 1*-cycle a connected component of τ_κ which consists of two leaves, which are connected by an element of κ, and the inner vertex, which is the parent of these leaves, as well as the three edges that connect these three vertices. We call the inner vertex of the cycle the basis of the* 1*-cycle.*

Given a contracted tree τ_κ with a 1-cycle $\mathcal{C}$, we write $(\tau \setminus \mathcal{C})_{\widetilde{\kappa}}$ for the contracted tree that is obtained by "removing" the cycle $\mathcal{C}$ from τ_κ. That is, the graph that remains after removing all edges and nodes that belong to $\mathcal{C}$ and replacing the remaining two edges which used to connect to the basis of the 1-cycle by a new, single edge which connects the only remaining descendant of the basis we have removed to its parent. The contraction $\widetilde{\kappa}$ is the one induced naturally on $\tau \setminus \mathcal{C}$ by κ after the removal of the element that connects the leaves of $\mathcal{C}$. The removal is simply described by the following picture:

$\tau_\kappa =$  $\longmapsto$ $(\tau \setminus \mathcal{C})_{\widetilde{\kappa}} =$  . (3.45)

Observe that if $\tau \in \mathcal{T}_3$, then also $\tau \setminus \mathcal{C} \in \mathcal{T}_3$. An important lemma is the following, which records the effect of the 1-cycle on the associated stochastic integrals.

Lemma 3.2.6. *Consider a rooted tree τ be of the form $\tau = [\tau_1 \cdots \tau_n]$, $\tau_i \in \mathcal{T}_3$. Let $\mathcal{C}$ be a 1-cycle in a contracted tree τ_κ with the coordinates of its root being*

$(t,x) \in (0,\infty)\times\mathbf{R}^2$. Denote by (s_1,y_1) the coordinates of its basis, by (s_0,y_0) the coordinates of the parent of (s_1,y_1) and by (s_2,y_2) the coordinates of the only descendant of (s_1,y_1) that does not belong to $\mathcal{C}$. Denote also by (r,z) the coordinates of the two identified leaves of the 1-cycle $\mathcal{C}$ (where we recall that the time coordinate r will coincide with 0)[2]*. Then*

$$\int_0^t \int_{(\mathbf{R}^2)^2} \widehat{K}^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{V}},y_{\mathcal{V}})\,\mathrm{d}y_1\,\mathrm{d}z\,\mathrm{d}r = \text{  }_\varepsilon(s_1)\mathbb{1}_{\{s_2\leqslant s_1\leqslant s_0\}}\widehat{K}^{t,x}_{\widetilde{\tau}_{\widetilde{\kappa}},\varepsilon}(s_{\mathcal{V}\setminus\mathcal{C}},y_{\mathcal{V}\setminus\mathcal{C}})\;, \tag{3.46}$$

where

$$\text{  }_\varepsilon(s_1) := \lambda_\varepsilon^2 e^{2\mathfrak{m}\,s_1}\,p_{2(s_1+\varepsilon^2)}(0)\;, \tag{3.47}$$

$\widetilde{\tau} := \tau\setminus\mathcal{C}$ and $\widetilde{\kappa}$ the contraction induced on $\widetilde{\tau}$ by the removal of $\mathcal{C}$ from τ_κ.

Proof. We start by performing the integration over the spatial coordinate z of the part of the kernel $\widehat{K}^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{V}},y_{\mathcal{V}})$ that depends on the variables (r,z). This corresponds to the following integral (recall form (3.40) that the kernel $\widehat{K}^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{V}},y_{\mathcal{V}})$ contains the factor $\delta_0(r)$):

$$\lambda_\varepsilon^2\int_{\mathbf{R}^2}\big(e^{\mathfrak{m}\,s_1}p_{s_1+\varepsilon^2}(z-y_1)\big)^2\,\mathrm{d}z = \lambda_\varepsilon^2 e^{2\mathfrak{m}\,s_1}\,p_{2(s_1+\varepsilon^2)}(0) = \text{  }_\varepsilon(s_1)\;. \tag{3.48}$$

Next, we integrate the remaining part of the kernel over y_1. This reduces to the Chapman–Kolmogorov identity (refer also to the picture in (3.44) for guidance):

$$\int_{\mathbf{R}^2} p^{(\mathfrak{m})}_{s_1-s_2}(y_2-y_1)p^{(\mathfrak{m})}_{s_0-s_1}(y_1-y_0)\;\mathrm{d}y_1 = p^{(\mathfrak{m})}_{s_0-s_2}(y_2-y_0)\;.$$

Combining the results of the two integrations above with the remaining components of the kernel $\widehat{K}^{t,x}_{\tau_\kappa,\varepsilon}(s_{\mathcal{V}},y_{\mathcal{V}})$, yields the expression on the right-hand side of (3.46). □

As it turns out, 1-cycles play an important role in our analysis. In particular, it will be a crucial ingredient to understand the small-ε behaviour of the

[2]Sometimes it may be more intuitive to think of purple contraction edges as heat kernels p_{ε^2}, while all black edges (even the ones connected to leaves) represent $p^{(\mathfrak{m})}$. In this case, each leaf (even contracted) corresponds an unique time-space point. However, in order to stay coherent with the kernel representation from Chapter 2, we will write them as actual identification (Dirac-δ's) of time-space points.

integral

$$\int_0^t \mathcal{V}_\varepsilon(s)\,\mathrm{d}s = \frac{\lambda_\varepsilon^2}{4\pi}\int_0^t \frac{e^{2\mathfrak{m}\,s}}{s+\varepsilon^2}\,\mathrm{d}s\,, \qquad t\in(0,\infty)\,,\ \varepsilon\in(0,\tfrac{1}{2})\,,$$

which we shall summarise in the following lemma.

Lemma 3.2.7. *Let $\mathfrak{m}\in\mathbf{R}$ and $t\in(0,\infty)$, then*

$$\left|\frac{1}{2\log\frac{1}{\varepsilon}}\int_0^t \frac{e^{2\mathfrak{m}\,s}}{s+\varepsilon^2}\,\mathrm{d}s - 1\right| \leqslant \frac{e^{2|\mathfrak{m}|\,t} + |\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}}\,,$$

for every $\varepsilon\in(0,\frac{1}{2})$, with $\overline{\mathfrak{m}} = \mathfrak{m}\vee 0$. In particular, for every $t\in(0,\infty)$

$$\lim_{\varepsilon\to 0}\int_0^t \mathcal{V}_\varepsilon(s)\,\mathrm{d}s = \frac{\hat{\lambda}^2}{2\pi}\,. \tag{3.49}$$

Proof. First expanding the exponential, we write

$$\int_0^t \frac{e^{2\mathfrak{m}\,s}-1}{s+\varepsilon^2}\,\mathrm{d}s = \int_0^t \frac{s}{s+\varepsilon^2}\sum_{k=1}^{\infty}\frac{(2\mathfrak{m})^k\,s^{k-1}}{k!}\,\mathrm{d}s\,.$$

Thus, we obtain

$$\left|\int_0^t \frac{e^{2\mathfrak{m}\,s}-1}{s+\varepsilon^2}\,\mathrm{d}s\right| \leqslant \sum_{k=1}^{\infty}\int_0^t \frac{|2\mathfrak{m}|^k\,s^{k-1}}{k!}\,\mathrm{d}s \leqslant e^{2|\mathfrak{m}|\,t}\,.$$

In addition, we have that

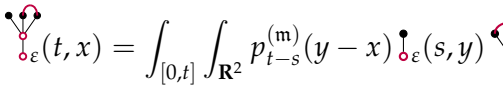
$$\left|\frac{1}{2\log\frac{1}{\varepsilon}}\int_0^t \frac{1}{s+\varepsilon^2}\,\mathrm{d}s - 1\right| = \frac{|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}}\,,$$

so that the statement follows from the triangle inequality. □

We will see that the contracted trees in the Wild expansion that contribute to the limiting fluctuations, are exactly those whose contraction creates only 1-cycles (which may also emerge in an iterative way, see the second example below). To get an idea of this phenomenon, let us look at the following examples.

Example 3.2.8. Consider the contracted tree $\mathcal{V}_\varepsilon$. Using Lemma 3.2.6 (or in

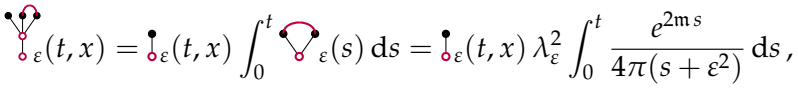
this case even a by-hand computation) we find that



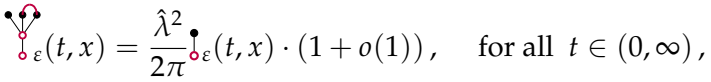
where inside the integral (s, y) are the time-space coordinates associated to the basis of the trident. We can next compute the spatial integral via Chapman–Kolmogorov as

$$\begin{aligned}\int_{\mathbf{R}^2} p^{(\mathfrak{m})}_{t-s}(y-x)\, \bullet_\varepsilon(s,y)\, \mathrm{d}y &= \lambda_\varepsilon \int_{(\mathbf{R}^2)^2} p^{(\mathfrak{m})}_{t-s}(y-x)\, e^{\mathfrak{m}\, s} p_{s+\varepsilon^2}(z-y)\, \eta(\,\mathrm{d}z)\, \mathrm{d}y \\ &= \lambda_\varepsilon e^{\mathfrak{m}\, t} \int_{\mathbf{R}^2} p_{t+\varepsilon^2}(z-x)\, \eta(\,\mathrm{d}z) = \bullet_\varepsilon(t,x)\,.\end{aligned}$$

Therefore, we obtain

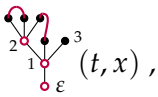


and hence, by Lemma 3.2.7, we find

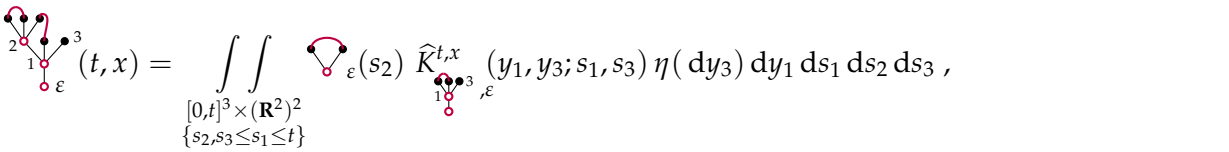


where the $o(1)$ is with respect to $\varepsilon \to 0$.

Example 3.2.9. This example demonstrates the iterative appearance of 1-cycles. Consider the contracted tree

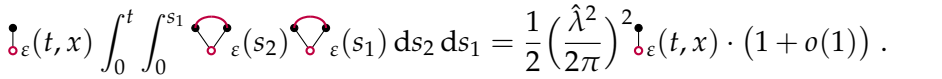


where we have tagged some vertices for reference in the following integrals. In particular, the coordinates of vertex i will be (s_i, y_i). Extracting first the 1-cycle with base vertex 2, we have that

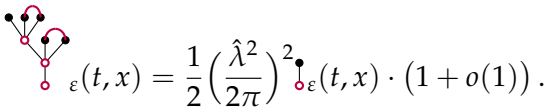


using Lemma 3.2.6. Now, applying once more Lemma 3.2.6, or just via the

previous example, this amounts to



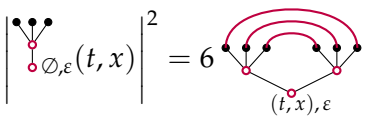
Thus, the contribution of this diagram is of the same order as in the previous example (albeit with a different constant) and will also contribute to the limiting Gaussian fluctuations. Following the same steps as above, we can determine similarly the contribution of the contracted tree



3.2.3 Existence of v-cycles in paired trees

Contrary to the above two examples, where only 1-cycles appeared, the next example will demonstrate a different cycle structure, which will lead to lower order contributions.

Example 3.2.10. Let us look at the order of magnitude of the term  . Its second moment has the diagrammatic representation in terms of the completely contracted tree



where the factor 6 counts the number of different pairings. Denoting by (s_1, y_1) and (s_2, y_2) the time-space coordinates of the bases of the left and right tridents, respectively, we can explicitly write the integral corresponding to the above diagram as

$$6\lambda_\varepsilon^6 \int_{D_t^2} p^{(\mathfrak{m})}_{t-s_1}(y_1 - x) \left(e^{\mathfrak{m}(s_1+s_2)} p_{s_1+s_2+2\varepsilon^2}(y_1 - y_2)\right)^3 p^{(\mathfrak{m})}_{t-s_2}(y_2 - x)\, \mathrm{d}y_{1,2}\, \mathrm{d}s_{1,2}$$

and then using the estimate

$$(e^{\mathfrak{m}(s_1+s_2)} p_{s_1+s_2+2\varepsilon^2}(y_1 - y_2))^2 \leqslant e^{4\overline{\mathfrak{m}}\, t} (2\pi(s_1 + s_2 + 2\varepsilon^2))^{-2}$$

together with Chapman–Kolmogorov, we can bound this by

$$
\begin{aligned}
&6\lambda_\varepsilon^6 e^{4\overline{\mathfrak{m}}\,t} \int_{D_t^2} \frac{p^{(\mathfrak{m})}_{t-s_1}(y_1 - x)\, e^{\mathfrak{m}\,(s_1+s_2)}\, p_{s_1+s_2+2\varepsilon^2}(y_1 - y_2)\, p^{(\mathfrak{m})}_{t-s_2}(y_2 - x)}{\big(2\pi(s_1+s_2+2\varepsilon^2)\big)^2}\, \mathrm{d}y_{1,2}\, \mathrm{d}s_{1,2} \\
&\qquad \leqslant \frac{6\lambda_\varepsilon^6}{(2\pi)^2} e^{6\overline{\mathfrak{m}}\,t}\, p_{2(t+\varepsilon^2)}(0) \int_{[0,t]^2} \frac{1}{(s_1+s_2+2\varepsilon^2)^2}\, \mathrm{d}s_1\, \mathrm{d}s_2 \\
&\qquad \leqslant \frac{6\lambda_\varepsilon^6}{(2\pi)^2} \log(1 + \tfrac{1}{2} t\varepsilon^{-2})\, e^{6\overline{\mathfrak{m}}\,t}\, p_{2(t+\varepsilon^2)}(0)\,.
\end{aligned}
$$

Since $\lambda_\varepsilon^6 = O((\log \frac{1}{\varepsilon})^{-3})$, we can conclude that

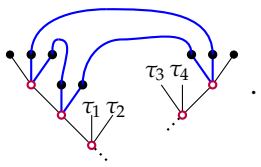
$$
\mathbb{E}\Big[\Big|\sqrt{\log \tfrac{1}{\varepsilon}}\, \mathcal{I}_{\emptyset,\varepsilon}(t,x)\Big|^2\Big] \leqslant \frac{C(t)}{\log \frac{1}{\varepsilon}}\,.
$$

In the last example there was no 1-cycle appearing. Instead, the contracted tree that emerged from the diagrammatic representation of the second moment, presented cycles containing *more than one* inner vertex, with every edge of the cycle incident to at least one leaf. We will call such cycles v**-cycles**. The emergence of v-cycle and the quantitative estimate of their contribution play a crucial role. The key observation is that contracted trees which do not consist of 1-cycles *only*, will have their second moment represented by a paired tree which necessarily contains a v-cycle. Such trees will turn out to have a lower order contribution.

Let us start with the formal definition of a v-cycle.

Definition 3.2.11. *For a given contracted tree τ_κ, a subgraph $\mathcal{C} = (\mathcal{V}_\mathcal{C}, \mathcal{E}_\mathcal{C}) \subseteq \tau_\kappa$ is a* v***-cycle*** *if it is a cycle in τ_κ (viewed as a graph) in which every edge is incident to at least one leaf of the tree τ. We define the* ***length*** *of a* v*-cycle to be the number of inner nodes of τ contained in $\mathcal{C}$ and we also denote by $\mathcal{I}_\mathcal{C}$ and $\mathcal{L}_\mathcal{C}$ the collection of the inner vertices and leaves of τ, respectively, that belong to $\mathcal{C}$. We will call a* v*-cycle of length $m \in \mathbf{N}$ a m-cycle for short.*

Note that a 1-cycle (Definition 3.2.5) is simply a special case of an m-cycle with $m = 1$. A pictorial example of a v-cycle is the one that appears in the following component of a contracted tree:



Consider a v-cycle of length m and denote its inner vertices by $v_1, ..., v_m$, where we will always keep the convention that in such an encoding we start from the left-most inner vertex of the cycle, in the graph picture of the tree. Let us also denote the time coordinates of $v_1, ..., v_m$ by $s_1, ..., s_m$, respectively, and introduce the kernel

$$\begin{aligned} \mathcal{K}_{\varepsilon}^{\otimes m}(s_1, \cdots, s_m) :=& \lambda_{\varepsilon}^{2m} \prod_{k=1}^{m} e^{\mathfrak{m}(s_k+s_{k+1})} p_{s_k+s_{k+1}+2\varepsilon^2}(0) \\ =& \prod_{k=1}^{m} \frac{\lambda_{\varepsilon}^2 e^{\mathfrak{m}(s_k+s_{k+1})}}{2\pi(s_k+s_{k+1}+2\varepsilon^2)}, \end{aligned} \tag{3.50}$$

with the convention that $s_{m+1} = s_1$. We notice that the above kernel is invariant under cyclic permutation of $s_1, s_2, ..., s_m$.

The following lemma establishes the existence of a v-cycle in a completely contracted tree of the form $[\tau, \tau']$.

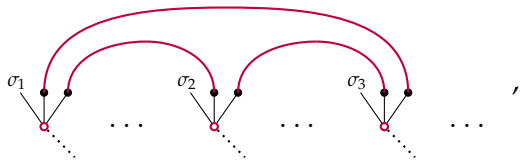
Lemma 3.2.12. *Let $\tau, \tau' \in \mathcal{T}_3$. Then, for every pairing $\gamma \in \mathcal{Y}(\tau, \tau')$, the paired tree $[\tau, \tau']_\gamma$ contains a* v*-cycle.*

First, we introduce some notation. For $\tau \in \mathcal{T}_{\leqslant 3}$, we partition the subset of inner nodes neighouring leaves in $\mathcal{L}(\tau)$ as follows: Let

- $\mathcal{V}_{\Yup}(\tau)$ be the subset of inner nodes $v \in \mathcal{I}(\tau)$ that is a **basis of a trident**, i.e. there exist exactly three $u_1, u_2, u_3 \in \mathcal{L}(\tau)$ such that $\mathfrak{p}(u_i) = v$.
- $\mathcal{V}_{\vee}(\tau)$ be the subset of inner nodes $v \in \mathcal{I}(\tau)$ that is a **basis of a cherry**, i.e. there exist exactly two $u_1, u_2 \in \mathcal{L}(\tau)$ such that $\mathfrak{p}(u_i) = v$.
- $\mathcal{V}_{\circ}(\tau)$ be the subset of inner nodes $v \in \mathcal{I}(\tau)$ that is a **basis of a lollipop**, i.e. there exist exactly one $u \in \mathcal{L}(\tau)$ such that $\mathfrak{p}(u) = v$. In the following, we will call elements in $\mathcal{V}_{\circ}(\tau)$ **dead-ends**.

Proof of Lemma 3.2.12. Let us start by noting that if $[\tau, \tau']$ does not contain any dead-ends, then the paired tree $[\tau, \tau']_\gamma$ contains a v-cycle. To see this, notice first that every leaf of $[\tau, \tau']$ belongs either to a cherry or to a trident. Now, consider an arbitrary leaf, call it v_0, and let v_1 be the unique leaf in $[\tau, \tau']$, which is connected to v_0 via γ. If v_1 is inside the same cherry or trident component as v_0, then we have already identified a v-cycle, which in this case is of length 1. If not, then let v_2 be a different leaf inside the same cherry or trident component as that of v_1, and denote by v_3 the leaf which is connected

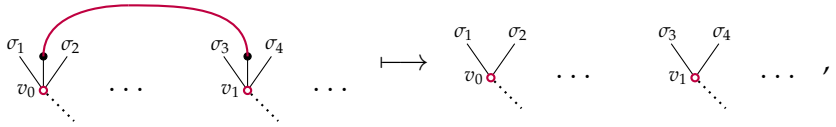
to v_2 via γ. Again, if v_2 and v_3 fall inside the same component (which in this case would necessarily be a trident), then a v-cycle comprising of leaves v_2, v_3 and the corresponding base point of the trident is identified. If not, then continue the procedure. Since there is only a finite number of leaves, we will either encounter somewhere in the process a v-cycle of length 1, or the path will return to a component previously visited, thus identifying a v-cycle. Diagrammatically, we have the following representation:



where sub-trees $\sigma_1, \sigma_1, \sigma_3$ may be identical to just a single leaf, i.e. $\bullet$, and even though we did not include them, there are γ–links emanating from the leaves of these trees.

We will next reduce the case that $[\tau, \tau']$ contains dead-ends to a situation of no dead-ends. Dead-ends present a problem: When tracing contractions in γ, we may hit a dead-end and thus are not able to continue to complete a v-cycle. What we will show is that by eliminating paths that start from a dead-end, the resulting sub-graph is one that consists of only cherries and tridents linked through γ. Thus, by the previous argument, a v-cycle exists within this sub-graph.

Let us start by picking an arbitrary dead-end of $[\tau, \tau']$. Call v_0 its associated leaf and suppose it connects to another leaf of $[\tau, \tau']$, which we call v_1. Now remove this connection as follows:



where we again understand that, despite not shown, there are γ–links emanating from trees $\sigma_1, \sigma_2, \sigma_3, \sigma_4$. Moreover, we agree that neither σ_1 nor σ_2 equals $\bullet$ (they might be $\emptyset$, though), so that this part of the tree corresponds to a dead-end, while σ_3, σ_4 might be comprising a $\bullet$.

In the resulting (contracted) tree on the right-hand side, we distinguish three cases:

(i) Neither of σ_3 and σ_4 are single leaves. In this case we have eliminated two dead-ends, while not affecting the number of cherries and tridents.

(ii) Only one of the σ_3 and σ_4 is a $\bullet$. In this case, we have eliminated a dead-end in the left part of the tree, while we have also created a new dead-end in the right part, corresponding to either σ_3 or σ_4, whichever happens to be the $\bullet$. In this case, we have also reduced by one the number of cherries (by eliminating the cherry that was present in the right part of the sub-tree) but, nevertheless, we have not reduced the number of tridents.

(iii) Both σ_3 and σ_4 are $\bullet$. In this case, in the right part of the tree we had, before the elimination, a trident. Thus, after the elimination we reduced both the number of dead-ends and tridents by one, while the number of cherries actually increased by one.

As we will prove in Lemma 3.2.13 below, the total number of tridents in $[\tau, \tau']$ is strictly larger than the number of dead-ends. In all three cases above, the elimination procedure preserves this inequality. Indeed, in (i) the number of dead-ends is reduced by 2 while the number of tridents remains the same, in (ii) both the number of tridents and dead-ends remain the same (the number of cherries is reduced by one but this has no effect) and in (iii) both the number of tridents and dead-ends is reduced by 1 (the number of cherries increases by 1). Thus, continuing to eliminate dead-ends, we necessarily end up with a sub-tree of $[\tau, \tau']_\gamma$, which will only contain cherries and tridents. We can now return to the beginning of the proof and the situation of a (sub-ternary) tree that consists of only tridents and cherries, which necessarily contains a v-cycle. □

Lemma 3.2.13. *Let $\tau \in \mathcal{T}_3 \setminus \{\bullet\}$, then $|\mathcal{V}_{\bullet}(\tau)| \leqslant |\mathcal{V}_{\bullet\bullet}(\tau)| - 1$.*

Proof. Clearly the statement is true for $\tau =$  . Any larger tree in $\mathcal{T}_3$ can be constructed from  by successively gluing tridents  onto leaves. Now, there are three possibilities for a trident to be glued onto an existing tree in $\tau \in \mathcal{T}_3$, as described in the table below.

Pre-gluing	Post-gluing	$\lvert\mathcal{V}_{\bullet}\rvert$	$\lvert\mathcal{V}_{\bullet\bullet}\rvert$
		$+0$	$+0$
		$+1$	$+1$
		-1	$+1$

Here, $\sigma_1, \sigma_2 \in \mathcal{T}_3 \setminus \{\bullet\}$ are placeholders for corresponding sub-trees. In all three cases the claimed inequality remains true as we can only create a new dead-end by creating a trident at the same time. □

3.2.4 Removal of v-cycles of arbitrary length

The estimates in this section provide a quantitative control on v-cycles, given in Lemmas 3.2.15 and 3.2.16 below. An important procedure that we will follow, is to spatially decouple v-cycles from the rest of the kernel encoded by the tree. We will be performing such decoupling estimates sequentially until we exhaust all v-cycles, including the v-cycles that will emerge through this process. We call this *cycle removal*, defined in the following in the instance of a single cycle removal.

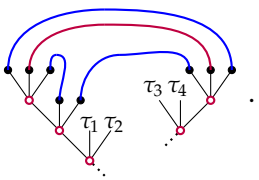
Definition 3.2.14 (Cycle removal)**.** *For any contracted tree τ_κ, with $\tau \in \mathcal{T} \setminus \{\mathbf{1}, \bullet\}$ and any* v*-cycle $\mathcal{C}$ in τ_κ, we define the contracted tree $(\tau \setminus \mathcal{C})_{\widetilde{\kappa}}$ obtained from τ_κ through the following procedure:*

(i) *Remove from τ_κ all the edges and vertices that belong to $\mathcal{C}$.*

(ii) *If v is an inner node that belongs to $\mathcal{C}$, we also remove all other edges connected to v. Moreover, for any descendant u of v, other than the one contained in $\mathcal{C}$, we replace $\{u, v\}$ by $\{u, w\}$, where w is the closest ancestor of u not contained in $\mathcal{C}$. If w does not exist, we simply remove the edge.*

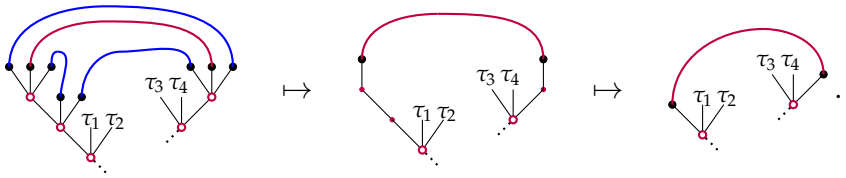
The result of this procedure is a tree $\tau \setminus \mathcal{C}$ with the contraction $\widetilde{\kappa}$ consisting of the remaining edges in κ after the above two steps.

For example, consider the following component of a tree



In this example there are three v-cycles: one whose contraction-edges consists of only the blue edges, one whose κ edges consists of the top blue edge and the red one and, finally, one whose κ edges consists of the red one and the two blue ones below the red. The process of removing the blue v-cycle is depicted below. The middle step shows the component after removing the leaves and edges that are part of the v-cycle. The rightmost tree is the final

outcome after also replacing the edges incident to the inner vertices of the v-cycle by a single edge.



The following lemma provides the central estimate that quantifies the contributions coming from v-cycles.

Lemma 3.2.15. *Let τ be of the form $\tau = [\tau_1 \cdots \tau_n]$, $\tau_i \in \mathcal{T}_3$, and $\kappa \in \mathcal{K}(\tau)$. Let $\mathcal{C}$ be a m-cycle in τ_κ, with inner vertices denoted by $v_1, ..., v_m$ associated to time-space coordinates $(s_{v_1}, y_{v_1}), ..., (s_{v_m}, y_{v_m})$. Recall the kernel $\widehat{K}^{t,x}_{\tau_{\kappa,\varepsilon}}(s_{\mathcal{V}}, y_{\mathcal{V}})$ from (3.40) associated to a contracted tree. Then*

$$\int_{(\mathbf{R}^2)^{\mathcal{I}_\mathcal{C}}} \int_{D_t^{\mathcal{L}_\mathcal{C}}} \widehat{K}^{t,x}_{\tau_{\kappa,\varepsilon}}(s_{\mathcal{V}}, y_{\mathcal{V}}) \, \mathrm{d}s_{\mathcal{L}_\mathcal{C}} \, \mathrm{d}y_{\mathcal{V}_\mathcal{C}} \tag{3.51}$$
$$\leqslant \; \varepsilon^{\otimes m}(s_{v_1}, ..., s_{v_m}) \left\{ \prod_{v \in \mathcal{I}_\mathcal{C}} \mathbb{1}_{s_{\mathfrak{d}_\kappa(v)} \leqslant s_v \leqslant s_{\mathfrak{p}(v)}} \right\} \widehat{K}^{t,x}_{\widetilde{\tau}_{\widetilde{\kappa},\varepsilon}}(s_{\mathcal{V}\setminus\mathcal{C}}, y_{\mathcal{V}\setminus\mathcal{C}}) \, ,$$

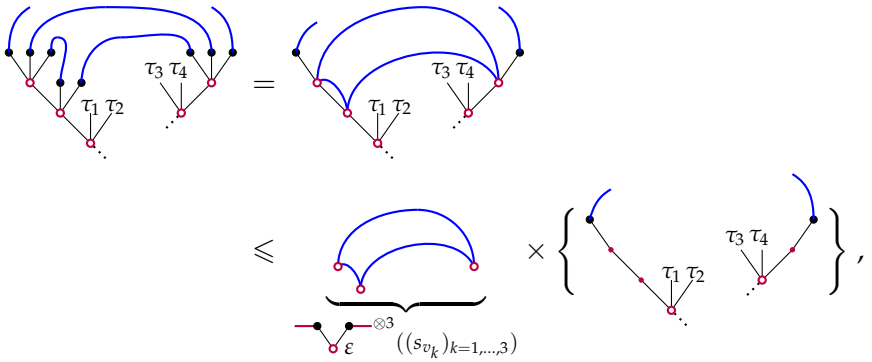
*where $\widetilde{\tau} = \tau \setminus \mathcal{C}$, $\widetilde{\kappa}$ is the contraction induced by κ on $\widetilde{\tau}$. The node $\mathfrak{p}(v)$ denotes the parent of a vertex v, in τ, and $\mathfrak{d}_\kappa(v)$ denotes the unique descendant of v which is **not** part of $\mathcal{C}$. The kernel $\varepsilon^{\otimes m}$ was defined in (3.50).*

Proof. The proof follows the steps of the computation in Example 3.2.10 by crucially applying Chapman–Kolmogorov to the integration over the space variables associated to the leaves of the v-cycle, combined with a uniform bound over the space variables on the resulting product of heat kernels. In particular, we have

$$\begin{aligned}
&\int_{(\mathbf{R}^2)^{\mathcal{L}_\mathcal{C}}} \prod_{w \in \mathcal{L}_\mathcal{C}} \lambda_\varepsilon e^{\mathfrak{m} s_{\mathfrak{p}(w)}} p_{s_{\mathfrak{p}(w)} + \varepsilon^2}(y_w - y_{\mathfrak{p}(w)}) \prod_{\{u,u'\} \in \kappa \cap \mathcal{E}_\mathcal{C}} \delta_0(y_u - y_{u'}) \, \mathrm{d}y_{\mathcal{L}_\mathcal{C}} \\
&\quad = \lambda_\varepsilon^{2m} \prod_{k=1}^{m} e^{\mathfrak{m}(s_{v_k} + s_{v_{k+1}})} p_{s_{v_k} + s_{v_{k+1}} + 2\varepsilon^2}(y_{v_{k+1}} - y_{v_k}) \\
&\quad \leqslant \lambda_\varepsilon^{2m} \prod_{k=1}^{m} e^{\mathfrak{m}(s_{v_k} + s_{v_{k+1}})} p_{s_{v_k} + s_{v_{k+1}} + 2\varepsilon^2}(0) \\
&\quad = \varepsilon^{\otimes m}((s_{v_k})_{k=1,...,m}) \, ,
\end{aligned}$$

where we omitted integration over the time variables associated to leaves on the left-hand side, as $s_w = 0$ for all $w \in \mathcal{L}_\mathcal{C}$ due to the Dirac–δ at zero.

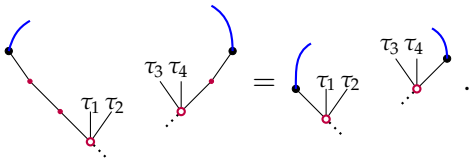
Inserting the above into the left-hand side of (3.51) we obtain the desired estimate. It is useful to have a pictorial representation of the estimate we have just performed:



with the equality depicting the application of Chapman-Kolmogorov and the inequality representing the application of the uniform bound

$$\prod_{k=1}^{m} e^{\mathfrak{m}(s_{v_k}+s_{v_{k+1}})} p_{s_{v_k}+s_{v_{k+1}}+2\varepsilon^2}(y_{v_{k+1}} - y_{v_k}) \leqslant \prod_{k=1}^{m} e^{\mathfrak{m}(s_{v_k}+s_{v_{k+1}})} p_{s_{v_k}+s_{v_{k+1}}+2\varepsilon^2}(0)\,.$$

The blue cycle appearing in the right-hand side represents the kernel $\overset{\otimes m}{\underset{\varepsilon}{\curlyvee}}((s_{v_k})_{k=1,\dots,m})$, which is spatially decoupled from the rest of the tree integrations. The small purple nodes indicate the remaining spatial integrals associated to inner nodes of the extracted cycle. By another application of Chapman–Kolmogorov, when integrating over the mentioned spatial variables associated to the small purple nodes, we obtain



This concludes the proof. □

The following lemma is crucial as it demonstrates that v-cycles of length larger than one have a vanishing contribution, as $\varepsilon \to 0$, and in fact, the contribution is smaller the larger the cycle.

Lemma 3.2.16. *The following bound holds for any $m \geqslant 1$*

$$\int_{[0,t]^m} \overset{\otimes m}{\underset{\varepsilon}{\curlyvee}}(s_1,\dots,s_m)\,\mathrm{d}s_{1,\dots,m} \leqslant \frac{(\lambda_\varepsilon e^{\overline{\mathfrak{m}}\,t})^{2m}}{2^m\,\pi} \log\left(1+\tfrac{t}{\varepsilon^2}\right), \tag{3.52}$$

where we recall that $\overline{\mathfrak{m}} = \mathfrak{m} \vee 0$.

Proof. In the case $m = 1$, the bound follows from (3.48), since

$$\int_0^t \text{  }_\varepsilon(s)\,\mathrm{d}s = \lambda_\varepsilon^2 \int_0^t \frac{e^{2\mathfrak{m}s}}{4\pi(s+\varepsilon^2)}\,\mathrm{d}s \leqslant \frac{\lambda_\varepsilon^2 e^{2\overline{\mathfrak{m}}t}}{2\pi} \log\left(1+\tfrac{t}{\varepsilon^2}\right). \tag{3.53}$$

Thus, we assume $m \geqslant 2$ for the remainder of the proof. First, note that since $a^2 + b^2 \geqslant 2ab$, we have

$$s_1 + s_m + 2\varepsilon^2 \geqslant 2\sqrt{(s_1+\varepsilon^2)(s_m+\varepsilon^2)} \geqslant \sqrt{(s_1+2\varepsilon^2)(s_m+2\varepsilon^2)}\,.$$

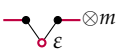
Therefore,

$$\begin{aligned}
&\int_{[0,t]^m} \text{  }^{\otimes m}(s_1,\dots,s_m)\,\mathrm{d}s_{1,\dots,m} \\
&= \frac{\lambda_\varepsilon^{2m}}{(2\pi)^m} \int_{[0,t]^m} \prod_{k=1}^m \frac{e^{\mathfrak{m}(s_k+s_{k+1})}}{s_k+s_{k+1}+2\varepsilon^2}\,\mathrm{d}s_{1,\dots,m} \\
&\leqslant \frac{(\lambda_\varepsilon e^{\overline{\mathfrak{m}}t})^{2m}}{(2\pi)^m} \int_{[0,t]^m} \frac{1}{\sqrt{s_1+2\varepsilon^2}} \frac{1}{\sqrt{s_m+2\varepsilon^2}} \prod_{k=1}^{m-1} \frac{1}{s_k+s_{k+1}+2\varepsilon^2}\,\mathrm{d}s_{1,\dots,m}\,,
\end{aligned} \tag{3.54}$$

where we additionally used that $e^{\mathfrak{m}s_k} \leqslant e^{\overline{\mathfrak{m}}t}$. Furthermore, for $2 \leqslant k \leqslant m$, using the change of variables $r = \frac{s_k}{s_{k-1}+2\varepsilon^2}$ together with the identity $\int_0^\infty \frac{1}{\sqrt{r}(1+r)}\,\mathrm{d}r = \pi$, we have that

$$\begin{aligned}
\int_0^t \frac{1}{\sqrt{s_k+2\varepsilon^2}} \frac{1}{s_{k-1}+s_k+2\varepsilon^2}\,\mathrm{d}s_k &\leqslant \frac{1}{\sqrt{s_{k-1}+2\varepsilon^2}} \int_0^{t/(s_{k-1}+2\varepsilon^2)} \frac{1}{\sqrt{r}} \frac{1}{1+r}\,\mathrm{d}r \\
&\leqslant \frac{\pi}{\sqrt{s_{k-1}+2\varepsilon^2}}\,.
\end{aligned} \tag{3.55}$$

Applying (3.55) $(m-1)$-times to (3.54), starting from $k = m$ and going down to $k = 2$, yields

$$\begin{aligned}
&\int_{[0,t]^m} \text{  }^{\otimes m}(s_1,\dots,s_m)\;\mathrm{d}s_{1,\dots,m} \\
&\leqslant \frac{(\lambda_\varepsilon e^{\overline{\mathfrak{m}}t})^{2m}\,\pi^{m-1}}{(2\pi)^m} \int_0^t \frac{1}{\sqrt{s_1+2\varepsilon^2}} \frac{1}{\sqrt{s_1+2\varepsilon^2}}\,\mathrm{d}s_1 \\
&= \frac{(\lambda_\varepsilon e^{\overline{\mathfrak{m}}t})^{2m}}{2^m\,\pi} \log\left(1+\tfrac{t}{2\varepsilon^2}\right),
\end{aligned}$$

which concludes the proof. □

Next we want to define an iterative process of extracting cycles from a paired tree and record this process via a mapping to an element of the permutation group. We will call this the *cycle extraction map* and define it below. To define such algorithm, it will be convenient to label vertices of trees. For a given tree $\tau \in \mathcal{T}_3$, we fix a representative ordered version of it and label the inner vertices of $[\tau, \tau]$ (excluding the root) with the numbers $\{1, \dots, 2i(\tau)\}$ in arbitrary order. Once we have labeled all inner nodes, we label its leaves with the integers $\{2i(\tau)+1, \dots, 4i(\tau)+2\}$. It will be convenient to define an ordering among all sequences of labels.

Definition 3.2.17 (Lexicographic ordering)**.** *For two vectors $\mathsf{V} = (v_1, ..., v_d) \in \mathbf{N}^d$ and $\mathsf{U} = (u_1, ..., u_{d'}) \in \mathbf{N}^{d'}$ with d and d' not necessarily equal, we say that V precedes U in lexicographic order and write $\mathsf{V} \prec \mathsf{U}$ if*

- *either there exists a $k \leqslant d \wedge d'$ such that $v_z = u_z$ for $z \leqslant k-1$ and $v_k < u_k$,*
- *or $d < d'$ and $v_z = u_z$ for $z \leqslant d$.*

The lexicographic order extends naturally to a total order on the set of v-cycles of a tree. Let $\mathcal{C}$ be a v-cycle, represented by the path-vector

$$\mathsf{V}_{\mathcal{C}} := (v_{i_1}, v_{j_1}, v_{j_2}, v_{i_2}, ..., v_{i_m}, v_{j_{2m-1}}, v_{j_{2m}}), \quad \text{with} \quad v_{i_k} \in \mathcal{I}_{\mathcal{C}}\,,\ v_{j_k} \in \mathcal{L}_{\mathcal{C}}\,, \tag{3.56}$$

with consecutive vertices in the vector being connected by an edge in the v-cycle. Here i_1 is the minimal label in $\mathcal{I}_{\mathcal{C}}$, and v_{j_1} the leaf with minimal label neighbouring v_{i_1}. This imposes a direction the path-vector is represented in. Now, let $\mathcal{C}'$ be a second v-cycle represented by the vector $\mathsf{V}_{\mathcal{C}'}$, then

$$\mathcal{C} \prec \mathcal{C}' \qquad \text{if} \qquad \mathsf{V}_{\mathcal{C}} \prec \mathsf{V}_{\mathcal{C}'} \text{ lexicographically.} \tag{3.57}$$

In this setting we can introduce the **cycle extraction map** Π_τ, for any $\tau \in \mathcal{T}_3$.

Definition 3.2.18 (Cycle extraction map)**.** *For any $\tau \in \mathcal{T}_3$ and $\gamma \in \mathcal{Y}(\tau, \tau)$, we will define inductively a sequence of* v*-cycles extracted from $[\tau, \tau]_\gamma$ as follows:*

(i) *Start by defining $\sigma_1 := [\tau, \tau]_\gamma$ and $\gamma_1 := \gamma$ and denote by $\mathcal{C}_1$ the minimal* v*-cycle in σ_1 (whose existence is guaranteed by Lemma 3.2.12) with respect to the lexicographic order. Define $\sigma_2 := \sigma_1 \setminus \mathcal{C}_1$ and on σ_2 the contraction γ_2 induced by γ_1 after the cycle removal of $\mathcal{C}_1$, according to Definition 3.2.14.*

(ii) *Assume that we have defined the contracted trees $(\sigma_i)_{\gamma_i}$, for $i = 1, ..., k$, as well as the* v*-cycles $\mathcal{C}_1, ..., \mathcal{C}_{k-1}$ belonging respectively to $(\sigma_1)_{\gamma_1}, ..., (\sigma_{k-1})_{\gamma_{k-1}}$.*

Then proceed by defining $\mathcal{C}_k$ to be the minimal v*-cycle belonging to σ_k, with respect to the lexicographic order. Further, define the contracted graph $(\sigma_{k+1})_{\gamma_{k+1}} := (\sigma_k \setminus \mathcal{C}_k)_{\tilde{\gamma}_k}$ via the cycle removal as in Definition 3.2.14.*

(iii) *Stop at $K(\tau,\gamma) := k$ if $\sigma_{k+1} =$*  .

Definition 3.2.19 (Permutation extraction map)**.** *Let $\tau \in \mathcal{T}_3$ and $i(\tau)$ denote the number of inner nodes of τ. We define as follows the permutation extraction map*

$$\Pi_\tau \colon \mathcal{Y}(\tau,\tau) \to S_{2i(\tau)}\,,$$

with S_n denoting the symmetric group of n elements. First, for any $\gamma \in \mathcal{Y}(\tau,\tau)$ consider the sequence of v*-cycles $(\mathcal{C}_k)_{k=1}^{K(\tau,\gamma)}$ constructed from $[\tau,\tau]_\gamma$ via the cycle extraction map from Definition 3.2.18. For any $\mathcal{C}_k$ belonging to this sequence let $v_{i_1^{(k)}},\dots,v_{i_{m_k}^{(k)}}$ be the vertices in $\mathcal{I}_{\mathcal{C}_k}$, listed in the same order as in (3.56). We then map every cycle to a permutation cycle*

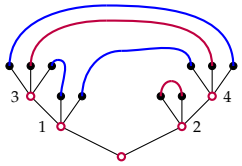
$$\mathcal{C}_k \mapsto \widehat{\mathcal{C}}_k := \big(i_1^{(k)}\ i_2^{(k)}\ \cdots\ i_{m_k}^{(k)}\big) \in S_{2i(\tau)}\,,$$

where we used the cycle notation $(i_1^{(k)}\ i_2^{(k)}\ \cdots\ i_{m_k}^{(k)})$ for the permutation $i_j^{(k)} \mapsto i_{j+1}^{(k)}$. The permutation extraction map $\Pi_\tau(\gamma)$ is then defined as the permutation $\pi \in S_{2i(\tau)}$ with cycle decomposition

$$\pi = \prod_{k=1}^{K(\tau,\gamma)} \big(i_1^{(k)}\ i_2^{(k)}\ \cdots\ i_{m_k}^{(k)}\big)\,.$$

To clarify the tools we have introduced so far, let us discuss an example.

Example 3.2.20. *Consider the following paired tree*

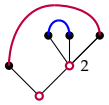
$[\tau,\tau]_\gamma =$ 

$= (\sigma_1)_{\gamma_1}\,,$

where we marked the minimal v*-cycle (w.r.t. lexicographic ordering) in blue, which will be removed in the first iteration of the cycle extraction. For the sake of clarity we ommited the labels of leaves in the diagaram above, and merely represent the*

corresponding v*-cycles by their inner nodes. This yields*

$$(\sigma_2)_{\gamma_2} = \quad \textit{and} \quad \widehat{\mathcal{C}}_1 = (1\,3\,4)\,,$$



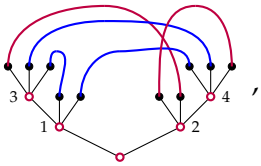
where once more we marked the new minimal v*-cycle in blue. The final iteration then yields*

$$(\sigma_3)_{\gamma_3} = \quad \textit{and} \quad \widehat{\mathcal{C}}_2 = (2)\,.$$

Overall, for the above example, the permutation extraction map Π_τ *yields*

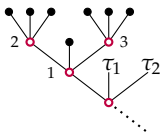
$$\gamma \mapsto (1\,3\,4)(2) \in S_4\,. \tag{3.58}$$

Some remarks on the permutation extraction map are due. First, we note that the mapping is well defined. This is because once a minimal v-cycle is to be extracted, the integers indexing its inner nodes are removed from the permutation and the new tree has inner nodes indexed by the remaining integers. A second observation is that it is not bijective. To see that the map is not injective, consider the tree from Example 3.2.20, however, now with the pairing



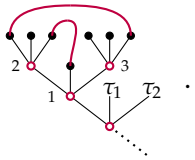
,

where again we marked the minimal v-cycles in blue. Then Π_τ maps the above pairing also to the permutation (3.58). Moreover, the map is not surjective: For example consider a paired tree $[\tau, \tau]$ containing the component

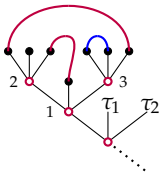


and a permutation $\pi \in S_n$ containing the permutation cycle $(1\,2\,3)$. Then there exists no pairing $\gamma \in \mathcal{Y}(\tau, \tau)$ such that $\Pi_\tau(\gamma) = \pi$, since it is not possible to construct a v-cycle according to this permutation cycle. Let us try to construct a corresponding v-cycle and see that this fails. Necessarily, γ would contain an edge contracting a leaf connected to v_2 and v_3, respectively. Moreover, we can only connect a single leaf neighbouring either v_2 or v_3 to the isolated leave neighbouring v_1, say we choose a leaf at v_2. Then the pairing γ

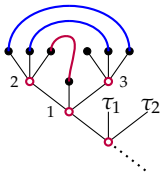
would contain the following edges (in purple):



Now, it is not possible to close a v-cycle with v_3, while also crossing v_1, v_2. Either we create a 1-cycle at v_3 or pair one of the leaves neighbouring v_3 with the remaining leave at v_2, which will create the v-cycle of length two that only contains v_2 and v_3. This is illustrated as follows, where we marked the smaller cycles in blue



and



Note that this construction does not depend on the specific leaves we chose. In particular, the roles of v_2 and v_3 can be reversed.

The main result of this section is an upper bound on the integral represented by a paired tree obtained via the permutation cycle sequence extracted by the map Π_τ. This upper bound will turn out to be sharp when $\Pi_\tau(\gamma) = \mathrm{Id}$, meaning that only cycles of length one are extracted. Before we state the result, let us introduce the following notation for the time-simplex induced by a tree. For a tree $\tau \in \mathcal{T}$ we define

$$\mathfrak{D}_\tau(t) := \{s \in [0,t]^{\mathcal{I}(\tau)\setminus \mathfrak{o}} \ : \ \text{if } \mathfrak{p}(u) = v \text{ , then } s_v \geq s_u\}\,, \tag{3.59}$$

with the usual convention that $s_{\mathfrak{o}} = t$.

Lemma 3.2.21. *Let $\tau \in \mathcal{T}_3$ and $i(\tau)$ be the number of internal nodes of τ. For $\pi \in S_{2i(\tau)}$ with permutation cycle decomposition $\{\widehat{\mathcal{C}}_i\}_{i=1}^{K(\pi)}$ and $\gamma \in \Pi_\tau^{-1}(\pi)$, we have*

$$[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \lambda_\varepsilon^2 \Big(\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}} \Big) e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0)\,,$$

where $\mathcal{I} := \mathcal{I}([\tau,\tau]) \setminus \mathfrak{o}$ *and*

$$\Psi_{\pi,\varepsilon}(s) := \prod_{i=1}^{K(\pi)} \varepsilon^{\otimes|\widehat{\mathcal{C}}_i|}(s_v; v \in \mathcal{I}_{\widehat{\mathcal{C}}_i}), \tag{3.60}$$

with $\varepsilon^{\otimes n}$ *defined in* (3.50). *Here* $\mathcal{I}_{\widehat{\mathcal{C}}}$ *denotes all those inner vertices whose labels lie in the permutation cycle* $\widehat{\mathcal{C}}$. *In the case* $\pi = \mathrm{Id}$, *the inequality can be replaced with an equality.*

Proof. Let $\pi \in S_{2i(\tau)}$ and $\gamma \in \Pi_\tau^{-1}(\pi)$. The proof works by extracting cycles successively from $[\tau,\tau]$ and using Lemma 3.2.15 to obtain a bound on $[\tau,\tau]_{\gamma,\varepsilon}$ in terms of these cycles. We write $\mathcal{V} = \mathcal{V}([\tau,\tau]) \setminus \mathfrak{o}$ and $\mathcal{I} = \mathcal{I}([\tau,\tau]) \setminus \mathfrak{o}$. Starting with the extraction of the minimal v-cycle $\mathcal{C}_1$ in $[\tau,\tau]_\gamma$, we have

$$\begin{aligned}[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \int_{D_t^{\mathcal{V}\setminus\mathcal{C}_1}} \widehat{K}^{t,x}_{([\tau,\tau]\setminus\mathcal{C}_1)_{\tilde{\gamma}}}(s_{\mathcal{V}\setminus\mathcal{C}_1}, y_{\mathcal{V}\setminus\mathcal{C}_1}) \\ \int_{[0,t]^{\mathcal{I}_{\mathcal{C}_1}}} \varepsilon^{\otimes|\mathcal{I}_{\mathcal{C}_1}|}(s_v; v \in \mathcal{I}_{\mathcal{C}_1}) \left\{ \prod_{v\in\mathcal{I}_{\mathcal{C}_1}} \mathbb{1}_{s_{\mathfrak{d}_{\tau,\gamma}(v)} \leqslant s_v \leqslant s_{\mathfrak{p}_\tau(v)}} \right\} \\ \mathrm{d}s_{\mathcal{I}_{\mathcal{C}_1}}\, \mathrm{d}s_{\mathcal{V}\setminus\mathcal{C}_1}\, \mathrm{d}y_{\mathcal{V}\setminus\mathcal{C}_1}\,,\end{aligned} \tag{3.61}$$

where $\widehat{K}^{t,x}_{([\tau,\tau]\setminus\mathcal{C}_1)_{\tilde{\gamma}}}$ denotes the kernel corresponding to the fully contracted tree $([\tau,\tau]\setminus\mathcal{C}_1)_{\tilde{\gamma}} = (\sigma_1)_{\tilde{\gamma}}$, following the notation in Definition 3.2.14. We can then proceed iteratively by extracting the v-cycles via the cycle extraction map from Definition 3.2.18, until we reach the tree $\bullet\!\!\frown\!\!\bullet$. In this way we obtain the upper bound, using Lemma 3.2.15,

$$\begin{aligned}[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \bullet\!\!\frown\!\!\bullet_\varepsilon(t,x) \\ \times \int_{[0,t]^{\mathcal{I}}} \prod_{i=1}^{K(\tau,\gamma)} \left\{ \varepsilon^{\otimes|\mathcal{I}_{\mathcal{C}_i}|}(s_v; v \in \mathcal{I}_{\mathcal{C}_i}) \left\{ \prod_{v\in\mathcal{I}_{\mathcal{C}_i}} \mathbb{1}_{s_{\mathfrak{d}_{\sigma_i,\gamma}(v)} \leqslant s_v \leqslant s_{\mathfrak{p}_{\sigma_i}(v)}} \right\} \right\} \mathrm{d}s_{\mathcal{I}}\,,\end{aligned} \tag{3.62}$$

with the sequence of v-cycles $(\mathcal{C}_k)_{k=1}^{K(\tau,\gamma)}$ and the sequence of reduced trees $(\sigma_i)_{i=1}^{K(\tau,\gamma)}$ from the cycle extraction map. Note that $K(\tau,\gamma)$ equals the number of permutation cycles $K(\pi)$. In order to avoid confusion, we added a subindex $\mathfrak{p}_{\sigma_i}$ to the parent map $\mathfrak{p}$ (and also to the descendent map $\mathfrak{d}_\gamma$), making clear with respect to which tree the map is to be interpreted. Now, by

application of Lemma 3.2.22 below, the inequality (3.62) reads

$$[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \text{  }_{\varepsilon}(t,x) \int_{\mathfrak{D}_{[\tau,\tau]}(t)} \prod_{i=1}^{K(\tau,\gamma)} \text{  }_{\varepsilon}^{\otimes|\mathcal{I}_{\mathcal{C}_i}|}(s_v; v\in\mathcal{I}_{\mathcal{C}_i})\,\mathrm{d}s_{\mathcal{I}}$$
$$= \lambda_\varepsilon^2 e^{2\mathfrak{m}\,t} p_{2(t+\varepsilon^2)}(0)\left(\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}}\right),$$

which yields the desired uppper bound. If $\pi = \mathrm{Id}$, then the inequality in the first line becomes an equality as we are successively removing 1-cycles and apply the identity in Lemma 3.2.6, rather than the upper bound in Lemma 3.2.15. The proof is complete. □

Lemma 3.2.22. *Let $t > 0$, $\tau \in \mathcal{T}_3$ and write $\mathcal{I} := \mathcal{I}([\tau,\tau]) \setminus \mathfrak{o}$. Then for every pairing $\gamma \in \mathcal{Y}(\tau,\tau)$, we have*

$$[0,t]^{\mathcal{I}} \cap \bigcap_{i=1}^{K(\tau,\gamma)} \bigcap_{v\in\mathcal{I}_{\mathcal{C}_i}} \{s_{\mathfrak{d}_{\sigma_i,\gamma}(v)} \leqslant s_v \leqslant s_{\mathfrak{p}_{\sigma_i}(v)}\} = \mathfrak{D}_{[\tau,\tau]}(t)\,, \tag{3.63}$$

where $(\mathcal{C}_k)_{k=1}^{K(\tau,\gamma)}$ denotes the sequence of v*-cycles and $(\sigma_i)_{i=1}^{K(\tau,\gamma)}$ the sequence of reduced trees, constructed from $[\tau,\tau]_\gamma$ via the cycle extraction map (Definition 3.2.18). The set $\mathfrak{D}_{[\tau,\tau]}(t)$ was defined in (3.59).*

Proof. Let $v \in \mathcal{I}$ be arbitrary. Then there exists an $i = 1,\dots,K(\tau,\gamma)$ such that $v \in \mathcal{I}_{\mathcal{C}_i}$ and we define $u_1 := \mathfrak{d}_{\sigma_i,\gamma}(v)$, $w_1 := \mathfrak{p}_{\sigma_i}(v)$. Notably, there exists a unique "ancestral" path $(u_1,\dots,u_m,v,w_{m'},\dots,w_1)$ in the tree $[\tau,\tau]$ with $u_j, w_j \in \mathcal{I}$, $j \geqslant 2$.

If $s_{\mathcal{I}} \in \mathfrak{D}_{[\tau,\tau]}(t)$, then

$$0 \leqslant s_{\mathfrak{d}_{\sigma_i,\gamma}(v)} = s_{u_1} \leqslant \cdots \leqslant s_{u_m} \leqslant s_v \leqslant s_{w_{m'}} \leqslant \cdots \leqslant s_{w_1} = s_{\mathfrak{p}_{\sigma_i}(v)} \leqslant t\,,$$

which implies in particular that $s_{\mathfrak{d}_{\sigma_i,\gamma}(v)} \leqslant s_v \leqslant s_{\mathfrak{p}_{\sigma_i}(v)}$.

On the other hand, note that $\mathfrak{p}(v) = \mathfrak{p}_{[\tau,\tau]}(v) = w_{m'}$, which lies in $\mathcal{I}_{\mathcal{C}_{i'}}$ for some $i' = 1,\dots,K(\tau,\gamma)$. Thus, if $s_{\mathcal{I}}$ lies in the left-hand side of (3.63), then

- either the parent of v has not been removed by the cycle extraction map in an earlier iteration, i.e. $i' \geqslant i$, in which case

$$s_v \leqslant s_{\mathfrak{p}_{\sigma_i}(v)} = s_{w_{m'}} = s_{\mathfrak{p}(v)}\,,$$

- or the parent of v has been removed in an a previous iteration, i.e. $i' < i$,

then

$$s_v = s_{\mathfrak{d}_{\sigma_{i'},\gamma}(w_{m'})} \leqslant s_{w_{m'}} = s_{\mathfrak{p}(v)} \,.$$

As the choice of v was arbitrary, this concludes the proof. □

3.3 Contributing and non-contributing trees

This section is dedicated to the proof of Proposition 3.1.2, for a single fixed tree $\tau \in \mathcal{T}_3$. Obtaining this result requires a precise quantitative control over the limiting behavior of contracted and paired trees. Such control will build on a systematic application of the bounds and ideas that we have introduced in Section 3.2, and in particular in Lemma 3.2.21.

In the previous section we analysed **paired** trees and found that it was possible to identify v-cycles and remove them iteratively to obtain an upper bound (or an exact estimate) on the integral associated to such a tree. In this section we start instead with an arbitrary **contracted** ternary tree $[\tau]_\kappa$. Our objective is to obtain a bound (or an exact estimate) on the second moment of the Wiener–Itô stochastic integral associated to such contracted tree. To obtain these estimates, we must sum over all possible pairings γ of $[\tau, \tau]$ which complete the contraction κ, and for each such pairing we can follow the procedure described in the previous section. One of the key points of this section is therefore to keep track of all the combinatorial factors that appear when counting pairings and contractions associated to arbitrarily large trees.

The remainder of this section will be split into two main parts. First, we study those contracted trees that do not vanish in the limit $\varepsilon \to 0$ and identify them using properties of the underlying graph τ_κ. We also determine their precise contribution and the size of the set of all such contractions. In the second part, we will instead state and prove a uniform upper bound for the rate of convergence of contracted trees that vanish as $\varepsilon \to 0$.

3.3.1 Contributing contractions

We start by studying those contracted trees that contribute to the fluctuations in the limit $\varepsilon \to 0$, and for which an exact estimate of the contribution is necessary. For this reason, let us define **contributing contractions** as follows.

Definition 3.3.1. *For any tree $\tau \in \mathcal{T}_3$ and contraction $\kappa \in \mathcal{K}(\tau)$, we say that κ*

contributes *if there exist* $(t,x) \in (0,\infty) \times \mathbf{R}^2$ *such that*

$$\limsup_{\varepsilon \to 0} \left(\log \tfrac{1}{\varepsilon}\right) \cdot \mathbb{E}\left[\left|[\tau]_{\varepsilon,\kappa}(t,x)\right|^2\right] > 0\,. \tag{3.64}$$

We denote the set of all contributing contractions by

$$C(\tau) := \{\kappa \text{ such that } \tau_\kappa \text{ contributes}\} \subseteq \mathcal{K}(\tau)\,.$$

Identifying contributing contractions

Before determining the precise contribution of contracted trees, we show that the abstract condition in (3.64) can be replaced by a condition on the underlying graph structure of the contracted tree. More precisely, we will see that the estimates of Section 3.2.4 imply that contracted trees contribute if and only if the corresponding integrals lie in the first homogeneous Wiener chaos and we can iteratively remove 1-cycles from them.

Lemma 3.3.2. *Let* $\tau \in \mathcal{T}_3$*, then* $C(\tau)$*, see Definition 3.3.1, satisfies:*

(i) $C(\tau) = \{\kappa \in \mathcal{K}(\tau) : \exists \gamma \in \mathcal{Y}(\tau_\kappa, \tau_\kappa) \text{ with } \Pi_\tau(\gamma) = \mathrm{Id}\}$.

(ii) *If* $\kappa \in C(\tau)$*, then* τ_κ *has a single uncontracted leaf. That is,* $[\tau]_\kappa$ *lies in the first Wiener–Itô chaos and is therefore Gaussian.*

The property of being contributing, associated to a contracted tree $[\tau]_\kappa$, is defined using the second moment condition (3.64). We can express (3.64) by summing $[\tau,\tau]_\gamma$ over all pairings in $\gamma \in \mathcal{Y}(\tau_\kappa,\tau_\kappa)$, recall (3.42). The paired tree obtained, once we fix an element of $\mathcal{Y}(\tau_\kappa,\tau_\kappa)$, can be treated via Lemma 3.2.21. In particular, it turns out that for contributing trees τ_κ and $\gamma \in \mathcal{Y}(\tau_\kappa,\tau_\kappa)$ the cycle extraction map satisfies $\Pi_\tau(\gamma) = \mathrm{Id}$. As a consequence, we can determine precisely the limiting behavior of such a paired trees through the last statement of Lemma 3.2.21, which is the key ingredient in the proof of Lemma 3.3.2 and the content of the following lemma.

Lemma 3.3.3. *Let* $\tau \in \mathcal{T}_3$ *and* $\pi \in S_{2i(\tau)}$*. Then for every* $\gamma \in \Pi_\tau^{-1}(\pi)$ *and all* $(t,x) \in (0,\infty) \times \mathbf{R}^2$

$$\lim_{\varepsilon \to 0} \left(\log \tfrac{1}{\varepsilon}\right) \cdot [\tau,\tau]_{\gamma,\varepsilon}(t,x) = \begin{cases} \hat{\lambda}^2 \left\{ \frac{1}{\tau!} \left(\frac{\hat{\lambda}^2}{2\pi}\right)^{|\tau|} \right\}^2 p_{2t}^{(\mathfrak{m})}(0) & \text{if } \pi = \mathrm{Id}\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

*If the above limit vanishes, we say γ is a **non-contributing pairing**, and call it a **contributing pairing** otherwise.*

The proof of Lemma 3.3.3 uses the following identity.

Lemma 3.3.4. *Let $\tau \in \mathcal{T}_3$ and $\bar{\tau} = \mathscr{T}(\tau)$ be the trimmed tree as in (2.82), then*

$$\int_{\mathfrak{D}_{[\tau]}(t)} \prod_{v \in \mathcal{I}(\tau)} \frac{\lambda_\varepsilon^2 e^{2\mathfrak{m}\, s_v}}{4\pi(s_v + \varepsilon^2)} \, \mathrm{d}s_{\mathcal{I}(\tau)} = \frac{1}{\bar{\tau}!} \left(\frac{\lambda_\varepsilon^2}{4\pi} \int_0^t \frac{e^{2\mathfrak{m}\, s}}{s + \varepsilon^2} \, \mathrm{d}s \right)^{|\tau|} =: [\tau]_\varepsilon(t) \,, \tag{3.65}$$

with the tree-time-simplex $\mathfrak{D}_{[\tau]}(t)$ introduced in (3.59).

This, on the other hand, is a consequence of the fact that the integrand on the left-hand side in (3.65) is a symmetric function in the variables $s_{\mathcal{I}(\tau)} = s_{\mathcal{V}(\tau)}$:

Lemma 3.3.5. *For a tree $\tau \in \mathcal{T}$ of the form $\tau = [\tau_1 \cdots \tau_n]$ such that $\mathcal{I} := \mathcal{I}(\tau) \setminus \mathfrak{o}_\tau \neq \emptyset$ and any symmetric function $\Psi \colon \mathbf{R}^{\mathcal{I}} \to \mathbf{R}$, we have that*

$$\int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} = \frac{1}{\bar{\tau}_1! \cdots \bar{\tau}_n!} \int_{[0,t]^{\mathcal{I}}} \Psi(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} \,,$$

with the domain $\mathfrak{D}_\tau(t)$ defined in (3.59) and $\bar{\tau} = \mathscr{T}(\tau)$ denoting the trimmed tree defined in (2.82).

Proof. We prove the statement by induction over $i(\tau) \geqslant 2$. Note that the case $i(\tau) = 1$ is excluded as the single inner node must necessarily be the root. If $i(\tau) = 2$, then τ must be of the form $\tau = [$  $\bullet \cdots \bullet]$ and $\mathcal{I} = \{v\}$ (denoting the root of ). Then

$$\int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} = \int_0^t \Psi(s_v) \, \mathrm{d}s_v \,,$$

which is the desired statement since $\mathscr{T}($  $)! = 1$. Now assume that the statement holds true for any tree $\tau' = [\tau'_1 \cdots \tau'_{n'}]$ such that $i(\tau') \leqslant N$, for some $N \geqslant 2$, and write $\mathcal{I}' := \mathcal{I}(\tau') \setminus \mathfrak{o}'$ with $\mathfrak{o}' := \mathfrak{o}_{\tau'}$. Let $\tau \in \mathcal{T}$ be of the form $\tau = [\tau' \bullet \cdots \bullet]$, so that $\bar{\tau} = [\bar{\tau}']$. Then

$$\begin{aligned} \int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} &= \int_0^t \int_{\mathfrak{D}_{\tau'}(s_{\mathfrak{o}'})} \Psi(s_{\mathfrak{o}'}, s_{\mathcal{I}'}) \, \mathrm{d}s_{\mathcal{I}'} \, \mathrm{d}s_{\mathfrak{o}'} \\ &= \frac{1}{\bar{\tau}'_1! \cdots \bar{\tau}'_{n'}!} \int_0^t \int_{[0,s_{\mathfrak{o}'}]^{\mathcal{I}'}} \Psi(s_{\mathfrak{o}'}, s_{\mathcal{I}'}) \, \mathrm{d}s_{\mathcal{I}'} \, \mathrm{d}s_{\mathfrak{o}'} \,, \end{aligned}$$

where we used the induction hypothesis and the fact that $\Psi(s_{\mathfrak{o}'}, \cdot)\colon \mathbf{R}^{\mathcal{I}'} \to \mathbf{R}$ is a symmetric function. Using the symmetry of the function Ψ, we further see that the identification of variable $s_{\mathfrak{o}'}$ as the maximum variable is irrelevant with regards to the integration, and the assignment of any of the variables $s_{\mathcal{I}}$ as being the maximum would result in the same value. Thus, the above is equal to

$$\int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}} = \frac{1}{|\mathcal{I}| \cdot \tau_1'! \cdots \tau_{n'}'!} \int_{[0,t]^{\mathcal{I}}} \Psi(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}}\,,$$

and the statement follows now for τ because $\tau'! = |\tau'| \cdot \tau_1'! \cdots \tau_{n'}'! = |\mathcal{I}| \cdot \tau_1'! \cdots \tau_{n'}'!$. Furthermore, we notice that

$$\int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}} = \int_{\mathfrak{D}_{[\tau']}(t)} \Psi(s_{\mathcal{I}(\tau')})\, \mathrm{d}s_{\mathcal{I}(\tau')}\,,$$

meaning that the additional occurrences of $\bullet$ in the grafting of τ do not affect the integral. Finally, consider the case $\tau \in \mathcal{T}$ such that $i(\tau) = N+1$, with $\tau = [\tau_1 \ \cdots \ \tau_k \bullet \cdots \bullet]$, $k \geqslant 2$ and $\tau_i \neq \bullet$. Again the extra occurrences of $\bullet$ in the grafting of τ do not affect the value of the integral. Moreover, the integration domain $\mathfrak{D}_\tau(t)$ can be written as a disjoint union of sub-tree-simplices:

$$\int_{\mathfrak{D}_\tau(t)} \Psi(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}} = \int_{\mathfrak{D}_{[\tau_1]}(t)} \cdots \int_{\mathfrak{D}_{[\tau_k]}(t)} \Psi(s_{\mathcal{I}})\, \mathrm{d}s_{\mathcal{I}(\tau_k)} \cdots \mathrm{d}s_{\mathcal{I}(\tau_1)}\,.$$

The statement follows from the induction hypothesis, since the restriction of Ψ to a subset of variables remains a symmetric function. □

Proof of Lemma 3.3.3. Consider $\tau \in \mathcal{T}_3$, $\pi \in S_{2i(\tau)}$ and $\gamma \in \Pi_\tau^{-1}(\pi)$. Then for all $(t,x) \in (0,\infty) \times \mathbf{R}^2$, Lemma 3.2.21 implies

$$(\log \tfrac{1}{\varepsilon}) \cdot [\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \hat{\lambda}^2 \left\{ \int_{\mathfrak{D}_{[\tau,\tau]}(t)} \Psi_{\pi,\varepsilon}(s)\, \mathrm{d}s_{\mathcal{I}} \right\} e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0)\,,$$

where we remind that $\mathcal{I} := \mathcal{I}([\tau,\tau]) \setminus \mathfrak{o}$. Extending the domain of integration from $\mathfrak{D}_{[\tau,\tau]}(t)$ to $[0,t]^{\mathcal{I}}$, the right-hand side can be factorised

$$(\log \tfrac{1}{\varepsilon}) \cdot [\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \hat{\lambda}^2 \left\{ \prod_{i=1}^{K(\tau,\gamma)} \int_{[0,t]^{\mathcal{I}_{\mathcal{C}_i}}} \Big(\varepsilon \Big)^{\otimes |\mathcal{I}_{\mathcal{C}_i}|} (s_v; v \in \mathcal{I}_{\mathcal{C}_i})\, \mathrm{d}s_{\mathcal{I}_{\mathcal{C}_i}} \right\} e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0)\,,$$

where $(\mathcal{C}_i)_{i=1}^{K(\tau,\gamma)}$ denotes the sequence of v-cycles constructed from $[\tau,\tau]_\gamma$ via

the cycle extraction map (Definition 3.2.18). For each of the integrals we have

$$\int_{[0,t]^{\mathcal{I}_{\mathcal{C}_i}}} \; \varepsilon^{\otimes|\mathcal{I}_{\mathcal{C}_i}|}(s_v; v \in \mathcal{I}_{\mathcal{C}_i})\, \mathrm{d}s_{\mathcal{I}_{\mathcal{C}_i}} \begin{cases} = \frac{\lambda_\varepsilon^2}{4\pi} \int_0^t \frac{e^{2\mathfrak{m}\, s}}{s+\varepsilon^2}\, \mathrm{d}s & \text{if } |\mathcal{I}_{\mathcal{C}_i}| = 1\,, \\ \leqslant \frac{(\lambda_\varepsilon e^{\overline{\mathfrak{m}}\, t})^{2|\mathcal{I}_{\mathcal{C}_i}|}}{2^{|\mathcal{I}_{\mathcal{C}_i}|}\,\pi} \log\left(1 + \frac{t}{\varepsilon^2}\right) & \text{if } |\mathcal{I}_{\mathcal{C}_i}| \geqslant 2\,, \end{cases}$$

where for the case $|\mathcal{I}_{\mathcal{C}_i}| \geqslant 2$ we used Lemma 3.2.16. Note that the right-hand side in the second case vanishes in the limit $\varepsilon \to 0$, since $\lambda_\varepsilon \sim (\log \frac{1}{\varepsilon})^{-1/2}$. In particular, if $\pi \neq \mathrm{Id}$ at least one v-cycle $\mathcal{C}_i$ must satisfy $|\mathcal{I}_{\mathcal{C}_i}| \geqslant 2$, which yields that $[\tau, \tau]_{\gamma,\varepsilon}(t, x)$ vanishes as $\varepsilon \to 0$.

On the other hand, if $\pi = \mathrm{Id}$ all the v-cycles $\mathcal{C}_i$ are 1-cycles and we can replace all inequalities with equalities to obtain

$$\begin{aligned} &(\log \tfrac{1}{\varepsilon}) \cdot [\tau, \tau]_{\gamma,\varepsilon}(t, x) \\ &= \hat{\lambda}^2 \left\{ \int_{\mathfrak{D}_{[\tau,\tau]}(t)} \prod_{v \in \mathcal{I}} \varepsilon^{\otimes 1}(s_v)\, \mathrm{d}s_{\mathcal{I}} \right\} e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0) \\ &= \hat{\lambda}^2 \left\{ \int_{\mathfrak{D}_{[\tau]}(t)} \prod_{v \in \mathcal{I}(\tau)} \frac{\lambda_\varepsilon^2 e^{2\mathfrak{m}\, s_v}}{4\pi(s_v + \varepsilon^2)}\, \mathrm{d}s_{\mathcal{I}(\tau)} \right\}^2 e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0)\,. \end{aligned}$$

Hence, we deduce from Lemma 3.3.4 and Lemma 3.2.7 that

$$\lim_{\varepsilon \to 0} (\log \tfrac{1}{\varepsilon}) \cdot [\tau, \tau]_{\gamma,\varepsilon}(t, x) = \hat{\lambda}^2 \left\{ \frac{1}{\tau!} \left(\frac{\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \right\}^2 p_{2t}^{(\mathfrak{m})}(0)\,,$$

which concludes the proof. □

Finally, we are ready to show that contributing contractions can be identified as those that have a single uncontracted leaf and allow for an iterative removal of 1-cycles.

Proof of Lemma 3.3.2. Let us start by recalling from (3.42) that for any tree $\tau \in \mathcal{T}_3$ and any contraction $\kappa \in \mathcal{K}(\tau)$

$$(\log \tfrac{1}{\varepsilon}) \cdot \mathbb{E}\left[[\tau]_{\kappa,\varepsilon}^2(t, x) \right] = \sum_{\gamma \in \mathcal{Y}(\tau_\kappa, \tau_\kappa)} (\log \tfrac{1}{\varepsilon}) \cdot [\tau, \tau]_{\gamma,\varepsilon}(t, x)\,.$$

Therefore, to prove the first statement of the lemma, it suffices to prove that

$$\lim_{\varepsilon \to 0} (\log \tfrac{1}{\varepsilon}) \cdot [\tau, \tau]_{\gamma,\varepsilon}(t, x) > 0\,,$$

for all $(t,x) \in (0,\infty) \times \mathbf{R}^2$ if and only if $\Pi_\tau(\gamma) = \mathrm{Id}$, which is implied by Lemma 3.3.3.

For the second statement we instead proceed by induction over the number of inner nodes $i(\tau)$. We can check that the statement is true for $i(\tau) = 0$, i.e. $\tau = \bullet$, since $|\mathcal{Y}(\tau,\tau)| = 1$ and

$$\limsup_{\varepsilon\to 0} (\log \tfrac{1}{\varepsilon}) \cdot \mathbb{E}\big[\big|\mathcal{I}_{\circ_\varepsilon}(\bullet)\big|^2(t,x)\big] > 0\,.$$

Now, let $m \in \mathbf{N}$ and assume that the statement holds for all $\hat\tau \in \mathcal{T}_3$ satisfying $i(\hat\tau) \leqslant m$. Choose $\tau \in \mathcal{T}_3$ with $i(\tau) = m+1$ and let $\kappa \in C(\tau)$. By the first point of the present Lemma 3.3.2, which we have just proven, we know there exists a $\gamma \in \mathcal{Y}(\tau_\kappa, \tau_\kappa) \cap \Pi_\tau^{-1}(\mathrm{Id})$. In particular, let $\mathcal{C}_1$ be the first 1-cycle that is extracted by the permutation-extraction map applied to the pairing γ, and write $\sigma_2 = ([\tau,\tau] \setminus \mathcal{C}_1)_{\gamma_2} = [\hat\tau,\tau]_{\gamma_2}$, with $\hat\tau := \tau \setminus \mathcal{C}_1$, assuming without loss of generality that we have removed the cycle from the left tree.

Then via (3.61), we deduce that

$$\begin{aligned}[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant \int_{D_t^{\mathcal{V}\setminus\mathcal{C}_1}} \widehat{K}^{t,x}_{([\tau,\tau]\setminus\mathcal{C}_1)_{\gamma_2}}(s_{\mathcal{V}\setminus\mathcal{C}_1}, y_{\mathcal{V}\setminus\mathcal{C}_1}) \\ \int_0^t \mathcal{I}_{\circ_\varepsilon}^{\otimes 1}(s_v; v \in \mathcal{I}_{\mathcal{C}_1}) \left\{ \prod_{v\in\mathcal{I}_{\mathcal{C}_1}} \mathbb{1}_{s_{\mathfrak{d}_\gamma(v)} \leqslant s_v \leqslant s_{\mathfrak{p}(v)}} \right\} \\ \mathrm{d}s_{\mathcal{I}_{\mathcal{C}_1}}\, \mathrm{d}s_{\mathcal{V}\setminus\mathcal{C}_1}\, \mathrm{d}y_{\mathcal{V}\setminus\mathcal{C}_1}\,,\end{aligned} \tag{3.66}$$

where we again used $\mathcal{V} = \mathcal{V}([\tau,\tau]) \setminus \mathfrak{o}$. Note that the product in the expression above only consists of a single term, because $\mathcal{C}_1$ is a 1-cycle. Dropping the time-constraint encoded by $\mathbb{1}_{s_{\mathfrak{d}_\gamma(v)} \leqslant s_v \leqslant s_{\mathfrak{p}(v)}}$, we therefore obtain

$$[\tau,\tau]_{\gamma,\varepsilon}(t,x) \leqslant [\hat\tau,\tau]_{\gamma_2,\varepsilon}(t,x) \left\{ \int_0^t \mathcal{I}_{\circ_\varepsilon}^{\otimes 1}(s_v; v \in \mathcal{I}_{\mathcal{C}_1})\, \mathrm{d}s_{\mathcal{I}_{\mathcal{C}_1}} \right\}. \tag{3.67}$$

Taking the lim sup over $\varepsilon \to 0$ after multiplying both sides with $(\log \frac{1}{\varepsilon})$ yields

$$\begin{aligned}0 &< \limsup_{\varepsilon\to 0} (\log \tfrac{1}{\varepsilon})\, [\tau,\tau]_{\gamma,\varepsilon}(t,x) \\ &\leqslant \frac{\hat\lambda^2}{2\pi} \limsup_{\varepsilon\to 0} (\log \tfrac{1}{\varepsilon})\, [\hat\tau,\tau]_{\gamma_2,\varepsilon}(t,x)\,.\end{aligned} \tag{3.68}$$

Here, the first inequality holds since $\kappa \in C(\tau)$, while the second inequality is a consequence of (3.67) and Lemma 3.2.7. In particular, the limit on

the right-hand side must be positive. Next, by Definition 3.2.4 there exists $(\kappa_1(\gamma_2), \kappa_2(\gamma_2))$ such that $\gamma_2 \in \mathcal{Y}(\hat{\tau}_{\kappa_1(\gamma_2)}, \tau_{\kappa_2(\gamma_2)})$, with $\kappa_2(\gamma_2) = \kappa$. Via an application of the Cauchy–Schwartz inequality we then obtain

$$[\hat{\tau}, \tau]_{\gamma_2,\varepsilon} \leqslant \mathbb{E}\left[[\hat{\tau}]_{\kappa_1(\gamma_2),\varepsilon}\,[\tau]_{\kappa,\varepsilon}\right] \leqslant \left(\mathbb{E}\left[[\hat{\tau}]^2_{\kappa_1(\gamma_2),\varepsilon}\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[[\tau]^2_{\kappa,\varepsilon}\right]\right)^{\frac{1}{2}}. \tag{3.69}$$

Together with (3.68), this implies $\kappa_1(\gamma_2) \in C(\hat{\tau})$. By the induction assumption, $\hat{\tau}_{\kappa_1(\gamma_2)}$ (note that $i(\hat{\tau}) = m$) has a single uncontracted leaf, which implies that also τ_κ has a single uncontracted leaf. This concludes the desired result. □

Lemma 3.3.2, together with the identity from Lemma 3.3.3, implies that if $\kappa \in C(\tau)$ then $[\tau]_{\varepsilon,\kappa}$ is (and converges to) a mean-zero Gaussian with limiting fluctuation

$$\lim_{\varepsilon\to 0} \sqrt{\log \tfrac{1}{\varepsilon}} \cdot \left\|[\tau]_{\varepsilon,\kappa}(t,x)\right\|_{L^2(\mathbb{P})} = \frac{\hat{\lambda}}{\tau!}\left(\frac{\hat{\lambda}^2}{2\pi}\right)^{|\tau|} \sqrt{p^{(\mathfrak{m})}_{2t}(0)}\,.$$

In the next subsection, we will see that a stronger statement holds true, as we will be able to identify $[\tau]_{\kappa,\varepsilon}$ with $\bullet_\varepsilon$ up to a multiplicative factor.

Determining contributions

In the previous section, we identified contributing pairings (and contractions) to be the ones mapped by Π_τ to the identity permutation, i.e. the algorithm defined in Definition 3.2.18 only extracts cycles of length one. Precisely this fact will turn out to be useful, when determining the following identity for contributing contracted trees.

Lemma 3.3.6. *For every $\tau \in \mathcal{T}_3, \kappa \in C(\tau), \varepsilon \in (0, \frac{1}{2})$ and $(t,x) \in (0,\infty)\times\mathbf{R}^2$ we have that*

$$[\tau]_{\kappa,\varepsilon}(t,x) = \frac{1}{\tau!}\left(\frac{\lambda_\varepsilon^2}{4\pi}\int_0^t \frac{e^{2\mathfrak{m}s}}{s+\varepsilon^2}\,\mathrm{d}s\right)^{|\tau|} \bullet_\varepsilon(t,x) = [\tau]_\varepsilon(t)\,\bullet_\varepsilon(t,x)\,.$$

Note that the right-hand side of the identity in the lemma above does not depend on the particular contraction $\kappa \in C(\tau)$.

Proof. Fix $\tau \in \mathcal{T}_3$ and let $\kappa \in C(\tau)$. Choose $\gamma \in \Pi_\tau^{-1}(\mathrm{Id}) \cap \mathcal{Y}(\tau_\kappa, \tau_\kappa)$, which exists by Lemma 3.3.2(i). Since $\gamma \in \Pi_\tau^{-1}(\mathrm{Id})$, the cycle-extraction map (Definition 3.2.18) associates to γ a sequence of 1-cycles $(\mathcal{C}_i)_{i=1}^{2i(\tau)}$. In order to dis-

tinguish between the two trees generating $[\tau,\tau]$, let us write $[\tau_1,\tau_2] := [\tau,\tau]$. Now, let $(\mathcal{C}'_i)_{i=1}^{i(\tau)}$ be the subset of cycles whose bases belong to the tree τ_1. In other words, $(\mathcal{C}'_i)_{i=1}^{i(\tau)}$ contains all cycles in $(\mathcal{C}_i)_{i=1}^{2i(\tau)}$ such that $\mathcal{I}_{\mathcal{C}_i} \subset \mathcal{I}(\tau_1)$. This yields an iterative rule of removing 1-cycles from $[\tau_1]_\kappa$, which is identical to $[\tau]_\kappa$. In particular, each $\widehat{\mathcal{C}}_i$ corresponds to a unique inner node v_i of τ.

Recall that for $v \in \mathcal{I}(\tau)$ we write $\mathfrak{p}(v) \in \mathcal{I}([\tau])$ for the parent of v and also recall the representation of the stochastic integrals with respect to (3.40), which allows us to write

$$[\tau]_{\kappa,\varepsilon}(t,x) = \int_{D_t^{\mathcal{V}(\tau)}} \widehat{K}^{t,x}_{[\tau]_\kappa,\varepsilon}(s_{\mathcal{V}(\tau)}, y_{\mathcal{V}(\tau)})\ \mathrm{d}s_{\mathcal{V}(\tau)}\ \mathrm{d}y_{\mathcal{V}(\tau)\setminus\ell}\ \eta(\mathrm{d}y_\ell)\,,$$

where we denote by (s_ℓ, y_ℓ) the space-time point associated to the single uncontracted leaf in τ_κ, cf. Lemma 3.3.2. We apply Lemma 3.2.6 with respect to the 1-cycle $\mathcal{C}'_1$, which yields

$$[\tau]_{\kappa,\varepsilon}(t,x) = \int_{D_t^{\mathcal{V}(\tau)\setminus\mathcal{C}'_1}} \Bigg\{ \int_0^t \; \overset{\varepsilon}{\bullet}\!\!\vee\!\!\bullet^{\otimes 1}(s_{v_1}) \mathbb{1}_{\{s_{\mathfrak{d}_\kappa(v_1)} \leqslant s_{v_1} \leqslant s_{\mathfrak{p}(v_1)}\}}\, \mathrm{d}s_{v_1} \Bigg\}$$
$$\widehat{K}^{t,x}_{[\widetilde{\tau}]_{\widetilde{\kappa}},\varepsilon}(s_{\mathcal{V}(\tau)\setminus\mathcal{C}'_1}, y_{\mathcal{V}(\tau)\setminus\mathcal{C}'_1})\, \mathrm{d}s_{\mathcal{V}(\tau)\setminus\mathcal{C}'_1}\, \mathrm{d}y_{\mathcal{V}(\tau)\setminus(\ell\cup\mathcal{C}'_1)}\ \eta(\mathrm{d}y_\ell)\,.$$

Now, by applying Lemma 3.2.6 successively another $i(\tau)-1$ times with respect to each of the 1-cycles $(\mathcal{C}'_i)_{i=2}^{i(\tau)}$, we obtain

$$[\tau]_{\kappa,\varepsilon}(t,x) = \Bigg\{ \int_{[0,t]^{\mathcal{I}(\tau)}} \prod_{v\in\mathcal{I}(\tau)} \; \overset{\varepsilon}{\bullet}\!\!\vee\!\!\bullet^{\otimes 1}(s_v) \mathbb{1}_{\{s_{\mathfrak{d}_\kappa(v)} \leqslant s_v \leqslant s_{\mathfrak{p}(v)}\}}\, \mathrm{d}s_{\mathcal{I}(\tau)} \Bigg\}$$
$$\times\, \lambda_\varepsilon e^{\mathfrak{m}\,t} \int_{\mathbf{R}^2} p_{t+\varepsilon^2}(y_\ell - x)\ \eta(\mathrm{d}y_\ell)\,.$$

The stochastic integral in the second line equals $\overset{\varepsilon}{\bullet}(t,x)$, whereas the time integral in the brackets can be written as

$$\int_{[0,t]^{\mathcal{I}(\tau)}} \prod_{v\in\mathcal{I}(\tau)} \; \overset{\varepsilon}{\bullet}\!\!\vee\!\!\bullet^{\otimes 1}(s_v) \mathbb{1}_{\{s_{\mathfrak{d}_\kappa(v)} \leqslant s_v \leqslant s_{\mathfrak{p}(v)}\}}\, \mathrm{d}s_{\mathcal{I}(\tau)}$$
$$= \int_{\mathfrak{D}_{[\tau]}(t)} \prod_{v\in\mathcal{I}(\tau)} \frac{\lambda_\varepsilon^2\, e^{2\mathfrak{m}\,s_v}}{4\pi(s_v+\varepsilon^2)}\, \mathrm{d}s_{\mathcal{I}(\tau)}\,,$$

arguing along the same lines as in Lemma 3.2.22. Together with Lemma 3.3.4, this concludes the proof. □

Counting contributing contractions

In Lemma 3.3.6, we saw that the limit of a contributing contracted tree $[\tau]_\kappa$ is independent of the chosen contraction in $C(\tau)$. Hence, in order to conclude the limit of X_ε^τ, it is only left to determine the size of $C(\tau)$.

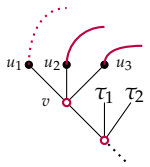
Lemma 3.3.7. *Let $\tau \in \mathcal{T}_3$, then $|C(\tau)| = 3^{i(\tau)}$.*

In order to prove Lemma 3.3.7, we first need the following result.

Lemma 3.3.8. *Let $\tau \in \mathcal{T}_3$ and $\kappa \in C(\tau)$, then every trident in τ_κ has an internal contraction. More precisely, for every $v \in \mathcal{V}_{\Psi}(\tau)$ there exists an edge $\{u, u'\} \in \kappa$ such that $v = \mathfrak{p}(u) = \mathfrak{p}(u')$. The set $\mathcal{V}_{\Psi}(\tau)$ was introduced below the statement of Lemma 3.2.12.*

Proof. Let $\tau \in \mathcal{T}_3$, $\kappa \in C(\tau)$ and consider $[\tau, \tau]_\gamma$ for the unique $\gamma \in \mathcal{Y}(\tau_\kappa, \tau_\kappa)$, see Lemma 3.3.2(ii). For any $v \in \mathcal{V}_{\Psi}(\tau)$, we write $u_1(v), u_2(v), u_3(v) \in \mathcal{L}(\tau)$ for the three leaves it is connected to (indexing them with 1 to 3 from left to right).

Now, assume there exists a $v \in \mathcal{V}_{\Psi}(\tau)$ without an internal contraction, i.e. $\{u_i(v), u_j(v)\} \notin \kappa$ for $1 \leqslant i, j \leqslant 3$:



, (3.70)

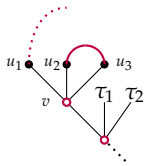
for some $\tau_1, \tau_2 \in \mathcal{T}_3$. Note that possibly one of the leaves could be uncontracted, which is indicated in the example above by the dotted purple line. It is an immediate consequence that $\Pi_\tau(\gamma) \neq \mathrm{Id}$, since otherwise $\{u_i(v), u_j(v)\} \in \kappa \subset \gamma$ for some $1 \leqslant i, j \leqslant 3$. However, this implies that

$$\lim_{\varepsilon \to 0} (\log \tfrac{1}{\varepsilon}) \cdot [\tau, \tau]_{\gamma,\varepsilon}(t, x) = 0\,, \quad \forall (t, x) \in (0, \infty) \times \mathbf{R}^2\,,$$

by Lemma 3.3.3, thus, contradicting the assumption $\kappa \in C(\tau)$. This finishes the proof. □

Proof of Lemma 3.3.7. We prove the statement by induction over the number of inner nodes $i(\tau)$, starting with $i(\tau) = 0$, i.e. $\tau = \bullet$. In this case, $|C(\tau)| = |\mathcal{K}(\tau)| = |\mathcal{Y}(\tau, \tau)| = 1$ and the claim holds. Now assume the statement holds true for any $\widehat{\tau} \in \mathcal{T}_3$ satisfying $i(\widehat{\tau}) \leqslant m$.

Let $\tau \in \mathcal{T}_3$ with $i(\tau) = m+1$ and $v \in \mathcal{V}_{\curlyvee}(\tau)$. We denote its neighbouring leaves by $u_1(v), u_2(v), u_3(v) \in \mathcal{L}(\tau)$ (indexing them with 1 to 3 from left to right):



, (3.71)

for some $\tau_1, \tau_2 \in \mathcal{T}_3$. Again note that possibly one of the leaves could be uncontracted. Using Lemma 3.3.8, we can partition $C(\tau)$ into three sets $C_{1,2}(\tau), C_{1,3}(\tau), C_{2,3}(\tau)$ with

$$C_{i,j}(\tau) := \big\{\kappa \in C(\tau) \,:\, \{u_i(v), u_j(v)\} \in \kappa\big\} \,.$$

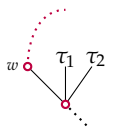
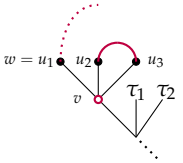
For any contraction $\kappa \in C(\tau)$ we define the tree resulting from τ_κ after removing the 1-cycle $\mathcal{C}$ with $\mathcal{I}_\mathcal{C} = \{v\}$:

$$\widehat{\tau}_{\tilde{\kappa}} := (\tau \setminus \mathcal{C})_{\tilde{\kappa}} \,, \tag{3.72}$$

using the cycle removal from Definition 3.2.14. Because v and thus $\widehat{\tau}$ (up to labelling of one leaf via $u_i(v)$) do not depend on the choice of $\kappa \in C(\tau)$, (3.72) defines a map $\mathfrak{K}_v : C(\tau) \to \mathcal{K}(\widehat{\tau})$ with $\mathfrak{K}_v(\kappa) := \tilde{\kappa}$. Moreover, we have $\tilde{\kappa} = \mathfrak{K}_v(\kappa) \in C(\hat{\tau})$ since

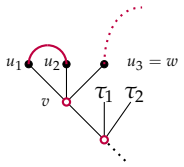
$$\begin{aligned} 0 &< \limsup_{\varepsilon \to 0} \, \big(\log \tfrac{1}{\varepsilon}\big) \cdot \mathbb{E}\big[|[\tau]_{\kappa,\varepsilon}(t,x)|^2\big] \\ &\leqslant \left(\frac{\hat{\lambda}^2}{2\pi}\right)^2 \limsup_{\varepsilon \to 0} \, \big(\log \tfrac{1}{\varepsilon}\big) \cdot \mathbb{E}\big[|[\hat{\tau}]_{\tilde{\kappa},\varepsilon}(t,x)|^2\big] \,, \end{aligned}$$

following the same lines as in (3.68). In fact, for any choice $1 \leqslant i < j \leqslant 3$, the map $\mathfrak{K}_v|_{C_{i,j}(\tau)}$ maps onto $C(\widehat{\tau})$ and defines a bijection. To see this, consider an arbitrary contraction $\widehat{\kappa} \in C(\widehat{\tau})$ (with the labeling of $\widehat{\tau}$ induced by τ) and define $\kappa := \widehat{\kappa} \cup \{u_i(v), u_j(v)\}$. For example we have the following reconstruction of a contraction in $C_{2,3}(\tau)$ using the inverse $(\mathfrak{K}_v|_{C_{2,3}(\tau)})^{-1}$:

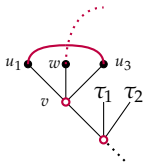
$[\widehat{\tau}]_{\widehat{\kappa}} =$  $\mapsto$  $= [\tau]_\kappa \,.$ (3.73)

On the other hand, for the same $\widehat{\kappa}$ we can also reconstruct the following two

contractions in $C_{1,2}(\tau)$ and $C_{1,3}(\tau)$, respectively:



and



.

In particular, for each set $C_{i,j}(\tau)$ there exists a unique $\kappa \in C_{i,j}(\tau)$ such that $\mathfrak{K}_v(\kappa) = \widehat{\kappa}$. As a consequence all three sets $C_{i,j}(\tau)$ have the same cardinality, which agrees with $|C(\widehat{\tau})|$. Lastly, applying the induction hypothesis to $|C(\widehat{\tau})|$, yields

$$|C(\tau)| = |C_{1,2}(\tau)| + |C_{1,3}(\tau)| + |C_{2,3}(\tau)| = 3|C(\widehat{\tau})| = 3^{m+1}\,.$$

This concludes the proof. □

3.3.2 Non-contributing trees

Up to now, we have identified contributing pairings (and contractions) to be the ones that lie in the pre-image $\Pi_\tau^{-1}(\mathrm{Id})$, when considering a fixed tree $\tau \in \mathcal{T}_3$. Moreover, we determined their exact contribution. Now, it is only left to control the overall contribution of the remaining contractions, which we will prove to have no contribution. We summarise this sections main finding in the following lemma:

Lemma 3.3.9. *Let $T > 0$, then uniformly over any $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$, $\tau \in \mathcal{T}_3^{N_\varepsilon, \circ}$, for $N_\varepsilon = \lfloor \log \frac{1}{\varepsilon} \rfloor$, $x \in \mathbf{R}^2$ and uniformly for all $t \in [0, T]$, we have*

$$\Big\| \sum_{\kappa \notin C(\tau)} \sqrt{\log \tfrac{1}{\varepsilon}}\, [\tau]_{\kappa,\varepsilon}(t, x) \Big\|_{L^2(\mathbb{P})} \leqslant \frac{1}{\sqrt{4 \log \frac{1}{\varepsilon}}} \frac{1}{\tau!} \left(\frac{6 e^{2+2\pi} \hat{\lambda}^2 e^{2\overline{\mathfrak{m}} t}}{\pi} \right)^{|\tau|} \frac{\hat{\lambda} e^{\mathfrak{m} t}}{\sqrt{4(t + \varepsilon^2)}},$$

where τ denotes the trimmed tree $\mathscr{T}(\tau)$ as in (2.82). In particular, for a fixed $\tau \in \mathcal{T}_3$ and $t > 0$, the right-hand side vanishes in the small-ε limit.

For the proof of Lemma 3.3.9 we need the following two lemmas. The first is an upper bound of a "symmetrised" integral over v-cycles.

Lemma 3.3.10. *Let $T > 0$. Then uniformly over any $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$, $t \in [0, T]$ and*

$\tau \in \mathcal{T}_3^{N_\varepsilon,\circ}$, $N_\varepsilon = \lfloor \log \frac{1}{\varepsilon} \rfloor$, *we have*

$$\sum_{\pi \in S_{2i(\tau)} \setminus \{\mathrm{Id}\}} \Big(\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} \Big) \leqslant \frac{1}{\log \frac{1}{\varepsilon}} \frac{1}{(\tau!)^2} \frac{(\hat{\lambda} e^{\overline{\mathfrak{m}} t})^{4i(\tau)}}{4\pi^{2i(\tau)-1}} e^{2(2+2\pi)\, i(\tau)} ,$$

where $\mathcal{I} := \mathcal{I}([\tau,\tau]) \setminus \mathfrak{o}$, *and* $\tau = \mathscr{T}(\tau)$ *denotes the trimmed tree* (2.82). *The function* $\Psi_{\pi,\varepsilon}$ *was defined in* (3.60).

The next result guarantees that the number of pairings γ of a tree $[\tau,\tau]$, which correspond to a permutation $\pi \in S_{2i(\tau)}$, grows at most exponentially in the number of inner nodes of a tree.

Lemma 3.3.11. *Let* $\tau \in \mathcal{T}_3$ *and* $\pi \in S_{2i(\tau)}$, *then* $\big|\Pi_\tau^{-1}(\pi)\big| \leqslant 6^{2i(\tau)}$.

Having both Lemma 3.3.10 and 3.3.11 at hand, we can now prove Lemma 3.3.9. The proofs of Lemma 3.3.10 and 3.3.11 are instead deferred to the end of this section.

Proof of Lemma 3.3.9. Consider $T > 0$, $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and $(t,x) \in [0,T] \times \mathbf{R}^2$. By representing second moments of contracted trees in terms of paired trees, cf. (3.42), we have

$$\begin{aligned} \Big\| \sum_{\kappa \notin C(\tau)} \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) \Big\|_{L^2(\mathbb{P})}^2 &= \log \tfrac{1}{\varepsilon} \sum_{\kappa,\kappa' \notin C(\tau)} \sum_{\gamma \in \mathcal{Y}(\tau_\kappa, \tau_{\kappa'})} [\tau,\tau]_{\gamma,\varepsilon}(t,x) \\ &\leqslant \log \tfrac{1}{\varepsilon} \sum_{\pi \in S_{2i(\tau)} \setminus \{\mathrm{Id}\}} \sum_{\gamma \in \Pi_\tau^{-1}(\pi)} [\tau,\tau]_{\gamma,\varepsilon}(t,x) , \end{aligned}$$

where we used additionally Lemma 3.3.3 to identify non-contributing pairings as precisely the ones that do not map onto Id under Π_τ, and the fact that

$$\exists \gamma \in \mathcal{Y}(\tau_\kappa, \tau_{\kappa'}) \text{ s.t. } \limsup_{\varepsilon \to 0} \big(\log \tfrac{1}{\varepsilon}\big) \, [\tau,\tau]_{\gamma,\varepsilon}(t,x) > 0 \quad \Leftrightarrow \quad \kappa, \kappa' \in C(\tau) ,$$

which is a consequence of the Cauchy-Schwartz inequality, cf. (3.69), and Lemma 3.3.2. Thus, Lemmas 3.2.21 and 3.3.11 imply

$$\begin{aligned} &\Big\| \sum_{\kappa \notin C(\tau)} \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) \Big\|_{L^2(\mathbb{P})}^2 \\ &\leqslant \big(\log \tfrac{1}{\varepsilon}\big) \, \lambda_\varepsilon^2 \, 6^{2i(\tau)} \sum_{\pi \in S_{2i(\tau)} \setminus \{\mathrm{Id}\}} \Big(\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}}) \, \mathrm{d}s_{\mathcal{I}} \Big) \, e^{2\mathfrak{m} t} p_{2(t+\varepsilon^2)}(0) . \end{aligned}$$

By application of Lemma 3.3.10, this can be further upper bounded by

$$\lambda_\varepsilon^2\, 6^{2i(\tau)}\, \frac{1}{(\tau!)^2}\, \frac{(\hat\lambda e^{\overline{\mathfrak{m}}\, t})^{4i(\tau)}}{4\pi^{2i(\tau)-1}} e^{2(2+2\pi)\, i(\tau)}\, e^{2\mathfrak{m}\, t} p_{2(t+\varepsilon^2)}(0)$$
$$= \frac{1}{4\, \log \frac{1}{\varepsilon}}\, \frac{1}{(\tau!)^2} \left(\frac{6e^{2+2\pi}\hat\lambda^2 e^{2\overline{\mathfrak{m}}\, t}}{\pi} \right)^{2i(\tau)}\, \frac{\hat\lambda^2 e^{2\mathfrak{m}\, t}}{4(t+\varepsilon^2)}\,.$$

Recalling the fact that $i(\tau) = |\tau|$, the result follows by taking the square root. □

Now we pass to the proof of Lemma 3.3.10, which is a quantitative version of Lemma 3.3.3. Instead of extending the integration domain from $\mathfrak{D}_{[\tau,\tau]}(t)$ to the box $[0,t]^{\mathcal{I}([\tau,\tau])\setminus \mathfrak{o}}$, as it was done in the proof of the latter lemma, we will make use of the fact that summation over all permutations in $\pi \in S_n \setminus \{\mathrm{Id}\}$ has a symmetrising effect that allows for a more precise control of the integral.

Lemma 3.3.12. *Fix $\varepsilon \in (0, \frac{1}{2})$ and $n \in \mathbf{N}$. Consider for any $\pi \in S_n$ the function $\Psi_{\pi,\varepsilon}$ introduced in (3.60). Then the function $\overline{\Psi}_\varepsilon : [0,\infty)^n \to \mathbf{R}$ defined by*

$$\overline{\Psi}_\varepsilon(s_1, \dots s_n) := \sum_{\pi \in S_n \setminus \{\mathrm{Id}\}} \Psi_{\pi,\varepsilon}(s_1, \dots, s_n)\,,$$

is symmetric in the variables $s_1, \cdots, s_n$.

Proof. It suffices to consider the function $\overline{\varphi}_\varepsilon$ given by

$$(s_1, ..., s_n) \mapsto \sum_{\pi \in S_n} \prod_{i=1}^n \frac{1}{s_i + s_{\pi(i)} + 2\varepsilon^2}\,,$$

since the term corresponding to the identity partition is symmetric itself. Now, for $\sigma \in S_n$, if we indicate $s_\sigma = (s_{\sigma(i)})_{i=1}^n$, we have

$$\overline{\varphi}_\varepsilon(s_\sigma) = \sum_{\pi \in S_n} \prod_{i=1}^n \frac{1}{s_{\sigma(i)} + s_{\sigma(\pi(i))} + 2\varepsilon^2} = \sum_{\pi \in S_n} \prod_{i=1}^n \frac{1}{s_i + s_{\sigma\pi\sigma^{-1}(i)} + 2\varepsilon^2}$$
$$= \sum_{\pi \in S_n} \prod_{i=1}^n \frac{1}{s_i + s_{\pi(i)} + 2\varepsilon^2} = \overline{\varphi}_\varepsilon(s)\,,$$

by performing an index change and using the fact that $\{\sigma\pi\sigma^{-1} : \pi \in S_n\} = S_n$. This concludes the proof. □

Having additionally the symmetrising effect of summing over all permutations (apart from the identity), we can prove the desired upper bound.

Proof of Lemma 3.3.10. Fix $T > 0$, $t \in [0,T]$ and define $n := 2i(\tau)$. Using Lemma 3.3.5 and the definition of $\overline{\Psi}_\varepsilon$ in Lemma 3.3.12, we find that

$$\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \overline{\Psi}_\varepsilon(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}} = \frac{1}{(\tau!)^2} \sum_{\pi \in S_n \setminus \{\mathrm{Id}\}} \int_{[0,t]^{\mathcal{I}}} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}}\,, \tag{3.74}$$

since $\overline{\Psi}_\varepsilon$ is symmetric by Lemma 3.3.12. Recall from (3.60) that $\Psi_{\pi,\varepsilon}$ is a product over cycles in the permutation π, which allows us to factorise the integral

$$\int_{[0,t]^{\mathcal{I}}} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}} = \prod_{i=1}^{K(\pi)} \int_{[0,t]^{\widehat{\mathcal{C}}_i}} \; \varepsilon^{\otimes|\widehat{\mathcal{C}}_i|}(s_v; v \in \widehat{\mathcal{C}}_i)\,\mathrm{d}s_{\widehat{\mathcal{C}}_i}\,,$$

with $K(\pi)$ denoting the number of cycles in the permutation π. Applying Lemma 3.2.16 to each term in the product yields

$$\begin{aligned}
\int_{[0,t]^{\mathcal{I}}} \Psi_{\pi,\varepsilon}(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}} &\leqslant \prod_{i=1}^{K(\pi)} \left(\frac{(\lambda_\varepsilon e^{\overline{\mathrm{m}}\, t})^{2|\widehat{\mathcal{C}}_i|}}{2^{|\widehat{\mathcal{C}}_i|}\,\pi} \log\left(1 + \tfrac{t}{\varepsilon^2}\right) \right) \\
&= \frac{(\lambda_\varepsilon e^{\overline{\mathrm{m}}\, t})^{2n}}{2^n} \left(\frac{1}{\pi} \log\left(1 + \tfrac{t}{\varepsilon^2}\right) \right)^{K(\pi)} \\
&\leqslant \frac{(\lambda_\varepsilon e^{\overline{\mathrm{m}}\, t})^{2n}}{2^n} M_\varepsilon(t)^{K(\pi)}\,,
\end{aligned}$$

where we introduced

$$M_\varepsilon(t) := \left\lceil \frac{1}{\pi} \log\left(1 + \tfrac{t}{\varepsilon^2}\right) \right\rceil .$$

Therefore, we obtain the following upper bound to (3.74):

$$\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \overline{\Psi}_\varepsilon(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}} \leqslant \frac{n!}{(\tau!)^2} \frac{(\lambda_\varepsilon e^{\overline{\mathrm{m}}\, t})^{2n}}{2^n} \left(\mathrm{E}_{S_n}\left[M_\varepsilon(t)^{K(\pi)} \right] - \frac{M_\varepsilon(t)^n}{n!} \right) . \tag{3.75}$$

Here we used the identity

$$\mathrm{E}_{S_n}\left[M_\varepsilon(t)^{K(\pi)} \right] = \frac{M_\varepsilon(t)^n}{n!} + \sum_{\pi \in S_n \setminus \{\mathrm{Id}\}} \frac{M_\varepsilon(t)^{K(\pi)}}{n!}\,,$$

with the expectation taken with respect to the uniform distribution on S_n, which has probability mass function $\frac{1}{n!}$. Hence, we have reduced the problem to studying the generating function of a discrete random variable, namely the total number of cycles in a uniformly at random chosen permutation. Its

distribution is a well studied object and we have the explicit identity

$$\mathrm{E}_{S_n}\Big[M_\varepsilon(t)^{K(\pi)}\Big] = \binom{n+M_\varepsilon(t)-1}{n}$$

at hand, see e.g. [Uel17, Display (5.14)]. Thus, we can rewrite (3.75) as

$$\int_{\mathfrak{D}_{[\tau,\mathcal{T}]}(t)} \overline{\Psi}_\varepsilon(s_\mathcal{I})\,\mathrm{d}s_\mathcal{I} \leqslant \frac{1}{(\tau!)^2}\frac{(\hat{\lambda}e^{\overline{\mathfrak{m}}\,t})^{2n}}{(2\log\frac{1}{\varepsilon})^n}\left(\frac{(n+M_\varepsilon(t)-1)!}{(M_\varepsilon(t)-1)!} - M_\varepsilon(t)^n\right). \tag{3.76}$$

Now, we expand the difference on the right-hand side using a telescopic sum which leads to

$$\frac{(n+M_\varepsilon(t)-1)!}{(M_\varepsilon-1)!} - M_\varepsilon(t)^n = \sum_{j=1}^{n-1} j\,M_\varepsilon(t)^j \prod_{k=j+1}^{n-1}(M_\varepsilon(t)+k)\,. \tag{3.77}$$

Note that for every $k = 0,\dots,n-1$

$$\begin{aligned}\frac{M_\varepsilon(t)+k}{2\log\frac{1}{\varepsilon}} &\leqslant \frac{\frac{1}{\pi}\log\left(1+\frac{t}{\varepsilon^2}\right)+k+1}{2\log\frac{1}{\varepsilon}} \leqslant \frac{1}{\pi}\Big(1+\frac{|\log(t+\varepsilon^2)|+\pi(k+1)}{2\log\frac{1}{\varepsilon}}\Big)\\ &\leqslant \frac{1}{\pi}\exp\left(\frac{|\log(t+\varepsilon^2)|+\pi(k+1)}{2\log\frac{1}{\varepsilon}}\right),\end{aligned}$$

thus, together with (3.76) and (3.77)

$$\begin{aligned}&\int_{\mathfrak{D}_{[\tau,\mathcal{T}]}(t)} \overline{\Psi}_\varepsilon(s_\mathcal{I})\,\mathrm{d}s_\mathcal{I} \qquad (3.78)\\ &\leqslant \frac{1}{(\tau!)^2}\frac{(\hat{\lambda}e^{\overline{\mathfrak{m}}\,t})^{2n}}{2\pi^{n-1}\log\frac{1}{\varepsilon}}\sum_{j=1}^{n-1} j\exp\left((n-1)\frac{|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}} + \pi\frac{j+\sum_{k=j+1}^{n-1}(k+1)}{2\log\frac{1}{\varepsilon}}\right)\\ &\leqslant \frac{1}{\log\frac{1}{\varepsilon}}\frac{1}{(\tau!)^2}\frac{(\hat{\lambda}e^{\overline{\mathfrak{m}}\,t})^{2n}}{2\,\pi^{n-1}}\sum_{j=1}^{n-1} j\exp\left((n-1)\frac{|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}} + \pi\sum_{k=j}^{n-1}\frac{k+1}{2\log\frac{1}{\varepsilon}}\right).\end{aligned}$$

The terms in the exponent can be estimated uniformly over ε and $t \in [0,T]$. For the first term, recall (3.33). Moreover, since $n \leq 2N_\varepsilon = 2\lfloor\log\frac{1}{\varepsilon}\rfloor$,

$$\sum_{k=j}^{n-1}\frac{k+1}{2\log\frac{1}{\varepsilon}} \leqslant \sum_{k=1}^{n}\frac{k}{2\log\frac{1}{\varepsilon}} = \frac{n(n+1)}{4\log\frac{1}{\varepsilon}} \leqslant n\,.$$

Combining the two estimates with (3.78) yields

$$\begin{aligned}\int_{\mathfrak{D}_{[\tau,\tau]}(t)} \overline{\Psi}_\varepsilon(s_{\mathcal{I}})\,\mathrm{d}s_{\mathcal{I}} &\leqslant \frac{1}{\log\frac{1}{\varepsilon}}\frac{1}{(\tau!)^2}\frac{(\hat{\lambda}e^{\overline{\mathrm{m}}\,t})^{2n}}{2\,\pi^{n-1}}\frac{n(n-1)}{2}e^{(2+\pi)n}\\ &\leqslant \frac{1}{\log\frac{1}{\varepsilon}}\frac{1}{(\tau!)^2}\frac{(\hat{\lambda}e^{\overline{\mathrm{m}}\,t})^{2n}}{4\,\pi^{n-1}}e^{(2+2\pi)n},\end{aligned}$$

where we used $n(n-1) \leqslant e^{2n} \leqslant e^{\pi n}$ in the last step. This concludes the proof. □

Proof of Lemma 3.3.11. For convenience let us write $n = 2i(\tau)$. Our aim is to obtain an upper bound on the total number of parings γ that give rise to a given permutation $\pi \in S_n$, via the map $\Pi_\tau(\gamma)$ from Definition 3.2.19. It will be convenient to represent π as a product of permutation cycles $\pi = \prod_{i=1}^{K(\pi)} \widehat{\mathcal{C}}_i$, for some $K(\pi) \in \{1, \cdots, n\}$. Here we slightly abuse the notation $\widehat{\mathcal{C}}_i$, which is already used in Definition 3.2.19. Indeed, while the decomposition of a permutation into a product of cycles is unique, the order in which these cycles are chosen is arbitrary. This is not the case in Definition 3.2.19, as for example the cycle $\widehat{\mathcal{C}}_1$ is necessarily a cycle among bases of tridents or cherries (because it is the first cycle we extract from a paired tree). In our setting, since we start from an arbitrary permutation π, we assume nothing further on $\widehat{\mathcal{C}}_i$ other than that they are cycles that decompose π. Then we provide the desired upper bound via the steps that follow.

(i) By Lemma 3.2.12 we know that any pairing γ in $\Pi_\tau^{-1}(\pi)$ must contain a v-cycle that alternates between leaves and inner nodes of $[\tau, \tau]$. This is because the first v-cycle we extract must be a v-cycle in the paired tree $[\tau, \tau]_\gamma$ (later ones belong instead to trees that are derived from $[\tau, \tau]_\gamma$ by extracting v-cycles). Therefore, there must exist a cycle $\widehat{\mathcal{C}}_{i_1}$ that only runs through inner vertices in $V_{\curlyvee} \cup \mathcal{V}_{\curlyvee\!\cdot\!\cdot} =: W^{(1)}$, with $V_{\curlyvee}$ and $\mathcal{V}_{\curlyvee\!\cdot\!\cdot}$ denoting the sets of *cherries* and *tridents* in $[\tau, \tau]$, respectively, recall the definition from below Lemma 3.2.12.

Indeed, if no such cycle exists, then the given permutation cannot arise from any paired tree $[\tau, \tau]_\gamma$ using the extraction algorithm Π_τ and the pre-image is the empty set, so that our upper bound holds true. See the discussion below Example 3.2.20 for an example of this kind.

(ii) Since $\widehat{\mathcal{C}}_{i_1} \subset V_{\curlyvee} \cup \mathcal{V}_{\curlyvee\!\cdot\!\cdot}$, to construct a v-cycle corresponding to $\widehat{\mathcal{C}}_{i_1}$ we must choose for every vertex v with label in $\widehat{\mathcal{C}}_{i_1}$ an (outgoing) leaf that

connects to the next vertex of the v-cycle and an (incoming) leaf that connects to the previous vertex of the v-cycle. Here, which leaf is outgoing and which leaf is incoming matters, leading to at most $2\binom{3}{2} = 6$ choices for every vertex (for nodes in $\mathcal{V}_{\Psi}$ we have six choices, for nodes in $V_{\curlyvee}$ only two). Thus, there are at most $6^{|\widehat{\mathcal{C}}_{i_1}|}$ ways to construct a v-cycle through inner nodes labelled by $\widehat{\mathcal{C}}_{i_1}$.

(iii) Now proceed iteratively. For $j > 1$, we define the set $W^{(j)} \subset \mathcal{I}([\tau, \tau]) \setminus \mathfrak{o}$ as follows: An inner node v lies in $W^{(j)}$, if and only if

- there exist at least two distinct paths from v to leaves in $\mathcal{L}([\tau, \tau])$, which only run through descendants of v (away from the root), that have been previously extracted, i.e. they lie in $\bigcup_{k=1}^{j-1} \mathcal{I}_{\widehat{\mathcal{C}}_{i_k}}$,
- and v has not been extracted previously, i.e. $v \notin \bigcup_{k=1}^{j-1} \mathcal{I}_{\widehat{\mathcal{C}}_{i_k}}$.

The set $W^{(j)}$ describes the nodes that became "admissible" after extracting the cycles $\{\widehat{\mathcal{C}}_i\}_{i \in \{i_1, \cdots, i_{j-1}\}}$, meaning that the inner nodes of the next cycle that we extract must belong to $W^{(j)}$.

(iv) Proceed by choosing a cycle $\widehat{\mathcal{C}}_{i_j}$ with nodes in $W^{(j)}$ and counting all possible choices to create a v-cycle with the corresponding nodes as inner nodes.

(v) We are done once all cycles have been removed, or we cannot find a cycle that runs through vertices in $W^{(j)}$. In the latter case, we again found a permutation that cannot be obtained as image of the map Π_τ (so the pre-image is empty and the bound trivially true).

In this way, we count all possible pairings that lead to π. Overall, either the pre-image is empty and the stated bound is trivially true, or it is bounded by

$$6^{\sum_{i=1}^{K(\pi)} |\widehat{\mathcal{C}}_i|} = 6^n \,,$$

which is the desired bound and concludes the proof.

□

3.3.3 Proof of Proposition 3.1.2

Proof of Proposition 3.1.2. Fix any $T > 0$, $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$ and $\tau \in \mathcal{T}_3^{N_\varepsilon, \circ}$. Then by Lemma 3.2.1 and identity (3.39), we can write

$$\begin{aligned} X_\varepsilon^\tau &= \frac{h^{(\tau)}(1)}{s(\tau)} [\tau]_\varepsilon = \frac{h^{(\tau)}(1)}{s(\tau)} \sum_{\kappa \in \mathcal{K}(\tau)} [\tau]_{\kappa,\varepsilon} \\ &= \frac{h^{(\tau)}(1)}{s(\tau)} \left\{ \sum_{\kappa \in C(\tau)} [\tau]_{\kappa,\varepsilon} + \sum_{\kappa \notin C(\tau)} [\tau]_{\kappa,\varepsilon} \right\} . \end{aligned}$$

Thus, by the triangle inequality and the fact that $3^{|\tau|} = \sum_{\kappa \in C(\tau)} 1$, see Lemma 3.3.7, we have for every $(t, x) \in (0, T] \times \mathbf{R}^2$

$$\begin{aligned} &\left\| \sqrt{\log \tfrac{1}{\varepsilon}} \cdot X_\varepsilon^\tau(t,x) - \frac{h^{(\tau)}(1)}{\tau!\, s(\tau)} \left(\frac{3\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \hat{\lambda} e^{\mathfrak{m} t} P_{t+\varepsilon^2} \eta(x) \right\|_{L^2(\mathbb{P})} \\ &\leqslant \frac{|h^{(\tau)}(1)|}{s(\tau)} \sum_{\kappa \in C(\tau)} \left\| \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) - \frac{1}{\tau!} \left(\frac{\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \hat{\lambda} e^{\mathfrak{m} t} P_{t+\varepsilon^2} \eta(x) \right\|_{L^2(\mathbb{P})} \\ &\quad + \frac{|h^{(\tau)}(1)|}{s(\tau)} \left\| \sum_{\kappa \notin C(\tau)} \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) \right\|_{L^2(\mathbb{P})} . \end{aligned} \tag{3.79}$$

The second term on the right-hand side can be directly estimated as

$$\begin{aligned} &\left\| \sum_{\kappa \notin C(\tau)} \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) \right\|_{L^2(\mathbb{P})} \\ &\leqslant \frac{1}{\tau!} \frac{1}{\sqrt{4 \log \frac{1}{\varepsilon}}} \left(\frac{6 e^{2+2\pi} \hat{\lambda}^2 e^{2\overline{\mathfrak{m}} t}}{\pi} \right)^{|\tau|} \frac{\hat{\lambda} e^{\mathfrak{m} t}}{\sqrt{4(t+\varepsilon^2)}} , \end{aligned} \tag{3.80}$$

where we used Lemma 3.3.9. For the first term, we obtain

$$\begin{aligned} &\sum_{\kappa \in C(\tau)} \left\| \sqrt{\log \tfrac{1}{\varepsilon}} \cdot [\tau]_{\kappa,\varepsilon}(t,x) - \frac{1}{\tau!} \left(\frac{\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \hat{\lambda} e^{\mathfrak{m} t} P_{t+\varepsilon^2} \eta(x) \right\|_{L^2(\mathbb{P})} \\ &= \frac{1}{\tau!} \left(\frac{3\hat{\lambda}^2}{2\pi} \right)^{|\tau|} \left\| \left(\left(\frac{1}{2 \log \frac{1}{\varepsilon}} \int_0^t \frac{e^{2\mathfrak{m} s}}{s+\varepsilon^2} \, \mathrm{d}s \right)^{|\tau|} - 1 \right) \hat{\lambda} e^{\mathfrak{m} t} P_{t+\varepsilon^2} \eta(x) \right\|_{L^2(\mathbb{P})} \\ &\leqslant \frac{1}{\tau!} \left(\frac{9\hat{\lambda}^2 e^{2\overline{\mathfrak{m}} t + 1}}{2\pi} \right)^{|\tau|} \frac{e^{2|\mathfrak{m}| t} + |\log(t+\varepsilon^2)|}{2 \log \frac{1}{\varepsilon}} \frac{\hat{\lambda} e^{\mathfrak{m} t}}{\sqrt{4\pi(t+\varepsilon^2)}} , \end{aligned} \tag{3.81}$$

where in the first equality we applied Lemma 3.3.6 and Lemma 3.3.7, together with the identity $\bullet\!\!\circ_\varepsilon(t,x) = \hat\lambda(\log\frac{1}{\varepsilon})^{-\frac{1}{2}} e^{\mathfrak{m}\,t} P_{t+\varepsilon^2}\eta(x)$. In the last step, we moreover applied Corollary 3.3.14, proven below, which yields

$$\begin{aligned}\left|\left(\frac{1}{2\log\frac{1}{\varepsilon}}\int_0^t \frac{e^{2\mathfrak{m}\,s}}{s+\varepsilon^2}\,\mathrm{d}s\right)^{|\tau|} - 1\right| &\leqslant |\tau|\ \left(3e^{2\overline{\mathfrak{m}}\,t}\right)^{|\tau|-1} \frac{e^{2|\mathfrak{m}|\,t} + |\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}} \\ &\leqslant \left(3e^{2\overline{\mathfrak{m}}\,t+1}\right)^{|\tau|} \frac{e^{2|\mathfrak{m}|\,t} + |\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}}\,,\end{aligned}$$

where used additionally the bound $|\tau| \leqslant 3e^{|\tau|}$ in the last inequality. [3]

Now, combining (3.79), (3.80) and (3.81) yields

$$\begin{aligned}&\left\|\sqrt{\log\tfrac{1}{\varepsilon}}\cdot X_\varepsilon^\tau(t,x) - \frac{h^{(\tau)}(1)}{\tau!\,s(\tau)}\left(\frac{3\hat\lambda^2}{2\pi}\right)^{|\tau|}\hat\lambda e^{\mathfrak{m}\,t}P_{t+\varepsilon^2}\eta(x)\right\|_{L^2(\mathbb{P})} \\ &\leqslant \frac{|h^{(\tau)}(1)|}{\tau!\,s(\tau)}(\hat\lambda e^{\overline{\mathfrak{m}}\,t})^{2|\tau|}\left(\left(\frac{9\,e}{2\pi}\right)^{|\tau|}\frac{e^{2|\mathfrak{m}|\,t}+|\log(t+\varepsilon^2)|}{2\log\frac{1}{\varepsilon}}\frac{1}{\sqrt{\pi}}\right. \\ &\qquad\left. + \frac{1}{2\sqrt{\log\frac{1}{\varepsilon}}}\left(\frac{6e^{2+2\pi}}{\pi}\right)^{|\tau|}\right)\frac{\hat\lambda e^{\mathfrak{m}\,t}}{\sqrt{4(t+\varepsilon^2)}} \\ &\leqslant \frac{|h^{(\tau)}(1)|}{\tau!\,s(\tau)}\left(C\,\hat\lambda^2 e^{2\overline{\mathfrak{m}}\,t}\right)^{|\tau|}\frac{e^{2|\mathfrak{m}|\,t}+|\log(t+\varepsilon^2)|+\sqrt{\log\frac{1}{\varepsilon}}}{2\log\frac{1}{\varepsilon}}\frac{\hat\lambda e^{\mathfrak{m}\,t}}{\sqrt{4(t+\varepsilon^2)}}\,,\end{aligned}$$

with C being a new constant given by

$$C := \frac{6e^{2+2\pi}}{\pi} \geqslant \left(\frac{9\,e}{2\pi}\right) \vee \left(\frac{6e^{2+2\pi}}{\pi}\right)\,. \tag{3.82}$$

This concludes the proof. □

Remark 3.3.13. *The constant C (3.82) restricts the range of $\hat\lambda$ in Theorem 3.0.1, however, we stress that it is not optimal. We refrained from giving a sharper estimate, since the choice $\hat\lambda$ is restricted by the nature of our approach (radius of convergence of the power series). Because we do not expect a phase transition in $\hat\lambda$, we chose a tidier proof over optimising the constant.*

Let us mention that a smaller C can be obtained by sharpening the estimates below (3.76) and in (3.33). In particular, for the limiting result to hold, we don't re-

[3] The estimates in the display above vastly overestimate the quantity on the left-hand side. However, we will see below that the main restriction will come in fact from the term (3.80). Thus, the crude upper bounds here do not play a role.

quire estimates uniformly in $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$*, but can restrict ourselves to an arbitrary small interval* $(0, \delta)$.

We finish by providing the last ingredient of the proof above, which is a consequence of Lemma 3.2.7 and Taylor's approximation theorem.

Corollary 3.3.14. *Let* $\mathfrak{m} \in \mathbf{R}$, $T > 0$ *and* $k \in \mathbf{N}$, *then*

$$\left| \left(\frac{1}{2 \log \frac{1}{\varepsilon}} \int_0^t \frac{e^{2\mathfrak{m} s}}{s + \varepsilon^2} \, \mathrm{d}s \right)^k - 1 \right| \leqslant k \, \left(3 e^{2\overline{\mathfrak{m}} t} \right)^{k-1} \frac{e^{2|\mathfrak{m}| t} + |\log (t + \varepsilon^2)|}{2 \log \frac{1}{\varepsilon}} ,$$

for every $t \in [0, T]$ *and* $\varepsilon \in (0, \frac{1}{T} \wedge \frac{1}{2})$.

Proof. Using a first order Taylor expansion of the monomial of order k around 1, we have

$$\begin{aligned} &\left| \left(\frac{1}{2 \log \frac{1}{\varepsilon}} \int_0^t \frac{e^{2\mathfrak{m} s}}{s + \varepsilon^2} \, \mathrm{d}s \right)^k - 1 \right| \\ &\leqslant k \left\{ \sup_{x \in [1, b_\varepsilon(t) \vee 1]} x^{k-1} \right\} \left| \frac{1}{2 \log \frac{1}{\varepsilon}} \int_0^t \frac{e^{2\mathfrak{m} s}}{s + \varepsilon^2} \, \mathrm{d}s - 1 \right| , \end{aligned} \tag{3.83}$$

with

$$b_\varepsilon(t) := \frac{1}{2 \log \frac{1}{\varepsilon}} \int_0^t \frac{e^{2\mathfrak{m} s}}{s + \varepsilon^2} \, \mathrm{d}s \leqslant e^{2\overline{\mathfrak{m}} t} \left(1 + \frac{|\log (t + \varepsilon^2)|}{2 \log \frac{1}{\varepsilon}} \right) .$$

In particular, we find that

$$\sup_{x \in [1, b_\varepsilon(t) \vee 1]} x^{k-1} \leqslant e^{2(k-1)\overline{\mathfrak{m}} t} \left(1 + \frac{|\log (t + \varepsilon^2)|}{2 \log \frac{1}{\varepsilon}} \right)^{k-1} \leqslant e^{2(k-1)\overline{\mathfrak{m}} t} 3^{k-1} ,$$

where we used (3.33) in the last step. The statement follows now by upper bounding the last term in (3.83) via Lemma 3.2.7. □

Chapter 4

Central limit theorems for the 2D random directed polymer

The directed random polymer model is given by an exponentially tilt of the simple random walk measure $\mathbf{P}_N$. We restate the definition from the introduction: For a fixed realisation of the disorder ω, the directed polymer measure of length N and disorder strength $\beta \geq 0$ is defined using the following change of measure

$$\mathbf{P}^{\omega}_{\beta,N}(\,\mathrm{d}S) := \frac{1}{Z_{\beta,N}(0,0,N,\star)} \exp\left(\sum_{n=1}^{N}(\beta\,\omega_{n,S_n} - \lambda(\beta))\right)\mathbf{P}_N(\,\mathrm{d}S),$$

where $\lambda(\beta)$ is a positive constant, which we will fix below. The (point-to-plane) partition function is given by

$$Z_{\beta,N}(0,0,N,\star) := \mathbf{E}_N\left[\exp\left(\sum_{n=1}^{N}(\beta\,\omega_{n,S_n} - \lambda(\beta))\right)\right],$$

making $\mathbf{P}^{\omega}_{\beta,N}$ a probability measure. Here, $\star$ denotes the free boundary condition of the endpoint S_N to take arbitrary values in $\mathbf{Z}^d$. The notation for the partition function might seem overloaded at this point, but will become clear below.

The aim of this chapter, based on the article [Gab23], is to prove an invariance principle for the 2D-DRPM, in the full subcritical intermediate disorder regime. More precisely, the distribution of diffusively rescaled polymer paths converges in probability to the Wiener measure, when taking the weak disorder limit. Previosuly, analogous results were only obtained for $d \neq 2$.

Throughout this chapter we assume that $\omega = (\omega_{n,z})_{(n,z)\in\mathbb{N}\times\mathbb{Z}^d}$ is a collection of i.i.d. real random variables with law $\mathbb{P}$ (independent of $\mathbf{P}_N$), satisfying

$$\mathbb{E}[\omega_{n,z}] = 0, \qquad \mathbb{E}[\omega_{n,z}^2] = 1,$$
$$\lambda(\beta) := \log \mathbb{E}[e^{\beta\omega_{n,z}}] < \infty \quad \forall \beta > 0 \text{ small enough.}$$

For technical reasons, we also require a concentration inequality for the law $\mathbb{P}$. More precisely, we assume the existence of $\gamma > 1$ and $C_1, C_2 \in (0,\infty)$ such that for every $n \in \mathbb{N}$ and convex, 1-Lipschitz $f : \mathbf{R}^n \mapsto \mathbf{R}$ we have

$$\mathbb{P}(|f(\omega_1,\dots,\omega_n) - M_f| \geq t) \leq C_1 \exp\left(-\frac{t^\gamma}{C_2}\right), \tag{4.1}$$

where $(\omega_i)_{1\leq i\leq n}$ is a subset of the family of random variables introduced above and M_f is a median of $f(\omega_1,\dots,\omega_n)$. Condition (4.1) guarantees control on the negative tail of the environment and, for example, is satisfied whenever ω is bounded or Gaussian. See [Led01] for an even wider class of potential laws and more details. We also refer to Remark 4.2.5 for a discussion and possible approach to weaken this assumption.

Results and comparison to the literature

We will consider the DRPM in two space-dimensions and study the asymptotic behaviour of paths under the measure $\mathbf{P}^\omega_{\beta,N}$ in the large N limit. The random measure is supported on the space $\{(S_n)_n \in (\mathbb{Z}^2)^{N+1}\}$, more precisely, its support is given by the subset $\Omega_{0,N}$ of nearest neighbour paths starting at the origin. Naturally, the space $C[0,1] := C([0,1],\mathbf{R}^2)$ equipped with the supremum-norm is considered as common reference space for the paths. For this, we introduce the mapping $\overline{\pi} = \overline{\pi}_N : \Omega_{0,N} \to C[0,1]$ given by

$$X_t^{(N)} := \overline{\pi}(S)_t = \frac{1}{\sqrt{N}}\left(S_{\lfloor tN\rfloor} + (tN - \lfloor tN\rfloor)(S_{\lfloor tN\rfloor+1} - S_{\lfloor tN\rfloor})\right), \tag{4.2}$$

which embeds discrete nearest-neighbour paths $\Omega_{0,N}$ in the space of continuous functions, by linearly interpolating between integer points and rescaling space-time diffusively. Furthermore, we equip $C[0,1]$ with the Borel σ-algebra $\mathcal{F}$, implying measurability of the projection maps $\overline{\pi}$.

The chapter's main contribution is an invariance principle for rescaled polymer paths in the weak disorder limit, recall (1.9). This is the large N limit

when scaling $\beta = \beta_N \sim \widehat{\beta}\sqrt{\pi(\log N)^{-1}}$, with $\widehat{\beta} \in (0,1)$. To be more precise, we consider

$$\beta_N := \frac{\widehat{\beta}}{\sqrt{R_N}}, \quad \text{where} \quad R_N := \sum_{n=1}^{N} \sum_{z \in \mathbf{Z}^2} \mathbf{P}_N(S_n = z)^2 = \tfrac{\log N}{\pi} + O(1) \tag{4.3}$$

denotes the **replica overlap** of the simple random walk. The chapter's main result then states that the quenched polymer paths $\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}$ converge to the law of Brownian motion in $\mathbb{P}$-probability, in the full intermediate subcritical regime.

Theorem 4.0.1 ([Gab23]). *Let $\widehat{\beta} \in (0,1)$ and β_N be as in (4.3). Then*

$$\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N} \xrightarrow{d} \mathbf{P}\big(\tfrac{1}{\sqrt{2}}W \in \cdot\,\big), \qquad \text{as } N \to \infty, \quad \text{in } \mathbb{P}\text{-probability},$$

where $\mathbf{P}$ denotes the Wiener measure on $C[0,1]$. We wrote $\overline{\pi}_{\#}$ for the push-forward operation under $\overline{\pi}$, cf. (4.2).

Despite the random polymer converging to a stochastic process (which is independent of the disorder) on a large scales, the disorder influences the behaviour of the polymer on small scales. On the microscopic level, the disorder prevails and we can deduce a local limit theorem. This allows to compare the microscopic polymer transition probabilities to the ones of Brownian motion, weighted by random multiplicative factors which depend on the rescaled transition space-time points.

Proposition 4.0.2 ([Gab23]). *Let $\widehat{\beta} \in (0,1)$, β_N as in (4.3), $(z_j)_{j=1}^k = (z_j(N))_{j=1}^k \in \mathbf{Z}^2$ and $(m_j)_{j=1}^k = (m_j(N))_{j=1}^k \in \mathbb{N}$ such that the following limits exist*

$$\lim_{N\to\infty} \tfrac{z_j}{\sqrt{N}} =: x_j \quad \text{and} \quad \lim_{N\to\infty} \tfrac{m_j}{N} =: t_j \in (0,1) \text{ with } 0 < t_1 < \cdots < t_k < 1,$$

and $\mathbf{P}_N(S_{m_j(N)} = z_j(N)) > 0$. Then,

$$\begin{aligned} \Big(\frac{N}{2}\Big)^k \mathbf{P}^{\omega}_{\beta_N,N}(S_{m_1} = z_1, \dots, S_{m_k} = z_k) \\ \xrightarrow{d} \prod_{j=1}^{k} :e^{Y^-(t_j,x_j)}: :e^{Y^+(t_j,x_j)}: \prod_{j=1}^{k} p_{\frac{1}{2}(t_j - t_{j-1})}(x_j - x_{j-1}), \end{aligned} \tag{4.4}$$

where $Y_j^{\pm}$'s are i.i.d. centred Gaussians with variance $\log(1-\widehat{\beta}^2)^{-1}$. We used the shorthand notation $:e^Y: := e^{Y - \frac{1}{2}\mathbb{E}[Y^2]}$ for the Wick exponential.

The local limit theorem above reinforces the picture, that the considered intermediate disorder regime is indeed the region where effects of disorder start to emerge in a non-trivial way.

Remark 4.0.3. *We excluded the case $t_k = 1$ from Proposition 4.0.2, since it would only give rise to a single factor $:e^{Y^-(1,x)}:$. However, the proof can be repeated almost verbatim to include this case.*

Let us give a short overview on the literature of diffusivity of directed random polymers. For the two dimensional case with β_N scaled as in (4.3), it was proven in [CSZ17b] that the diffusively rescaled field

$$\left\{ Z_{\beta_N,N}(0, \lfloor \sqrt{N}x \rfloor, \lfloor tN \rfloor, \star) \, : \, t > 0, \, x \in \mathbf{R}^2 \right\}$$

converges to the solution of the stochastic heat equation with additive space-time white noise. However, to the author's best knowledge, there are no results on diffusivity of the polymer paths in this case. Subcritical scalings $\beta_N^2 \ll R_N^{-1}$ in $d = 2$ were considered by Feng [Fen12], who proved diffusivity of the polymer endpoint. However, under such rescaling the partition function's variance vanishes in the large N limit, which essentially brings us to the situation of setting $\widehat{\beta} = 0$. Theorem 4.0.1, on the other hand, considers a **critical scaling** under which the partition function converges to a non-trivial random variable and a transition (in $\widehat{\beta}$) of the polymer path behaviour is expected. Our result not only covers the diffusivity of the polymer endpoint, but fully determines the behaviour of the limiting polymer paths in the intermediate subcritical regime.

In dimension $d \geq 3$, diffusivity of the directed random polymer in the weak disorder regime was first proven to hold with probability one in [IS88], for sufficiently small disorder strength. The works [Bol89, Kif97] simplified and extended the original result further. First invariance principles were deduced in [AZ96, SZ96] where, for $\beta > 0$ small enough, almost sure convergence to the law of Brownian motion with dimension-dependent diffusion matrix was achieved. Later, it was extended to the full weak disorder regime by Comets and Yoshida [CY06] in the sense of a functional central limit theorem, which holds in probability. Theorem 4.0.1 can be viewed as the analogous result in $d = 2$.

Recently, Junk [Jun23, Jun21] gave an alternative proof of determining the limit of the polymer endpoint distribution (for bounded bond disorder) in $d \geq 3$, by introducing a comparison principle for partition functions of

distinct parameters β. This allows them to transform the polymer endpoint distribution into the one of the simple random walk with a multiplicative error, which converges to one.

For $d = 1$, it was shown that transition probabilities of the discrete polymer measure admit a random limit when space-time is scaled diffusively [AKQ14b]. Because every $\beta > 0$ lies in the strong disorder regime, they also relied on an intermediate disorder scaling: $\beta_N \sim \widehat{\beta} N^{-1/4}$. In [AKQ14a], the same authors constructed the corresponding continuum polymer measure. In contrast to $d \geq 3$, the limiting distribution of polymer paths is singular with respect to the Wiener measure, while maintaining the same basic properties as Brownian motion.

An analogous result was also shown for the continuum disordered pinning model in [CSZ14]. The advances, both for the pinning model and the $(1 + 1)$-dimensional directed polymer, then motivated to provide a general skeleton for the study of weak disorder scaling limits of discrete systems, see [CSZ17a].

Naturally, one would like to strengthen the invariance principle in Theorem 4.0.1 to $\mathbb{P}$-a.s. convergence, similar to the results in $d \geq 3$ (with $\widehat{\beta}$ small enough) [AZ96, SZ96]. However, these results exploit the fact that the sequence of partition functions $(Z_{\beta,N})_N$ forms a martingale. In the two-dimensional case, this property is lost due to the rescaling $\beta = \beta_N$, which is why we do not expect our methods to yield an almost sure invariance principle.

Lastly, let us mention that the statement of the local limit theorem (Proposition 4.0.2) is reminiscent of the construction in [AKQ14a], where the limiting field of partition functions was used to construct the continuum directed polymer in $(1 + 1)$-dimension. Because in $d = 1$ the random (macroscopic) field induced by the limiting partition functions is continuous in its time and space points, the constructed polymer measure is the correct limiting object. However, for $d = 2$ this is not the case anymore and leads to different behaviour of the polymer paths, both on a microscopic and macroscopic level. This (rough) structure of the partition function requires substantial work, in order to treat the limiting polymer measure.

Remark 4.0.4. *Theorem 4.0.1 should hold for a larger class of symmetric random walks which satisfy a local limit theorem in the sense of [CSZ17b, Hypothesis 2.4] and their replica overlap fulfills $R_N \to \infty$, as a slowly varying function.*

Remark 4.0.5. *Instead of studying discrete polymers, we could have similarly worked with polymers in the continuous space-time domain* $[0,1] \times \mathbf{R}^2$, *cf.* (1.16), *with disorder strength* $\beta_\varepsilon = \widehat{\beta}\sqrt{2\pi/\log \varepsilon^{-1}}$. *The corresponding version of Theorem 4.0.1 then reads: For every* $\widehat{\beta} \in (0,1)$

$$\mathbf{P}^{\xi}_{\beta_\varepsilon,\varepsilon} \xrightarrow{d} \mathbf{P}, \qquad \text{as } \varepsilon \to 0 \quad \text{in } \mathbb{P}\text{-probability}. \tag{4.5}$$

Since all properties of the DRPM's partition function also hold in the continuum, see [CSZ17b, CSZ20], the proof of the above fact should follow along the same lines.

4.1 Background and outline of the proof

Let us motivate the choice of the weak disorder scaling β_N. A second moment calculation of the partition function, see for example [IS88], provides the following heuristic: First, we write

$$\begin{aligned}\mathbb{E}[Z_{\beta,N}(0,0,N,\star)^2] &= \mathbf{E}_N^{\otimes 2}\Big[\prod_{n=1}^{N} e^{(\lambda(2\beta)-2\lambda(\beta))\mathbb{1}_{S_n=S'_n}}\Big] = \mathbf{E}_N^{\otimes 2}\Big[\prod_{n=1}^{N}(1+\sigma^2\mathbb{1}_{S_n=S'_n})\Big] \\ &= \sum_{k=0}^{N}\sigma^{2k}\sum_{1\le n_1<\cdots<n_k\le N}\mathbf{E}_N^{\otimes 2}\Big[\prod_{i=1}^{k}\mathbb{1}_{S_{n_i}=S'_{n_i}}\Big],\end{aligned} \tag{4.6}$$

where S and S' are two independent random walks of length N and σ is given by

$$\sigma = \sigma(\beta) := \sqrt{e^{\lambda(2\beta)-2\lambda(\beta)}-1}. \tag{4.7}$$

Note that $\lambda(2\beta) - 2\lambda(\beta) \sim \beta^2$ for small $\beta > 0$ and therefore $\lim_{\beta\to 0}\beta/\sigma(\beta) = 1$ [CSZ20, Equation (2.15)]. Next, we upper bound the sums in (4.6) by ignoring the ordering of n_i's, which yields

$$\mathbb{E}[Z_{\beta,N}(0,0,N,\star)^2] \le \sum_{k=0}^{N}\sigma^{2k}\Big(\sum_{n=1}^{N}\sum_{z\in\mathbf{Z}^2}\mathbf{P}_N(S_n=z)^2\Big)^k = \sum_{k=0}^{N}\big(\sigma^2 R_N\big)^k, \tag{4.8}$$

recalling the replica overlap R_N (4.3). Now, considering the fact that $\sigma(\beta) \sim \beta$ for small $\beta > 0$, (4.8) suggests that the correct rescaling is given by $\beta = \beta_N := \widehat{\beta}/\sqrt{R_N}$, whenever $\widehat{\beta} \in (0,1)$, such that (4.8) converges in the large N limit. Indeed, it was proven in [CSZ17b] that under β_N the partition function $Z_{\beta_N,N}(0,0,N,\star)$ converges to a non-trivial (random) limit, whenever $\widehat{\beta} \in (0,1)$. See also (4.14) below. Moreover, they noticed the

existence of a transition on the finer scale $\widehat{\beta}$, with $\widehat{\beta}_c = 1$ denoting the critical point where the L^2-norm of the partition function explodes.

Theorem 4.0.1 is not a straightforward consequence of the positivity of the limiting partition function, but requires precise estimates that quantify the correlation structure of the limiting field. The first step towards Theorem 4.0.1 is to show convergence of finite-dimensional distributions to the ones of Brownian motion.

Proposition 4.1.1. *Let $\widehat{\beta} \in (0,1)$ and β_N be as in (4.3), then for any $0 \leq t_1 < \ldots < t_k \leq 1$ we have*

$$\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}\big((X_{t_1},\ldots,X_{t_k}) \in \cdot\big) \xrightarrow{d} \mathbf{P}\big(\tfrac{1}{\sqrt{2}}(W_{t_1},\ldots,W_{t_k}) \in \cdot\big), \quad \text{in } \mathbb{P}\text{-probability},$$

where $\mathbf{P}$ denotes the Wiener measure on $C[0,1]$. We wrote $\overline{\pi}_{\#}$ for the push-forward operation under $\overline{\pi}$, cf. (4.2).

The proof of the Proposition 4.1.1 will be given in Section 4.5. We begin by observing that for $m_1,\ldots,m_k \in \mathbb{N}$ and $z_1,\ldots,z_k \in \mathbb{Z}^2$

$$\begin{aligned}&\mathbf{P}^{\omega}_{\beta_N,N}(S_{m_1} = z_1,\ldots,S_{m_k} = z_k)\\ &= \frac{1}{Z_{\beta_N,N}(0,0,N,\star)}\prod_{j=1}^{k+1}\mathbf{E}_N\left[e^{\sum_{n=m_{j-1}+1}^{m_j}(\beta_N w_{n,S_n}-\lambda(\beta_N))}\mathbb{1}_{S_{m_j}=z_j}\Big|S_{m_{j-1}} = z_{j-1}\right],\end{aligned}$$

where $m_0 = z_0 = 0$, $m_{k+1} = N$ and $z_{k+1} = \star$. For our purposes, it will be convenient to rewrite the above expression in terms of partition functions, conditioned on both the start and end point, i.e.

$$\begin{aligned}&\mathbf{P}^{\omega}_{\beta_N,N}(S_{m_1} = z_1,\ldots,S_{m_k} = z_k) \qquad (4.9)\\ &\quad = \frac{1}{Z_{\beta_N,N}(0,0,N,\star)}\prod_{j=1}^{k+1}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\,q_{m_j-m_{j-1}}(z_j-z_{j-1})\,.\end{aligned}$$

Here, we introduced the point-to-point partition functions

$$\begin{aligned}&\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\\ &:= \mathbf{E}_N\left[e^{\sum_{n=m_{j-1}+1}^{m_j}(\beta_N\omega_{n,S_n}-\lambda(\beta_N))}\Big|S_{m_{j-1}} = z_{j-1},\,S_{m_j} = z_j\right]\end{aligned}$$

and the shorthand $q_n(z)$, denoting the transition probability $\mathbf{P}_N(S_n = z)$ of the simple-random walk.

Remark 4.1.2. *The point-to-point partition function $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ also takes the disorder at the endpoint into consideration. However, it is more natural to compare the product of point-to-plane partition functions to point-to-point partition functions of the form*

$$Z_{\beta_N,N}(0,0 \mid N,z) := \mathbf{E}_N\left[e^{\sum_{n=1}^{N-1}(\beta_N\omega_{n,S_n}-\lambda(\beta_N))}\middle| S_0 = 0,\, S_N = z\right], \tag{4.10}$$

not taking the endpoint-disorder into account. The distinction in notation may be very subtle, but so is the difference between them. In fact, the difference between (4.10) and $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ vanishes in $L^2(\mathbb{P})$:

$$\begin{aligned}
&\|\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z) - Z_{\beta_N,N}(0,0 \mid N,z)\|^2_{L^2(\mathbb{P})}\\
&= \|Z_{\beta_N,N}(0,0 \mid N,z)\|^2_{L^2(\mathbb{P})}\mathbb{E}[(e^{\beta_N\omega_{N,z}-\lambda(\beta_N)}-1)^2]\\
&= \|Z_{\beta_N,N}(0,0 \mid N,z)\|^2_{L^2(\mathbb{P})}(e^{\lambda(2\beta_N)-2\lambda(\beta_N)}-1),
\end{aligned}$$

where we used the independence property of the disorder in the first equality. Because the first term on the right-hand side is uniformly bounded in N and $\lambda(2\beta_N) - 2\lambda(\beta_N) \sim \beta_N^2$, the L^2-difference vanishes. We will refer to both $Z_{\beta_N,N}(0,0 \mid N,z)$ and $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ as point-to-point partition functions, whenever the meaning is clear from the context.

With a slight abuse of notation, we will write for $0 \le s < t \le 1$

$$\mathfrak{Z}_{\beta_N,N}(sN,y \mid tN,z)\,,$$

instead of $\mathfrak{Z}_{\beta_N,N}(\lfloor sN\rfloor,y \mid \lfloor tN\rfloor,z)$. Similarly, for the point-to-plane partition function. Furthermore, for future reference, we introduce the plane-to-point partition function, which is defined as

$$Z_{\beta_N,N}(\widetilde{m},\star,m,z) := \mathbf{E}_N\left[e^{\sum_{n=\widetilde{m}}^{m-1}(\beta w_{n,S_n}-\lambda(\beta))}\middle| S_m = z\right]. \tag{4.11}$$

One can think of it as the partition function of a polymer starting in (m,z) and evolving backwards in time. We will refer to both "point-to-plane" and "plane-to-point" as "point-to-plane", whenever the context is clear.

Having representation (4.9) at hand, we see the necessity to understand the limiting behaviour of point-to-point partition functions, in order to analyse the finite-dimensional distributions of the polymer measure. To do so, we first prove a local limit theorem (Proposition 4.2.1), which allows to approxi-

mate point-to-point partition functions by the product of two point-to-plane partition functions:

$$\mathfrak{Z}_{\beta_N,N}(0,0\,|\,N,z) = Z_{\beta_N,N}(0,0,\tfrac{N}{2},\star)Z_{\beta_N,N}(\tfrac{N}{2},\star,N,z) + \varepsilon_N\,, \tag{4.12}$$

with ε_N vanishing in $L^2(\mathbb{P})$. Factorisations of this nature were considered previously in $d \geq 3$ [Sin95, Var06, CNN22, LZ22]. During completion of this work, Nakajima and Nakashima [NN23] proved independently a result similar to Proposition 4.2.1 in the continuum, see also Remark 4.2.3.

We want to put particular emphasis on [Sin95]. Their proof of the local limit theorem (4.12) in $d \geqslant 3$ uses the fact that polynomial chaos components (4.21) can be factorised, using a single random walk transition probability which is of order $q_N(z)$ [Sin95, Theorem 2]. Our proof of Lemma 4.2.6 resembles this approach. Additionally, they explain how a central limit theorem for the end-point distribution can be obtained from the above factorisation [Sin95, Theorem 4]. It is interesting to note, that [Sin95] only considers β's satisfying

$$\sigma(\beta)^2 R_\infty < 1\,, \tag{4.13}$$

where $R_\infty := \lim_{N\to\infty} R_N$, which is only finite in dimension $d \geq 3$. Formally generalising the condition with respect to the weak disorder limit in $d = 2$, (4.13) instead reads $\sigma(\beta_N)^2 R_N < 1$, which is equivalent to our assumption $\widehat{\beta} < 1$ due to $\sigma(\beta_N) \sim \beta_N$.

As a consequence of (4.12), the limiting distribution of point-to-point partition functions can be deduced from the corresponding point-to-plane partition functions approximating them. In [CSZ17b, Theorem 2.12] it was proved that finite families of partition functions of the 2D-DRPM converge jointly to a multivariate log-normal distribution. More precisely, consider the collection of space-time points $((n_i,z_i))_{1\leq i\leq k} = ((n_i(N),z_i(N)))_{1\leq i\leq k}$ such that for every $1 \leq i,j \leq k$

$$\begin{aligned} R_{N-n_i}/R_N &\to 1, \quad \text{as } N\to\infty, \\ \text{and} \quad \lim_{N\to\infty} R_{|n_i-n_j|\vee|z_i-z_j|^2}/R_N &= \zeta_{i,j} \in [0,1] \quad \text{exists,} \end{aligned}$$

then

$$\big(Z_{\beta_N,N}(n_i,z_i,N,\star)\big)_{1\leq i\leq k} \to (:e^{Y_i}:)_{1\leq i\leq k}, \tag{4.14}$$

where $(Y_i)_{1\leq i\leq k}$ is a multivariate Gaussian with

$$\mathbb{E}[Y_i] = 0 \quad \text{and} \quad \mathbb{E}[Y_i Y_j] = \log \frac{1-\widehat{\beta}^2\zeta_{i,j}}{1-\widehat{\beta}^2} \qquad \forall 1 \leq i,j \leq k. \tag{4.15}$$

Notably, the result holds for space-time points $((n_i, z_i))_{1\leq i\leq k}$ with positive macroscopic distance, in which case the corresponding tuple $(Y_i)_{1\leq i\leq k}$ consists of independent Gaussians because $\zeta_{i,j} = \mathbb{1}_{i\neq j}$.

Remark 4.1.3. *The correlation structure in* (4.15) *is precisely the reason for the different behaviour of the limiting polymer law in $d = 2$, compared to $d = 1$. In the one-dimensional case the partition functions have non-trivial dependency in the large N limit for macroscopically separated space-time points, see [AKQ14b], leading to a path measure singular to the Wiener measure. In the two-dimensional setting,* (4.15) *implies that partition functions started from macroscopically separated points will have independent limits, leading to an self-averaging effect for the polymer measure.*

After having dealt with the approximation of point-to-point partition functions, we can move on to the convergence of quenched polymer marginals. The greatest difficulty when dealing with the finite-dimensional marginals of the form (4.9) is the fact that none of the point-to-point partition functions is independent of the denominator $Z_{\beta_N,N}(0,0,N,\star)$. We now outline the approach taken in this chapter. For the sake of simplicity, we only explain the following for the end-point distribution.

First, we prove that the limiting annealed polymer marginal, i.e. $\lim_{N\to\infty}\mathbb{E}[\mathbf{P}^{\omega}_{\beta_N,N}(\frac{1}{\sqrt{N}}S_N \in \cdot)]$, agrees with the ones of Brownian motion, cf. Lemma 4.4.2. In fact, we show the much stronger result that the quenched marginal can be approximated in $L^1(\mathbb{P})$ by a simplified representation, without the partition function $Z_{\beta_N,N}(0,0,N,\star)$ in the denominator:

$$\lim_{N\to\infty}\left\|\mathbf{P}^{\omega}_{\beta_N,N}(\tfrac{1}{\sqrt{N}}S_N \in B) - \sum_{z\in\sqrt{N}B} Z_{\beta_N,N}(\tfrac{N}{2},\star,N,z)q_N(z)\right\|_{L^1(\mathbb{P})} = 0\,, \tag{4.16}$$

using the factorisation from (4.12). Here and throughout the chapter, $\sqrt{N}B$ denotes the set $\{z\in\mathbb{Z}^2 : \frac{z}{\sqrt{N}}\in B\}$. The expectation of the latter representation is immediate, which yields the annealed limit $\mathbf{P}(\frac{1}{\sqrt{2}}W_1 \in B)$.

In order to conclude convergence of the quenched marginal, the natural next step is to prove convergence of $\sum_{z\in\sqrt{N}B} Z_{\beta_N,N}(\frac{N}{2},\star,N,z)q_N(z)$ to its

mean in $L^1(\mathbb{P})$. Instead, we show convergence in $L^2(\mathbb{P})$:

$$\lim_{N\to\infty} \Big\| \sum_{z\in\sqrt{N}B} Z_{\beta_N,N}(\tfrac{N}{2},\star,N,z)q_N(z) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_1 \in B)\Big\|_{L^2(\mathbb{P})} = 0\,,$$

since this reduces to a simpler second-moment calculation. Estimating the second moment of $\sum_{z\in\sqrt{N}B} Z_{\beta_N,N}(\frac{N}{2},\star,N,z)q_N(z)$, requires careful evaluation of the partition functions' covariance structure, cf. Lemma 4.5.3, leading to a law-of-large number like behaviour. Together with (4.16), this concludes $L^1(\mathbb{P})$-convergence of the quenched end-point distribution.

After proving convergence of finite-dimensional distributions (Proposition 4.1.1), we show that for any fixed function $F \in C_b(C[0,1])$

$$\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)] \to \mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)], \qquad \text{as } N\to\infty, \quad \text{in } \mathbb{P}\text{-probability}, \tag{4.17}$$

by blending in ideas from the classical Donsker's invariance principle: Using tightness of the annealed polymer measure, it suffices to restrict the polymer paths to a compact set $K \subset C[0,1]$ when testing against a function $F \in C_b(C[0,1])$. The Stone-Weierstrass theorem then states that F can be approximated uniformly by cylinder functions on K, i.e. functions that only depend on finitely many marginals of the polymer path. Together with Proposition 4.1.1, this yields the functional central limit theorem (4.17).

Lastly, we prove equivalence of functional central limit theorem and invariance principle, using a countable weak convergence determining family of functions (Proposition 4.5.9). This yields weak convergence of the polymer measures as stated in Theorem 4.0.1. The same argument allows to rewrite the functional central limit theorem for $d \geq 3$ [CY06], in terms of an invariance principle, cf. Corollary 4.5.11.

Remark 4.1.4. *We want to stress that* (4.17) *is not yet a classical invariance principle, stating convergence of the polymer measure. Instead it is a central limit theorem, stating convergence when paths are tested against individual test functions. To emphasise this point, we note that a "true" invariance principle (in* $\mathbb{P}$*-probability) reads as follows: For every sequence* $(N_j)_{j\in\mathbf{N}}$ *in* $\mathbf{N}$ *there exists a subsequence* $(N_{j_m})_{m\in\mathbf{N}}$ *and a set* $\overline{\Omega} \subset \Omega$ *of full measure such that*

$$\pi^*_{N_{j_m}} \mathbf{E}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}[F(X)] \to \mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)] \quad \forall F \in C_b(C[0,1]),$$

for every $\omega \in \overline{\Omega}$. In (4.17) on the other hand, we fix a function $F \in C_b(C[0,1])$ for which there exists a subsequence and a set $\overline{\Omega}$ of full mass, on which the convergence holds. The dependency of $(N_{j_m})_{m\in\mathbf{N}}$ and $\overline{\Omega}$ on F does not allow us to exchange the order of quantifiers without further reasoning.

Structure of the chapter

The remainder of this chapter is structured as follows. In Section 4.2 we prove that point-to-point partition functions can be approximated by point-to-plane partition functions. In Section 4.4 we use this fact to prove convergence of the annealed finite-dimensional distributions of the polymer measure to the ones of a Brownian motion. Together with a tightness argument, this yields the annealed invariance principle (Proposition 4.4.1). Section 4.5 is divided into three parts. First, we prove Proposition 4.1.1, where we use the self-averaging behaviour described above. We then present the proof of the invariance principle (Theorem 4.0.1). Finally, we show the local limit theorem for the polymer marginals on microscopic scales (Proposition 4.0.2).

4.2 Approximating point-to-point partition functions

This section's main result is summarised in the following proposition. It states that a point-to-point partition function can be locally uniformly approximated by the product of a point-to-plane and plane-to-point partition function.

Proposition 4.2.1. *Let $\widehat{\beta} \in (0,1)$, $x \in \mathbf{R}^2$ and $r > 0$. Then for $0 < s^+ < t^- < 1$ we have*

$$\left\| \mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z) - Z_{\beta_N,N}(0,0,s^+N,\star) Z_{\beta_N,N}(t^-N,\star,N,z) \right\|_{L^2(\mathbb{P})} \to 0,$$

uniformly over all $z \in \sqrt{N}B(x,r)$ such that $q_N(z) > 0$. The statement remains true when replacing $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ with $Z_{\beta_N,N}(0,0 \mid N,z)$, which we introduced in (4.10).

Remark 4.2.2. *(i) Proposition 4.2.1 also holds for $s^+ = 1$ and $t^- = 0$ in terms of $L^{1+\delta}(\mathbb{P})$-convergence for some $\delta > 0$ small enough. This can be shown following the same steps in the proof of Proposition 4.2.1. After completion of this paper, it was proved that $\sup_{N\in\mathbf{N}} \mathbb{E}[\mathfrak{Z}_{\beta_N,N}(0,0,N,\star)^p] < \infty$ for arbi-*

trary $p > 0$ [CZ23, LZ23]. This stronger moment estimate allows to lift the mode of convergence from $L^{1+\delta}(\mathbb{P})$ to $L^2(\mathbb{P})$.

(ii) Note that Proposition 4.2.1 remains true when replacing the initial time $s = 0$ and final time $t = 1$ with arbitrary values $0 \leq s < t \leq 1$, i.e. when considering partition functions $Z_{\beta_N,N}(sN, 0 \mid tN, z)$ or $\mathfrak{Z}_{\beta_N,N}(sN, 0 \mid tN, z)$.

Throughout this section, the point $x \in \mathbf{R}^2$ plays the role of the macroscopic endpoint of the polymer path. In particular, $\sqrt{N}B(x, r)$ includes all microscopic points which are close to x on a macroscopic scale.

Remark 4.2.3. *A similar result to Proposition 4.2.1 was obtained in [NN23, Theorem 2.8]. They proved that the point-to-point partition function of a directed random polymer in the continuum can be approximated in $L^2(\mathbb{P})$ by the product of point-to-plane partition functions with mesoscopic time-horizon, if the distance between start and terminal space-point is not too large. Similar to the present paper, they show that contributions to the point-to-point partition function only come from the environment close to start and endpoint, before they replace the Brownian Bridge measure by two Brownian motions running independently forward and backward in time. For the partition function of a polymer of length N, an equivalent result to [NN23, Theorem 2.8] for the discrete case would read as follows: Uniformly over all $z \in \mathbf{Z}^2$ with $|z| \leq \sqrt{N \log N}$ and $q_N(z) > 0$*

$$\left\| \mathfrak{Z}_{\beta_N,N}(0, 0 \mid N, z) - Z_{\beta_N,N}(0, 0, l_N, \star) Z_{\beta_N,N}(N - l_N, \star, N, z) \right\|^2_{L^2(\mathbb{P})} \to 0 \,,$$

for $l_N = N^{1-(\log N)^{\gamma-1}}$ with $\gamma \in (0, 1)$. In contrast to Proposition 4.2.1, here the radius of uniformity is $\sqrt{N \log N}$. This is due to keeping track of vanishing rates in their proof, which allows to strengthen the result. Likewise, exact evaluation of the quantities in Lemma 4.2.7 using the local limit theorem should allow to increase the radius of uniformity in Proposition 4.2.1 to the same order. However, in regards of our main result this is not necessary.

Before continuing, we remind the reader that the partition function $Z_{\beta_N,N}(0, 0, N, \star)$ can be written in terms of a polynomial chaos decomposition, which is the discrete analogue of Proposition 2.1.10. More precisely, the partition function has an explicit discrete chaos expansion [CSZ17b, CSZ20]:

$$\begin{aligned} Z_{\beta_N,N}(0, 0, N, \star) &= \mathbf{E}_N \left[e^{\sum_{n=1}^N \sum_{z \in \mathbf{Z}^2} (\beta_N \omega_{n,z} - \lambda(\beta_N)) \mathbb{1}_{S_n = z}} \right] \\ &= \mathbf{E}_N \Big[\prod_{n=1}^N \prod_{z \in \mathbf{Z}^2} (1 + \sigma_N \eta_{n,z} \mathbb{1}_{S_n = z}) \Big] \end{aligned} \tag{4.18}$$

$$= 1 + \sum_{k=1}^{N} Z^{(k)}_{\beta_N,N}(0,0,N,\star),$$

where $Z^{(k)}_{\beta_N,N}(0,0,N,\star)$ is defined below and

$$\eta_{n,z} = \eta^{(N)}_{n,z} := \frac{1}{\sigma_N}(e^{\beta_N\omega_{n,z}-\lambda(\beta_N)} - 1)$$

being centred i.i.d. random variables with unit variance. We also wrote $\sigma_N := \sigma(\beta_N)$, which we shall use throughout the chapter. In the last equality of (4.18), the expansion of the products gives rise to the k-th homogeneous chaos denoted by

$$Z^{(k)}_{\beta_N,N}(0,0,N,\star) := \sigma_N^k \sum_{\substack{1\le n_1<\cdots<n_k\le N \\ z_1,\ldots,z_k\in\mathbf{Z}^2}} \left(\prod_{i=1}^{k} q_{n_i-n_{i-1}}(z_i - z_{i-1})\eta_{n_i,z_i}\right), \quad (4.19)$$

with (n_0, z_0) equal the origin $(0,0) \in \mathbf{N}\times\mathbf{Z}^2$. The terms in the series expansion above are orthogonal in the sense that $\mathbb{E}[Z^{(k)}_{\beta_N,N}(0,0,N,\star)Z^{(j)}_{\beta_N,N}(0,0,N,\star)] = 0$, whenever $k \neq j$, due to the different number of disorder-terms considered, cf. (2.11) in the continuum. Throughout the paper, we will use this fact in second moment computations without further explanation.

An analogous expansion holds for the plane-to-point partition function $Z_{\beta_N,N}(0,\star,N,z)$ with

$$Z^{(k)}_{\beta_N,N}(0,\star,N,z) \quad (4.20)$$
$$:= \sigma_N^k \sum_{\substack{0\le n_1<\cdots<n_k\le N-1 \\ z_1,\ldots,z_k\in\mathbf{Z}^2}} \left(\prod_{i=1}^{k} q_{n_{i+1}-n_i}(z_{i+1} - z_i)\eta_{n_i,z_i}\right),$$

where we assumed $(n_{k+1}, z_{k+1}) = (N, z)$ and used the symmetry of the transition probabilities of the simple random walk. Similarly, for the point-to-point partition function, we write

$$Z_{\beta_N,N}(0,0 \mid N,z) = 1 + \sum_{k=1}^{N} Z^{(k)}_{\beta_N,N}(0,0 \mid N,z)$$

and

$$Z^{(k)}_{\beta_N,N}(0,0 \mid N,z) \tag{4.21}$$
$$:= \sigma_N^k \sum_{\substack{1\le n_1<\cdots<n_k\le N-1 \\ z_1,\dots,z_k\in\mathbf{Z}^2}} \left(\prod_{i=1}^{k} q_{n_i-n_{i-1}}(z_i-z_{i-1})\eta_{n_i,z_i}\right) \frac{q_{N-n_k}(z-z_k)}{q_N(z)}.$$

We point out that the point-to-plane partition function can be recovered by taking the average over all possible endpoints and include the endpoint-disorder:

$$\begin{aligned} Z_{\beta_N,N}(0,0,N,\star) &= \sum_{z\in\mathbf{Z}^2} \mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z) q_N(z) \\ &= \sum_{z\in\mathbf{Z}^2} Z_{\beta_N,N}(0,0 \mid N,z) e^{\beta_N\omega_{N,z}-\lambda(\beta_N)} q_N(z). \end{aligned}$$

The proof of Proposition 4.2.1 relies on the fact that the main contribution of the point-to-point partition function comes from two mesoscopic sized subsets of the space-time domain around the start and terminal point. These space-time areas are the same as the ones giving main contribution to the partition functions $Z_{\beta_N,N}(0,0,N,\star)$, see [CSZ20]. For a microscopic reference point $(n,z) \in \mathbb{N}\times\mathbf{Z}^2$, we define such sets both forward and backward in time:

$$A_N^{\pm}(n,z) := \{(m,y) \,:\, |y-z| \le N^{1/2-a_N/4} \text{ and } 0 \le \pm(m-n) \le N^{1-a_N}\},$$

where $a_N := (\log N)^{\gamma-1}$ for some $\gamma \in (0,1)$.

As already mentioned above, only samples inside of $A_N^+(0,0)\cup A_N^-(N,z)$ will contribute to the L^2-limit of $Z^{(k)}_{\beta_N,N}(0,0 \mid N,z)$. On this account, we introduce the following decomposition

$$Z^{(k)}_{\beta_N,N}(0,0 \mid N,z) = Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z) + \widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z),$$

where $Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z)$ denotes the sum

$$\sigma_N^k \sum_{\substack{1\le n_1<\cdots<n_k\le N-1 \\ z_1,\dots,z_k\in\mathbf{Z}^2 \\ (n_i,z_i)\in A_N^+(0,0)\cup A_N^-(N,z)}} \left(\prod_{i=1}^{k} \eta_{n_i,z_i}\, q_{n_i-n_{i-1}}(z_i-z_{i-1})\right) \frac{q_{N-n_k}(z-z_k)}{q_N(z)}, \tag{4.22}$$

and $\widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z)$ the corresponding remainder. This restriction can be thought of as "turning off" the disorder outside the two boxes; for visualisation the reader may refer to Figure 4.2.1 below. Similarly, we separate

$$Z_{\beta_N,N}(0,0 \mid N,z) = Z^A_{\beta_N,N}(0,0 \mid N,z) + \widehat{Z}_{\beta_N,N}(0,0 \mid N,z)\,, \tag{4.23}$$

with $Z^A_{\beta_N,N}(0,0 \mid N,z) = \sum_{k=0}^N Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z)$. Sometimes we will also use this notation for point-to-plane partition functions. Then, $Z^{(j),A}_{\beta_N,N}(0,0,N,\star)$ and $Z^{(j),A}_{\beta_N,N}(0,\star,N,z)$ will stand for the multi-linear polynomials in (4.19) and (4.20) with disorder restricted to $A_N^+(0,0)$ or $A_N^-(N,z)$, respectively. Moreover, we want to mention that $Z^{(j),A}_{\beta_N,N}(0,0,N,\star) = 0$ for all $j > N^{1-a_N}$, hence,

$$Z^A_{\beta_N,N}(0,0,N,\star) = \sum_{j=0}^{\lfloor N^{1-a_N} \rfloor} Z^{(j),A}_{\beta_N,N}(0,0,N,\star)\,, \tag{4.24}$$

and similarly for the plane-to-point partition function. Lastly, note that the size of $A_N^\pm$ only depends on the level of approximation N and **not** the point-to-point partition function's final time (which sometimes happens to agree with N).

We will frequently make use of higher-order and negative moment estimates of the partition function. We summarise equations (3.12), (3.14) and (3.15) from [CSZ20] in the following lemma.

Lemma 4.2.4 ([CSZ20]). *Let $\widehat{\beta} \in (0,1)$.*

(i) Then there exists a $\delta = \delta(\widehat{\beta}) > 0$ and a constant $C'_{\widehat{\beta}} < \infty$ such that for every $p \in [2, 2+\delta]$

$$\sup_{N\in\mathbb{N}} \mathbb{E}\left[Z_{\beta_N,N}(0,0,N,\star)^p\right] \le C'_{\widehat{\beta}}\,, \quad \sup_{N\in\mathbb{N}} \mathbb{E}\left[Z^A_{\beta_N,N}(0,0,N,\star)^p\right] \le C'_{\widehat{\beta}}$$
$$\text{and} \quad \sup_{N\in\mathbb{N}} \mathbb{E}\left[\widehat{Z}_{\beta_N,N}(0,0,N,\star)^p\right] \le C'_{\widehat{\beta}}(a_N)^{p/2}.$$

(ii) For every $p > 0$, there exists a constant $\widehat{C}_{\widehat{\beta},p} > 0$ such that

$$\sup_{N\in\mathbb{N}} \mathbb{E}\left[Z_{\beta_N,N}(0,0,N,\star)^{-p}\right] \le \widehat{C}_{\widehat{\beta},p}$$
$$\text{and} \quad \sup_{N\in\mathbb{N}} \mathbb{E}\left[Z^A_{\beta_N,N}(0,0,N,\star)^{-p}\right] \le \widehat{C}_{\widehat{\beta},p}.$$

Lemma 4.2.4(i) is a consequence of hypercontractivity of (discrete) polynomial chaoses, recall Proposition 2.1.7 for the continuum version. Because $Z_{\beta_N,N}(0,0,N,\star)$ and $Z_{\beta_N,N}(0,\star,N,z)$ have the same distribution, all statements in Lemma 4.2.4 also hold for the plane-to-point partition function.

Remark 4.2.5. *That Lemma 4.2.4(ii) holds for arbitrary negative moments, is a consequence of the concentration inequality* (4.1) *and a left-tail estimate, see [CSZ20, Proposition 3.1]. We will use this control, in the proof of Lemma 4.4.2 and Proposition 4.0.2, to separate products of partition functions using Hölder's inequality, cf.* (4.52) *below. From the estimate* (4.52) *we will see that control of either large negative or positive moments of the partition function is sufficient. Hence, with an improved control on positive moments, it should be possible to push control from negative moments to positive ones. As mentioned in Remark 4.2.2, it was recently shown that all positive moments of the partition function are uniformly bounded in N [CZ23, LZ23] (without the need of assumption* (4.1)*), and it is expected that all statements in Lemma 4.2.4(i) continue to hold for arbitrary $p > 0$. Thus, assumption* (4.1) *could then be replaced by $\sup_{N\in\mathbb{N}} \mathbb{E}\left[Z_{\beta_N,N}(0,0,N,\star)^{-p}\right] < \infty$ for some $p > 2$.*

Although we will follow a similar approach in evaluating the limits of $Z^{A}_{\beta_N,N}(0,0 \mid N,z)$ and $\widehat{Z}_{\beta_N,N}(0,0 \mid N,z)$, we present their respective proofs separately for the sake of a more approachable presentation. We start with the chaos decomposition restricted to the macroscopically vanishing set $A^{+}_{N}(0,0) \cup A^{-}_{N}(N,z)$.

4.2.1 A single jump factorises the term

For the term $Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z)$ we only consider sample points $(n_i,z_i)_{i=1}^{k}$ that lie inside $A^{+}_{N}(0,0) \cup A^{-}_{N}(N,z)$, by restricting the sum in (4.22). Notably, this implies for every polymer path with k intermediate samples the existence of a single **"large" jump** across the valley separating $(0,0)$ and (N,z). This jump divides the polymer samples into two chains, one close to the starting point $(0,0)$, the other one close to the terminal point (N,z). These two chains can be analysed independently.

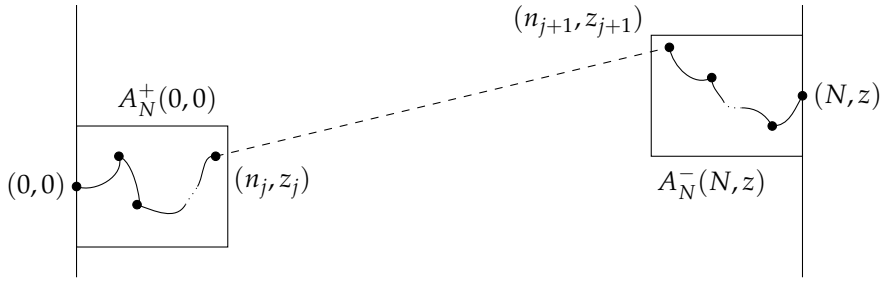


Figure 4.1: Restricting samples to the macroscopic vanishing boxes $A_N^{\pm}$ implies the existence of a large jump over time.

Following the above explanation, it is reasonable to rewrite (4.22) by summing over all possible positions of this "large" jump:

$$\begin{aligned} &Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z) \qquad (4.25)\\ &= \sigma_N^k \sum_{j=0}^{k} \sum_{\substack{(n_i,z_i)\in A_N^+(0,0)\forall i\le j\\ (n_i,z_i)\in A_N^-(N,y)\forall i>j\\ \text{s.t. } 1\le n_1<\cdots<n_k\le N-1}} \left(\prod_{i=1}^{k} q_{n_i-n_{i-1}}(z_i-z_{i-1})\eta_{n_i,z_i}\right) \frac{q_{N-n_k}(z-z_k)}{q_N(z)}, \end{aligned}$$

where j is the largest index before the jump, see also Figure 4.2.1. For the sake of clarity, we will abbreviate the conditions in the sum by $(n_i,z_i)_{i=1}^k \in \mathfrak{A}(j)$ and write $\sum_{(n_i,z_i)_{i=1}^k\in\mathfrak{A}(j)}$ coherently. Most notably, when omitting the ratio $q_{n_{j+1}-n_j}(z_{j+1}-z_j)/q_N(z)$ in (4.25), the right-hand side simplifies to

$$\sum_{j=0}^{k} Z^{(j),A}_{\beta_N,N}(0,0,N,\star) Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z),$$

and the two partition functions inside the sum are stochastically independent. Here, we used the notation defined in (4.19) and (4.20) and enhanced thereafter. This motivates the following lemma.

Lemma 4.2.6. *For all $x \in \mathbb{R}^2$ and $r > 0$, we have*

$$\sup_{\substack{z\in\sqrt{N}B(x,r)\\ \text{s.t. } q_N(z)>0}} \left\| Z^A_{\beta_N,N}(0,0 \mid N,z) - Z^A_{\beta_N,N}(0,0,N,\star) Z^A_{\beta_N,N}(0,\star,N,z) \right\|_{L^2(\mathbb{P})} \to 0,$$

as N tends to infinity.

Proof. Let $x \in \mathbb{R}^2$ and $r > 0$. We begin by noting that

$$Z^A_{\beta_N,N}(0,0,N,\star) Z^A_{\beta_N,N}(0,\star,N,z) \qquad (4.26)$$

$$= \sum_{k=0}^{N} \sum_{j=0}^{k} Z^{(j),A}_{\beta_N,N}(0,0,N,\star) Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z)\,.$$

This follows from the fact that $Z^{(j),A}_{\beta_N,N}(0,0,N,\star) = 0$ whenever $j > N^{1-a_N}$ (similarly for $Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z)$), thus,

$$\begin{aligned}
&\sum_{k=0}^{N} \sum_{j=0}^{k} Z^{(j),A}_{\beta_N,N}(0,0,N,\star) Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z) \\
&\qquad = \sum_{j=0}^{N^{1-a_N}} Z^{(j),A}_{\beta_N,N}(0,0,N,\star) \sum_{k=j}^{j+N^{1-a_N}} Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z)\,,
\end{aligned}$$

where we changed the order of the sums and added the restrictions $j, k-j \leq N^{1-a_N}$. After an index shift in the inner sum, we see that (4.26) holds true due to (4.24).

Now, expanding the right-hand side of (4.26), we have

$$\begin{aligned}
&Z^{A}_{\beta_N,N}(0,0,N,\star) Z^{A}_{\beta_N,N}(0,\star,N,z) \\
&\qquad = \sum_{k=0}^{N} \sum_{j=0}^{k} \sigma_N^{2j} \sum_{\substack{(n_i,z_i)_{i=1}^{j} \in A_N^+(0,0) \\ \text{s.t. } 1 \leq n_1 < \cdots < n_j}} \Big(\prod_{i=1}^{j} q_{n_i - n_{i-1}}(z_i - z_{i-1}) \eta_{n_i,z_i} \Big) \qquad (4.27) \\
&\qquad\qquad \times \sigma_N^{2(k-j)} \sum_{\substack{(\widetilde{n}_i,\widetilde{z}_i)_{i=1}^{k-j} \in A_N^-(N,z) \\ \text{s.t. } \widetilde{n}_1 < \cdots < \widetilde{n}_{k-j} \leq N-1}} \Big(\prod_{i=1}^{k-j} q_{\widetilde{n}_{i+1} - \widetilde{n}_i}(\widetilde{z}_{i+1} - \widetilde{z}_i) \eta_{\widetilde{n}_i,\widetilde{z}_i} \Big)\,.
\end{aligned}$$

Hence, using representations (4.25) and (4.27), we can estimate the second moment in the statement of the lemma, due to orthogonality, in terms of

$$\begin{aligned}
&\sum_{k=0}^{N} \mathbb{E}\Big[\Big(Z^{(k),A}_{\beta_N,N}(0,0 \mid N,z) - \sum_{j=0}^{k} Z^{(j),A}_{\beta_N,N}(0,0,N,\star) Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z) \Big)^2 \Big] \\
&\leq \sum_{k=0}^{N} \sum_{j=0}^{k} \mathbb{E}\Big[\big(Z^{(j),A}_{\beta_N,N}(0,0,N,\star) Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z) \big)^2 \qquad (4.28) \\
&\qquad\qquad \times \sup_{\substack{(n_j,z_j) \in A^+(0,0) \\ (n_{j+1},z_{j+1}) \in A^-(N,z)}} \Big(\frac{q_{n_{j+1}-n_j}(z_{j+1} - z_j)}{q_N(z)} - 1 \Big)^2 \Big].
\end{aligned}$$

Therefore, we only need to show that the ratio $q_{n_{j+1}-n_j}(z_{j+1} - z_j)/q_N(z)$ is negligible, which is intuitively clear since (n_j, z_j) and (n_{j+1}, z_{j+1}) are on a

macroscopic level close to $(0,0)$ and (N,z). Using the estimate

$$\sup_{\substack{(n_j,z_j)\in A^+(0,0)\\(n_{j+1},z_{j+1})\in A^-(N,z)}} \left|\frac{q_{n_{j+1}-n_j}(z_{j+1}-z_j)}{q_N(z)}-1\right| \leq \sup_{\substack{|N-n|<2N^{1-a_N}\\|z-y|<2N^{1/2-a_N/4}}} \left|\frac{q_n(y)}{q_N(z)}-1\right|, \tag{4.29}$$

we can upper bound (4.28) further (note that the right-hand side of (4.29) does not depend on k or j), which finally yields

$$\begin{aligned}
&\sum_{k=0}^{N}\mathbb{E}\Big[\Big(Z^{(k),A}_{\beta_N,N}(0,0\mid N,z)-\sum_{j=0}^{k}Z^{(j),A}_{\beta_N,N}(0,0,N,\star)Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z)\Big)^2\Big]\\
&\leq \sup_{\substack{|N-n|<2N^{1-a_N}\\|z-y|<2N^{1/2-a_N/4}}}\left|\frac{q_n(y)}{q_N(z)}-1\right|^2\\
&\qquad\times\sum_{k=0}^{N}\sum_{j=0}^{k}\mathbb{E}\left[\Big(Z^{(j),A}_{\beta_N,N}(0,0,N,\star)Z^{(k-j),A}_{\beta_N,N}(0,\star,N,z)\Big)^2\right].
\end{aligned}$$

Again, the sum of the right hand side is uniformly bounded in N and z, cf. Lemma 4.2.4. The last step consists of showing that the ratio of random-walk-transition-kernels is indeed uniformly close to one, i.e.

$$\sup_{\substack{z\in\sqrt{N}B(x,r)\\ \text{s.t. } q_N(z)>0}}\ \sup_{\substack{|N-n|<2N^{1-a_N}\\|z-y|<2N^{1/2-a_N/4}}}\left|\frac{q_n(y)}{q_N(z)}-1\right|\to 0,\quad \text{as } N\to\infty. \tag{4.30}$$

This follows directly from Lemma 4.2.7(i) below, and finishes the proof. □

The following lemma allows us to control the ratio of random walk transition probabilities, which is necessary for the previous proof and will be used throughout the chapter.

Lemma 4.2.7. *Let $x\in\mathbf{R}^2$ and $r>0$.*

(i) For $a_N:=(\log N)^{\gamma-1}$, $\gamma\in(0,1)$, we have

$$\sup_{\substack{z\in\sqrt{N}B(x,r)\\ s.t.\ q_N(z)>0}}\ \sup_{\substack{|N-n|<2N^{1-a_N}\\|z-y|<2N^{1/2-a_N/4}}}\left|\frac{q_n(y)}{q_N(z)}-1\right|\to 0\quad \text{as } N\to\infty. \tag{4.31}$$

(ii) There exists a constant $C > 0$ such that for all $k \in \mathbb{N}$

$$\sup_{\substack{z \in \sqrt{N}B(x,r) \\ \text{s.t. } q_N(z)>0}} \sup_{\substack{n \geq N/k \\ y \in \mathbf{Z}^2}} \frac{q_n(y)}{q_N(z)} \leq Ck\,, \tag{4.32}$$

for all N large enough.

Proof. (i) We want to apply the local limit theorem for simple random walks [LL10], which is why we write

$$\sup_{\substack{|N-n|<2N^{1-a_N} \\ |z-y|<2N^{1/2-a_N/4}}} \left| \frac{q_n(y)}{q_N(z)} - 1 \right| = \frac{1}{N\,q_N(z)} \sup_{\substack{|N-n|<2N^{1-a_N} \\ |z-y|<2N^{1/2-a_N/4}}} |N\,q_n(y) - N\,q_N(z)|\,.$$

Because $N\,q_N(z)$ converges uniformly in z and its limit is lower bounded by $2\,\inf_{\widetilde{x}\in B(x,r)} p_{1/2}(\widetilde{x})$ whenever $q_N(z) > 0$, we can ignore the factor in front of the supremum. Adding and subtracting $2\,p_{1/2}(z/\sqrt{N})$ yields

$$\begin{aligned}
&\sup_{\substack{|N-n|<2N^{1-a_N} \\ |z-y|<2N^{1/2-a_N/4}}} |N\,q_n(y) - N\,q_N(z)| \\
&\leq \sup_{\substack{|N-n|<2N^{1-a_N} \\ |z-y|<2N^{1/2-a_N/4}}} \Big(\Big|N\,q_n(y) - 2\,p_{\frac{1}{2}}(z/\sqrt{N})\Big| + \Big|2\,p_{\frac{1}{2}}(z/\sqrt{N}) - N\,q_N(z)\Big| \Big).
\end{aligned}$$

The second term on the right-hand side vanishes uniformly in z by the local limit theorem. Because y and z are arbitrarily close on the macroscopic scale, the first term vanishes for the same reason.

(ii) We begin by noting that for $z \in \sqrt{N}B(x,r)$ with $q_N(z) > 0$, we have

$$\sup_{\substack{n \geq N/k \\ y \in \mathbf{Z}^2}} \frac{q_n(y)}{q_N(z)} \leq k\,\frac{1}{N\,q_N(z)} \sup_{\substack{n \geq N/k \\ y \in \mathbf{Z}^2}} n\,q_n(y),$$

where we may again ignore the factor $N\,q_N(z)$ for the same reason as in (i). Hence, it suffices to prove the existence of a constant $C > 0$ such that

$$\sup_{\substack{n \geq N/k \\ y \in \mathbf{Z}^2}} n\,q_n(y) \leq C \qquad \forall\, k \in \mathbb{N}\,,$$

for N large enough. We make the following choice for C: Note that

$$\begin{aligned}\sup_{y\in\mathbf{Z}^2} n\, q_n(y) &\le \sup_{y\in\mathbf{Z}^2} \left(|n\, q_n(y) - 2p_{\frac{1}{2}}(y/\sqrt{n})| + |2p_{\frac{1}{2}}(y/\sqrt{n})| \right) \\ &\le \sup_{y\in\mathbf{Z}^2} |n\, q_n(y) - 2p_{\frac{1}{2}}(y/\sqrt{n})| + |2p_{\frac{1}{2}}(0)| =: C_n.\end{aligned}$$

Since the first term on the right-hand side converges in n by the local limit theorem, C_n is uniformly bounded in n. Thus, we have

$$\sup_{n\ge N/k} \sup_{y\in\mathbf{Z}^2} n\, q_n(y) \le \sup_{n\in\mathbb{N}} C_n =: C < \infty.$$

This finishes the proof. □

4.2.2 Multiple exceptional jumps are negligible

Lemma 4.2.6 states that instead of looking at the point-to-point partition function restricted to have a single large jump from $A_N^+(0,0)$ to $A_N^-(N,z)$, it suffices to look at the product of two point-to-plane partition functions, one looking forward the other one looking backward in time. In order to prove the stronger result in Proposition 4.2.1, it remains to show that samples with points outside of $A_N^+(0,0) \cup A_N^-(N,z)$ do not contribute to the L^2-limit of the partition function. We say that samples of this kind have an **exceptional jump**, if there exists an index $1 \le \widehat{j} \le k$ such that $(n_{\widehat{j}}, z_{\widehat{j}}) \notin A_N^+(0,0) \cup A_N^-(N,z)$. We will use $(n_i, z_i)_{i=1}^k \notin \mathfrak{A}$ as a shorthand.

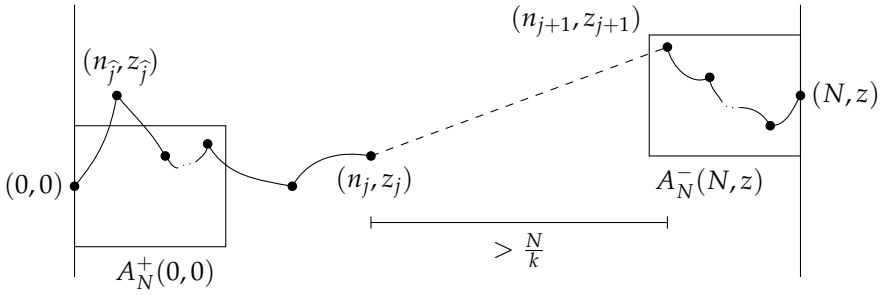


Figure 4.2: When considering k samples, the existence of a time-jump that is at least of length $\frac{N}{k}$, still allows us to split samples into two groups (while paying a multiplicative constant k). At least one of the two groups will contain samples outside the boxes $A_N^{\pm}$.

Lemma 4.2.8. *For every $x \in \mathbf{R}^2$ and $r > 0$, we have*

$$\lim_{N\to\infty} \sup_{z\in\sqrt{N}B(x,r)} \left\|\widehat{Z}_{\beta_N,N}(0,0 \mid N,z)\right\|_{L^2(\mathbb{P})} = 0.$$

Proof. Similar to the case where all samples are sufficiently close to the start and end point, we can again partition samples into two groups. First note that for every collection of k samples, there is at least one index $0 \le j \le k$ such that $n_{j+1} - n_j \ge \frac{N}{k}$, cf. Figure 4.2.2 (here we use again the notation $n_{k+1} = N$). Consequently, we may write

$$\begin{aligned}
&\mathbb{E}[(\widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z))^2] \\
&\le \sigma_N^{2k} \sum_{j=0}^{k} \sum_{(n_i,z_i)_{i=1}^k \notin \mathfrak{A}} \mathbb{1}_{n_{j+1}-n_j \ge \frac{N}{k}} \left(\prod_{i=1}^{j} q^2_{n_i-n_{i-1}}(z_i - z_{i-1})\right) \\
&\qquad\qquad \times \frac{q^2_{n_{j+1}-n_j}(z_{j+1}-z_j)}{q^2_N(z)} \left(\prod_{i=j+2}^{k+1} q^2_{n_i-n_{i-1}}(z_i - z_{i-1})\right).
\end{aligned} \tag{4.33}$$

Once more we want to separate indices into two groups, which requires the removal of a ratio of random walk transition kernels. First we apply Lemma 4.2.7(ii), which yields

$$\frac{q^2_{n_{j+1}-n_j}(z_{j+1}-z_j)}{q^2_N(z)} \le \sup_{\substack{z\in\sqrt{N}B(x,r)\\ \text{s.t. } q_N(z)>0}} \sup_{\substack{n\ge N/k\\ z_{j+1}-z_j\in\mathbf{Z}^2}} \frac{q^2_n(z_{j+1}-z_j)}{q^2_N(z)} \le Ck^2, \tag{4.34}$$

with C being a non-negative constant, independent of N and k.

In the remainder of the proof, we do not explicitly state results to hold locally uniformly in z. However, the reader should note that the statements remain true when adding the supremum over $\{z \in \sqrt{N}B(x,r) : q_N(z) > 0\}$ in front of all expressions below.

Combining (4.33) and (4.34), we just proved that

$$\begin{aligned}
&\mathbb{E}[(\widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z))^2] \\
&\le Ck^2\, \sigma_N^{2k} \sum_{j=0}^{k} \sum_{\substack{(n_i,z_i)_{i=1}^k \notin \mathfrak{A}\\ n_{j+1}-n_j \ge N/k}} \left(\prod_{i=1}^{j} q^2_{n_i-n_{i-1}}(z_i - z_{i-1})\right) \left(\prod_{i=j+2}^{k+1} q^2_{n_i-n_{i-1}}(z_i - z_{i-1})\right),
\end{aligned} \tag{4.35}$$

for some positive constant C. Since $(n_i, z_i)_{i=1}^k \notin \mathfrak{A}$, we know there exists a $\widehat{j} \in \{1, \dots, k\}$ such that $(n_{\widehat{j}}, z_{\widehat{j}}) \notin A_N^+(0,0) \cup A_N^-(N,z)$. Note that $\widehat{j}$ does not necessarily agree with j, and we consider the two cases $\widehat{j} \le j$ and $\widehat{j} > j$ separately. First assume that $\widehat{j} \le j$, then we can estimate the products in (4.35), which do not contain $\widehat{j}$, as follows

$$\sum_{\substack{0 \le n_{j+1} < \cdots < n_k \le N-1 \\ z_{j+1}, \dots, z_k \in \mathbf{Z}^2}} \left(\prod_{i=j+2}^{k+1} q^2_{n_i - n_{i-1}}(z_i - z_{i-1}) \right) \le \Big(\sum_{\substack{1 \le n \le N \\ \widetilde{z} \in \mathbf{Z}^2}} q_n^2(\widetilde{z}) \Big)^{k-j} = R_N^{k-j},$$

where we dropped the constraint of the n_i's being ordered and also used the symmetry of the simple random walk transition probabilities in the first step. Thus,

$$\begin{aligned} & k^2 \, \sigma_N^{2k} \sum_{j=0}^{k} \sum_{\substack{(n_i, z_i)_{i=1}^k \notin \mathfrak{A} \\ n_{j+1} - n_j \ge N/k}} \mathbb{1}_{\widehat{j} \le j} \left(\prod_{i=1}^{j} q^2_{n_i - n_{i-1}}(z_i - z_{i-1}) \right) \left(\prod_{i=j+2}^{k+1} q^2_{n_i - n_{i-1}}(z_i - z_{i-1}) \right) \\ & \le Ck^2 \sum_{j=0}^{k} (\sigma_N^2 R_N)^{k-j} \sigma_N^{2j} \sum_{\substack{1 \le n_1 < \cdots < n_j \le N \\ (z_i)_{i=1}^j \in (\mathbf{Z}^2)^j \\ \exists \widehat{j} \text{ s.t. } (n_{\widehat{j}}, z_{\widehat{j}}) \notin A_N^+(0,0)}} \left(\prod_{i=1}^{j} q^2_{n_i - n_{i-1}}(z_i - z_{i-1}) \right) \qquad (4.36) \\ & = Ck^2 \sum_{j=0}^{k} (\sigma_N^2 R_N)^{k-j} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N, N}(0,0,N,\star))^2]. \end{aligned}$$

We can perform the same estimates in the case of $\widehat{j} > j$ (in fact, we overestimate with $\widehat{j} \ge j$), with the roles of $A_N^+(0,0)$ and $A_N^-(N,z)$ reversed. Together with (4.35) and (4.36), this yields

$$\begin{aligned} \mathbb{E}[(\widehat{Z}^{(k)}_{\beta_N, N}(0,0 \mid N,z))^2] & \le Ck^2 \sum_{j=0}^{k} \Big((\sigma_N^2 R_N)^{k-j} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N, N}(0,0,N,\star))^2] \\ & \qquad + (\sigma_N^2 R_N)^{j} \mathbb{E}[(\widehat{Z}^{(k-j)}_{\beta_N, N}(0,\star,N,z))^2] \Big) \qquad (4.37) \\ & = 2Ck^2 \sum_{j=0}^{k} (\sigma_N^2 R_N)^{k-j} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N, N}(0,0,N,\star))^2], \end{aligned}$$

where we used $\mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N, N}(0,0,N,\star))^2] = \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N, N}(0,\star,N,z))^2]$ in the last step. Finally, using (4.37) and orthogonality of the polynomial chaos com-

ponents, we have the upper bound

$$\mathbb{E}[(\widehat{Z}_{\beta_N,N}(0,0 \mid N,z))^2] \leq 2C \sum_{k=1}^{N} k^2 \sum_{j=1}^{k} (\sigma_N^2 R_N)^{k-j} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2].$$

Switching the order of the two sums and performing an index shift, we have

$$\begin{aligned} &\sum_{k=1}^{N} k^2 \sum_{j=1}^{k} (\sigma_N^2 R_N)^{k-j} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2] \\ &= \sum_{j=1}^{N} \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2] \sum_{k=0}^{N-j} (k+j)^2 \, (\sigma_N^2 R_N)^k \,, \end{aligned}$$

which overall yields

$$\mathbb{E}[(\widehat{Z}_{\beta_N,N}(0,0 \mid N,z))^2] \leq C \sum_{j=1}^{N} j^2 \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2]. \tag{4.38}$$

Here we used the fact that, as $N \to \infty$,

$$\sum_{k=0}^{N-j} (k+j)^2 \, (\sigma_N^2 R_N)^k \leq j^2 \Big(1 + 4 \sum_{k=1}^{\infty} k^2 \, (\sigma_N^2 R_N)^k\Big) \to j^2 \Big(1 + 4 \sum_{k=1}^{\infty} k^2 \, \widehat{\beta}^{2k}\Big),$$

for all $j \geq 1$. Because the series converges, recall that $\lim_{N\to\infty} \sigma_N^2 R_N = \widehat{\beta}^2 < 1$, the series on the left-hand side is uniformly bounded in N and can be absorbed in the constant C.

In Lemma 4.3.1 below, we will deduce that

$$\mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2] \leq C \, j^2 \, (\sigma_N^2 R_N)^{\frac{j}{2}} \, a_N \sim C \, j^2 \, \widehat{\beta}^j \, a_N \,, \tag{4.39}$$

which is a direct implication of the estimates in [CSZ20, Section 3.4]. Hence, using (4.38) and (4.39), we have

$$\mathbb{E}[(\widehat{Z}_{\beta_N,N}(0,0 \mid N,z))^2] \leqslant C \sum_{j=1}^{N} j^2 \mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2] \leq C \, a_N \sum_{j=1}^{N} j^4 \widehat{\beta}^j \,,$$

which vanishes in the large N limit, since $a_N \to 0$. This completes the proof. □

It is worth pointing out, that the estimates used in (4.30) and (4.34) hold because $z \in \sqrt{N}B(x,r)$ is macroscopically bounded, which prevents the

divergence of the term $(N\, q_N(z))^{-1}$. For this reason, Proposition 4.2.1 only holds locally uniformly.

We are now ready to prove the section's main result, namely that the point-to-point partition function can be factorised into a product of two point-to-plane partition functions.

Proof of Proposition 4.2.1. First, we ignore the $\sup_{z\in\sqrt{N}B(x,r),\, q_N(z)>0}$ and only consider a single, fixed $z \in \mathbf{Z}^2$ such that $q_N(z) > 0$. As already mentioned in Remark 4.1.2, $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ can be approximated arbitrary well in $L^2(\mathbb{P})$ by $Z_{\beta_N,N}(0,0 \mid N,z)$ in the large N limit, which is why we can restrict ourselves to the latter in this proof.

We fix $0 < s^+ < t^- < 1$. The triangle inequality yields the estimate

$$
\begin{aligned}
&\left\| Z_{\beta_N,N}(0,0 \mid N,z) - Z_{\beta_N,N}(0,0,s^+N,\star) Z_{\beta_N,N}(t^-N,\star,N,z) \right\|_{L^2(\mathbb{P})} \\
&= \Big\| \big(\widehat{Z}_{\beta_N,N}(0,0 \mid N,z) + Z^A_{\beta_N,N}(0,0 \mid N,z)\big) \\
&\qquad\qquad - Z_{\beta_N,N}(0,0,s^+N,\star) Z_{\beta_N,N}(t^-N,\star,N,z) \Big\|_{L^2(\mathbb{P})} \\
&\le \left\| \widehat{Z}_{\beta_N,N}(0,0 \mid N,z) \right\|_{L^2(\mathbb{P})} \\
&\quad + \left\| Z^A_{\beta_N,N}(0,0 \mid N,z) - Z^A_{\beta_N,N}(0,0,s^+N,\star) Z^A_{\beta_N,N}(t^-N,\star,N,z) \right\|_{L^2(\mathbb{P})} \\
&\quad + \Big\| Z^A_{\beta_N,N}(0,0,s^+N,\star) Z^A_{\beta_N,N}(t^-N,\star,N,z) \\
&\qquad\qquad - Z_{\beta_N,N}(0,0,s^+N,\star) Z_{\beta_N,N}(t^-N,\star,N,z) \Big\|_{L^2(\mathbb{P})}.
\end{aligned}
$$

The first and second term vanish due to Lemma 4.2.8 and Lemma 4.2.6, respectively. Intuitively, the third term is negligible due to the results in [CSZ17b], which state that only samples inside the boxes $A^{\pm}_N$ contribute to the L^2-limit of the point-to-plane partition functions, see also Lemma 4.2.4. We begin by estimating

$$
\begin{aligned}
&\Big\| Z^A_{\beta_N,N}(0,0,s^+N,\star) Z^A_{\beta_N,N}(t^-N,\star,N,z) \\
&\qquad\qquad - Z_{\beta_N,N}(0,0,s^+N,\star) Z_{\beta_N,N}(t^-N,\star,N,z) \Big\|_{L^2(\mathbb{P})} \\
&\le \Big\| \Big(Z^A_{\beta_N,N}(0,0,s^+N,\star) - Z_{\beta_N,N}(0,0,s^+N,\star) \Big)\, Z^A_{\beta_N,N}(t^-N,\star,N,z) \Big\|_{L^2(\mathbb{P})} \\
&\quad + \Big\| Z_{\beta_N,N}(0,0,s^+N,\star) \Big(Z^A_{\beta_N,N}(t^-N,\star,N,z) - Z_{\beta_N,N}(t^-N,\star,N,z) \Big) \Big\|_{L^2(\mathbb{P})}
\end{aligned}
$$

$$= \left\| \widehat{Z}_{\beta_N,N}(0,0,s^+N,\star) \right\|_{L^2(\mathbb{P})} \left\| Z^A_{\beta_N,N}(t^-N,\star,N,z) \right\|_{L^2(\mathbb{P})} + \left\| Z_{\beta_N,N}(0,0,s^+N,\star) \right\|_{L^2(\mathbb{P})} \left\| \widehat{Z}_{\beta_N,N}(t^-N,\star,N,z) \right\|_{L^2(\mathbb{P})},$$

where we used independence of the disorder on the disjoint time intervals $[0,s^+]$ and $[t^-,1]$ in the last step. The second moments of $\widehat{Z}_{\beta_N,N}$ converge to zero, see Lemma 4.2.4, whereas the L^2-norms of $Z^A_{\beta_N,N}$ and $Z_{\beta_N,N}$ are uniformly bounded in N. Note that all statements in the proof hold uniformly on balls $\sqrt{N}B(x,r)$, therefore the local uniformity in the statement follows. This completes the proof. □

Before we move on to the next section, we state and prove the analogous result of Lemma 4.2.4(i) for the point-to-point partition function:

Corollary 4.2.9. *Let $x \in \mathbb{R}^2$ and $r > 0$ be arbitrary. For every $\widehat{\beta} < 1$ there exists a $\delta = \delta(\widehat{\beta}) > 0$ and $C > 0$ such that*

(i) $\sup_{N\in\mathbb{N}} \sup_{z\in\sqrt{N}B(x,r)} \mathbb{E}[Z_{\beta_N,N}(0,0 \mid N,z)^{2+\delta}] \leq C$,

(ii) $\sup_{N\in\mathbb{N}} \sup_{z\in\sqrt{N}B(x,r)} \mathbb{E}[\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)^{2+\delta}] \leq C$.

Proof. Let $x \in \mathbb{R}^2$ and $r > 0$. First, we prove statement (i) for the point-to-point partition function $Z_{\beta_N,N}(0,0 \mid N,z)$, ignoring the endpoint-disorder. From the proof of Proposition 4.2.1 and Lemma 4.2.4, we know that $\sup_{z\in\sqrt{N}B(x,r)} \mathbb{E}[Z_{\beta_N,N}(0,0 \mid N,z)^2]$ is uniformly bounded in N. Moreover, the point-to-point partition function still has the form of a multi-linear polynomial. In order to lift the boundedness to the $(2+\delta)$-moment for $\delta > 0$ sufficiently small, we apply the hypercontractivity property for polynomial chaoses, see for example [CSZ20, Appendix B]. It is easy to check that the necessary conditions are still satisfied. Thus, we may estimate the p-th moment, $p \geq 2$, by

$$\mathbb{E}[Z_{\beta_N,N}(0,0 \mid N,z)^p] \leq \Big(\sum_{k=0}^{N} c_p^{2k} \mathbb{E}[Z^{(k)}_{\beta_N,N}(0,0 \mid N,z)^2] \Big)^{\frac{p}{2}}. \tag{4.40}$$

Here, c_p is a constant which only depends on p and the law of the noise $\eta^{(N)}$. In [CSZ20, Theorem B.1] it was additionally shown that $\lim_{p\downarrow 2} c_p = 1$. In order to prove that the sum is uniformly bounded in N, we split it into the

two (by now well known) parts:

$$\begin{aligned}&\sum_{k=0}^{N} c_p^{2k}\mathbb{E}[Z^{(k)}_{\beta_N,N}(0,0 \mid N,z)^2] \\ &= \sum_{k=0}^{N} c_p^{2k}\mathbb{E}[Z^{A,(k)}_{\beta_N,N}(0,0 \mid N,z)^2] + \sum_{k=0}^{N} c_p^{2k}\mathbb{E}[\widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z)^2],\end{aligned} \tag{4.41}$$

due to orthogonality. The second term on the right-hand side can be upper bounded using (4.37) and Lemma 4.3.1

$$\begin{aligned}\sum_{k=0}^{N} c_p^{2k}\mathbb{E}[\widehat{Z}^{(k)}_{\beta_N,N}(0,0 \mid N,z)^2] &\le C\sum_{k=0}^{N} c_p^{2k}k^2\sum_{j=0}^{k}(\sigma_N^2 R_N)^{k-j}\mathbb{E}[(\widehat{Z}^{(j)}_{\beta_N,N}(0,0,N,\star))^2] \\ &\le a_N\, C\sum_{k=0}^{N} c_p^{2k}k^2\sum_{j=0}^{k}(\sigma_N^2 R_N)^{k-j}(\sigma_N^2 R_N)^{j/2}j^2 \\ &\le a_N\, C\sum_{k=0}^{N} k^5(c_p^4\sigma_N^2 R_N)^{k/2},\end{aligned}$$

and vanishes because $c_p^2\widehat{\beta} < 1$ for $p \ge 2$ small enough, where we used the fact that $c_p^2\sigma_N\sqrt{R_N} \sim c_p^2\widehat{\beta}$. On the other hand, the first term on the right-hand side of (4.41) can be estimated in terms of

$$\begin{aligned}&\sum_{k=0}^{N} c_p^{2k}\mathbb{E}[Z^{A,(k)}_{\beta_N,N}(0,0 \mid N,z)^2] \\ &\le \Bigg(\sup_{\substack{|N-n|<2N^{1-a_N}\\ |z-y|<2N^{1/2-a_N/4}}} \frac{q_n(y)}{q_N(z)}\Bigg) \\ &\qquad\qquad \times \sum_{k=0}^{N} c_p^{2k}\sum_{j=0}^{k}\mathbb{E}[Z^{A,(j)}_{\beta_N,N}(0,0,N,\star)^2]\mathbb{E}[Z^{A,(k-j)}_{\beta_N,N}(0,\star,N,z)^2] \\ &\le \Bigg(\sup_{\substack{|N-n|<2N^{1-a_N}\\ |z-y|<2N^{1/2-a_N/4}}} \frac{q_n(y)}{q_N(z)}\Bigg)\sum_{k=0}^{N}(k+1)c_p^{2k}(\sigma_N^2 R_N)^k,\end{aligned}$$

where we performed a similar estimate as in the proof of Lemma 4.2.6 in the first inequality, recall (4.25). In the second step, we simply used the fact that $\mathbb{E}[Z^{A,(j)}_{\beta_N,N}(0,0,N,\star)^2] \le (\sigma_N^2 R_N)^j$ (and similarly for the plane-to-point partition function). Again for $p \ge 2$ small enough, c_p can be absorbed into $\widehat{\beta}$, whereas Lemma 4.2.7(i) implies that the supremum in front of the sum converges to one. Overall, this yields that (4.40) is uniformly bounded in N.

The estimate for $\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)$ follows now from (i). Recall that

$$\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z) = e^{\beta_N \omega_{N,z} - \lambda(\beta_N)} Z_{\beta_N,N}(0,0 \mid N,z),$$

where we included again the disorder at the endpoint separately. We then have

$$\mathbb{E}[\mathfrak{Z}_{\beta_N,N}(0,0 \mid N,z)^p] = \mathbb{E}[(e^{\beta_N \omega_{N,z} - \lambda(\beta_N)})^p] \mathbb{E}[Z_{\beta_N,N}(0,0 \mid N,z)^p],$$

using the independence property of the disorder ω. Choosing now $p < 2 + \delta$, with $\delta = \delta(\widehat{\beta})$ from part (i), the second term is uniformly bounded. It remains to estimate $\mathbb{E}[(e^{\beta_N \omega_{N,z} - \lambda(\beta_N)})^p] = e^{\lambda(p\beta_N) - p\lambda(\beta_N)}$. By Taylor expansion, we can write $\lambda(p\beta_N) - p\lambda(\beta_N) \sim \frac{1}{2} p(p-1)\beta_N^2$, which yields boundedness of the exponential. This concludes the proof of (ii). □

4.3 Decay of remainders in polynomial chaoses

We still owe the reader a rigorous justification for the exponential decay of second moments of $\widehat{Z}^{(k)}_{\beta_N,N}$, which we use in the proofs of Lemma 4.2.8 and Corollary 4.2.9.

Lemma 4.3.1. *For any $k \in \mathbf{N}$, we have*

$$\mathbb{E}\left[\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)^2\right] \leq C\, k^2 \left(\sigma_N^2 R_N\right)^{\frac{k}{2}} a_N\,,$$

where C is a constant (independent of k and N) and $a_N = (\log N)^{\gamma-1}$, $\gamma \in (0,1)$, as in the definition of the sets $A_N^{\pm}$.

The proof of this statement can be found in [CSZ20, Section 3.4]. The original proof makes use of more precise estimates to show that $\mathbb{E}[\widehat{Z}_{\beta_N,N}(0,0,N,\star)^2]$ decays like a_N. One can follow the same steps using less sophisticated estimates to get an uniform bound on $\mathbb{E}[\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)^2]$ in terms of a_N instead, which yields the same qualitative bound. We sketch the argument for the sake of completeness:

Considering $\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)$ for some $k \leq N$, there is at least one sample

(n_j, z_j) outside the box $A_N^+(0,0)$. Thus,

$$\mathbb{E}\left[\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)^2\right] \\ \leq \sigma_N^{2k} \sum_{\substack{1\leq l_1,\dots,l_k\leq N \\ z_1,\dots,z_k\in\mathbf{Z}^2}} \sum_{j=1}^{k} \left(\mathbb{1}_{l_j>\frac{1}{k}N^{1-a_N}} + \mathbb{1}_{l_j\leq\frac{1}{k}N^{1-a_N},\,|z_j|\geq\frac{1}{k}N^{1/2-a_N/4}}\right) \prod_{i=1}^{k} q^2_{l_i}(z_i)\,,$$

where we extended the range of time-differences $l_i = n_i - n_{i-1}$ to all of $\{1,\dots,N\}$. Using once more the identity $\sum_{l=1}^{N}\sum_{z\in\mathbf{Z}^2} q_l^2(z) = R_N$, we have

$$\mathbb{E}\left[\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)^2\right] \\ \leq k\,\sigma_N^{2k} R_N^{k-1} \sum_{\substack{1\leq l\leq N \\ z\in\mathbf{Z}^2}} \left(\mathbb{1}_{l>\frac{1}{k}N^{1-a_N}} + \mathbb{1}_{l\leq\frac{1}{k}N^{1-a_N},|z|\geq\frac{1}{k}N^{1/2-a_N/4}}\right) q_l^2(z)\,. \tag{4.42}$$

Now, the contribution of the two indicator functions can be considered separately. We follow the exact same steps as in Section 3.4 of [CSZ20]:

- For the contribution of large time-jumps, we have

$$\frac{1}{R_N}\sum_{\frac{1}{k}N^{1-a_N}\leq l\leq N} q_{2l}(0) \leq C\frac{1}{R_N}\sum_{\frac{1}{k}N^{1-a_N}\leq l\leq N}\frac{1}{l} \\ \leq C'\,\frac{a_N\log N + \log k}{\log N} \leq 2C'\,k\,a_N\,,$$

  where we used additionally the crude estimates $\frac{\log k}{\log N} \leq k\,a_N$ in the last inequality.

- On the other hand, the contribution of the second term is upper bounded by

$$\frac{1}{R_N}\sum_{1\leq l\leq\frac{1}{k}N^{1-a_N}}\sum_{|z|>\frac{1}{k}N^{1/2-a_N/4}} q_l^2(z) \leq Ce^{-\eta\frac{N^{\frac{a_N}{2}}}{k}}\,,$$

  for some uniform $\eta > 0$, using Gaussian estimates for the simple random walk (we omit the details here and instead refer the reader to [CSZ20, Section 3.4]). Then, for N large enough, the term in (4.42) corresponding to the second indicator function, can be upper bounded in terms of

$$C\,k\,(\sigma_N^2 R_N)^k e^{-\eta\frac{N^{\frac{a_N}{2}}}{k}} \leq \begin{cases} C\,k(\sigma_N^2 R_N)^k e^{-\eta N^{\frac{a_N}{4}}}\,, & \text{if } k\leq (N^{a_N/2})^{\frac{1}{2}}\,, \\ C\,k(\sigma_N^2 R_N)^{\frac{k}{2}}(\widehat{\beta}+\delta)^{N^{\frac{a_N}{4}}}\,, & \text{if } k> (N^{a_N/2})^{\frac{1}{2}}\,, \end{cases}$$

where $\delta > 0$ small enough such that $\widehat{\beta} + \delta < 1$. Note that $c^{N^{\frac{a_N}{4}}} = o(a_N)$ for any $c \in (0,1)$, because as $N \to \infty$

$$(\log N)^{1-\gamma} c^{N^{\frac{a_N}{4}}} = (\log N)^{1-\gamma} c^{\exp(\frac{1}{4}(\log N)^\gamma)} \leq (\log N) c^{\frac{1}{4}(\log N)^\gamma} \to 0\,,$$

which includes in particular the case $c = \max\{\widehat{\beta} + \delta, e^{-\eta}\}$.

Adding up all estimates above yields the desired upper bound of $\mathbb{E}[\widehat{Z}^{(k)}_{\beta_N,N}(0,0,N,\star)^2]$ in Lemma 4.3.1.

4.4 The annealed polymer measure

After proving the factorisation of point-to-point partition functions, we can finally start analysing the limiting polymer measure. As dealing with the annealed polymer measures first will substantially simplify the required steps in the quenched case, we define the disorder-averaged measure $\mu_{\beta_N,N}$ on $(C[0,1], \mathcal{F})$, by

$$\mu_{\beta_N,N}(B) := \mathbb{E}[\overline{\pi}_\# \mathbf{P}^\omega_{\beta_N,N}(B)] \qquad \forall B \in \mathcal{F}. \tag{4.43}$$

This section's main result is an invariance principle for the paths of the annealed polymer measure.

Proposition 4.4.1 (Annealed invariance principle)**.** *For $\widehat{\beta} \in (0,1)$, we have*

$$\mu_{\beta_N,N} \xrightarrow{d} \mathbf{P}(\tfrac{1}{\sqrt{2}} W \in \cdot\,), \quad \text{as } N \to \infty,$$

where we recall that $\mathbf{P}$ denotes the Wiener measure on $C[0,1]$.

We begin by showing that the limiting finite-dimensional distributions of $\mu_{\beta_N,N}$ agree with the ones of a Brownian motion. In the section's second part, we prove the required tightness in $C[0,1]$. Together with the identification of finite-dimensional distributions, this yields Proposition 4.4.1.

Instead of determining the finite-dimensional distributions of the interpolated paths, it suffices to work with the corresponding starting point of the interpolation. To see this, recall that under $\overline{\pi}_\# \mathbf{P}^\omega_{\beta_N,N}$ we have

$$X_t = \frac{1}{\sqrt{N}} S_{\lfloor tN \rfloor} + \frac{1}{\sqrt{N}} (tN - \lfloor tN \rfloor)(S_{\lfloor tN \rfloor + 1} - S_{\lfloor tN \rfloor}). \tag{4.44}$$

Because the simple random walk has a finite range transtition kernel, the second term vanishes for N large and we are left with $\frac{1}{\sqrt{N}} S_{\lfloor tN \rfloor}$. Thus, the weak limits of $\frac{1}{\sqrt{N}} S_{\lfloor tN \rfloor}$ and X_t (under $\overline{\pi}_{\#} \mathbf{P}^{\omega}_{\beta_N, N}$) must coincide if they exist. We will assume $tN \in \mathbf{N}$ for the sake of notation.

Let $0 < t_1 < t_2 < \ldots < t_k \leq 1$. It will be convenient to fix a partitioning of the time intervals $(t_{i-1}, t_i]$ and $(t_k, 1]$ using $t_i^{\pm}$'s such that

$$t_{i-1}^{+} < t_i^{-} < t_i < t_i^{+}. \tag{4.45}$$

Now, recalling the representation in (4.9), the idea when proving convergence of the marginal distributions is to replace the first point-to-point partition function using Proposition 4.2.1:

$$\frac{\mathfrak{Z}_{\beta_N, N}(0, 0 \mid t_1 N, z_1)}{Z_{\beta_N, N}(0, 0, N, \star)} \simeq \frac{Z_{\beta_N, N}(0, 0, t_1 N, \star) Z_{\beta_N, N}(0, \star, t_1 N, z_1)}{Z_{\beta_N, N}(0, 0, N, \star)}, \tag{4.46}$$

where "$\simeq$" should be understood as approximation in $L^1(\mathbb{P})$ for large N. Because both the denominator and the first term in the numerator only depend on the disorder in a small neighbourhood around the starting point, they converge to the same limit and should cancel as N diverges. We are left with the remaining point-to-point partition functions, which can be analysed separately due to independence of the disorder on disjoint time intervals.

Having the above approach in mind, we introduce a shorthand notation for the following constellation of terms. For $0 < t_1 < t_2 < \ldots < t_k < 1$, we define

$$\begin{aligned} &Q^{\omega}_{\beta_N, N}\big((t_i, B_i)_{i=1}^{k}\big) \\ &:= \sum_{\substack{z_i \in \sqrt{N} B_i \\ 1 \leq i \leq k}} Z_{\beta_N, N}(t_1^{-} N, \star, t_1 N, z_1) \\ &\qquad \times \Big(\prod_{j=1}^{k-1} Z_{\beta_N, N}(t_j N, z_j, t_j^{+} N, \star) Z_{\beta_N, N}(t_{j+1}^{-} N, \star, t_{j+1} N, z_{j+1}) \Big) \\ &\qquad \times Z_{\beta_N, N}(t_k N, z_k, t_k^{+} N, \star) \prod_{j=1}^{k} q_{(t_j - t_{j-1}) N}(z_j - z_{j-1}). \end{aligned} \tag{4.47}$$

Whenever $t_k = 1$, the partition function $Z_{\beta_N, N}(t_k N, z_k, t_k^{+} N, \star)$ is dropped, since it depends on the disorder outside of $\{0, \ldots, N\} \times \mathbf{Z}^2$.

Lemma 4.4.2. *For $\widehat{\beta} \in (0,1)$ and $0 < t_1 < t_2 < \ldots < t_k \leq 1$, we have*

$$\lim_{N\to\infty} \Big\| \mathbf{P}^{\omega}_{\beta_N,N}\big(\tfrac{1}{\sqrt{N}} S_{t_1 N} \in B_1, \ldots, \tfrac{1}{\sqrt{N}} S_{t_k N} \in B_k\big) - Q^{\omega}_{\beta_N,N}\big((t_i, B_i)_{i=1}^k\big) \Big\|_{L^1(\mathbb{P})} = 0, \tag{4.48}$$

for every choice of bounded measurable sets $(B_i)_{i=1}^k \in (\mathbf{R}^2)^k$ satisfying $\lambda(\partial B_i) = 0$, where λ denotes the Lebesgue measure.[1]

In particular, this implies

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}\big[\mathbf{P}^{\omega}_{\beta_N,N}\big(\tfrac{1}{\sqrt{N}} S_{t_1 N} \in B_1, \ldots, \tfrac{1}{\sqrt{N}} S_{t_k N} \in B_k\big)\big] \\ &\qquad = \mathbf{P}\big(\tfrac{1}{\sqrt{2}} W_{t_1} \in B_1, \ldots, \tfrac{1}{\sqrt{2}} W_{t_k} \in B_k\big). \end{aligned} \tag{4.49}$$

Proof. We only elaborate the steps for the case $k = 2$ and $t_2 = 1$; the general statement follows along the same lines modulo more involved notation. Let $(B_i)_{1\leq i\leq 2}$ be bounded sets with boundary of Lebesgue-measure zero. Throughout the proof, we assume $B(x_i, r_i)$'s to be balls large enough such that they cover the bounded sets B_i. We write $t_0 = 0$ and $z_0 = 0$ for the starting point of the random polymer.

First, we will replace the point-to-point partition function $\mathbf{Z}_{\beta_N,N}(0,0 \mid t_1 N, z_1)$ in (4.9) by its point-to-plane counterparts, using Proposition 4.2.1. Afterwards, we exchange the arising term $Z_{\beta_N,N}(0,0,t_0^+ N, \star)$ with $Z_{\beta_N,N}(0,0,N,\star)$ to cancel the point-to-plane partition function in the denominator. In other words, we want to show that

$$\begin{aligned} \sup_{\substack{z_i \in \sqrt{N} B(x_i,r_i) \\ \text{s.t. } q_{t_1 N}(z_1) > 0}} \Big\| \frac{\mathbf{Z}_{\beta_N,N}(t_1 N, z_1 \mid N, z_2)}{Z_{\beta_N,N}(0,0,N,\star)} \Big(&\mathbf{Z}_{\beta_N,N}(0,0 \mid t_1 N, z_1) \\ &- Z_{\beta_N,N}(0,0,N,\star) Z_{\beta_N,N}(t_1^- N, \star, t_1 N, z_1) \Big) \Big\|_{L^1(\mathbb{P})} \end{aligned} \tag{4.50}$$

vanishes in the large N limits. Afterwards, the remaining point-to-point partition function can be replaced using once more Proposition 4.2.1, which yields (4.48).

Instead of showing the above convergence in (4.50) directly, we divide the statement into two, more manageable, terms by introducing the intermediate term $Z_{\beta_N,N}(0,0,t_0^+ N, \star) Z_{\beta_N,N}(t_1^- N, \star, t_1 N, z_1)$.

[1] We will often call a measurable set B satisfying $\lambda(B) = 0$, a *continuity set*.

- We start with the replacement of the point-to-point partition function following Proposition 4.2.1. First, we apply the Cauchy-Schwartz inequality which yields

$$\begin{aligned}
&\Big\| \frac{\mathcal{Z}_{\beta_N,N}(t_1 N, z_1 \mid N, z_2)}{Z_{\beta_N,N}(0,0,N,\star)} \Big(\mathcal{Z}_{\beta_N,N}(0,0 \mid t_1 N, z_1) \\
&\qquad - Z_{\beta_N,N}(0,0,t_0^+ N,\star) Z_{\beta_N,N}(t_1^- N,\star,t_1 N,z_1) \Big) \Big\|_{L^1(\mathbb{P})} \qquad (4.51) \\
&\leq \|\mathcal{Z}_{\beta_N,N}(0,0 \mid t_1 N, z_1) - Z_{\beta_N,N}(0,0,t_0^+ N,\star) Z_{\beta_N,N}(t_1^- N,\star,t_1 N,z_1)\|_{L^2(\mathbb{P})} \\
&\quad \times \|Z_{\beta_N,N}(0,0,N,\star)^{-1}\|_{L^2(\mathbb{P})} \|\mathcal{Z}_{\beta_N,N}(t_1 N, z_1 \mid N, z_2)\|_{L^2(\mathbb{P})}.
\end{aligned}$$

  Here we used independence of the disorder on the disjoint time intervals $(0, t_1 N]$ and $(t_1 N, N]$, in addition to the fact that partition functions $\mathcal{Z}_{\beta_N,N}(sN, 0 \mid tN, z)$ only depend on the disorder ω in the region $(sN, tN] \times \mathbf{Z}^2$. The last term in (4.51) vanishes uniformly in $z \in \sqrt{N} B(x_1, r_1)$, due to Proposition 4.2.1. The first and second term are uniformly bounded by Lemma 4.2.4 and Corollary 4.2.9, respectively.

- Lastly, we can estimate the remaining norm using again the fact, that we have arbitrary good control of negative moments of $Z_{\beta_N,N}(0,0,N,\star)$. Using

$$\begin{aligned}
&Z_{\beta_N,N}(0,0,t_0^+ N,\star) - Z_{\beta_N,N}(0,0,N,\star) \\
&= \widehat{Z}_{\beta_N,N}(0,0,t_0^+ N,\star) - \widehat{Z}_{\beta_N,N}(0,0,N,\star),
\end{aligned}$$

  the Cauchy-Schwarz inequality yields

$$\begin{aligned}
&\Big\| \frac{Z_{\beta_N,N}(0,0,t_0^+ N,\star) - Z_{\beta_N,N}(0,0,N,\star)}{Z_{\beta_N,N}(0,0,N,\star)} \\
&\qquad \times Z_{\beta_N,N}(t_1^- N,\star,t_1 N,z_1) \mathcal{Z}_{\beta_N,N}(t_1 N, z_1 \mid N, z_2) \Big\|_{L^1(\mathbb{P})} \\
&\leq \Big\| \frac{\widehat{Z}_{\beta_N,N}(0,0,t_0^+ N,\star) - \widehat{Z}_{\beta_N,N}(0,0,N,\star)}{Z_{\beta_N,N}(0,0,N,\star)} \Big\|_{L^2(\mathbb{P})} \\
&\qquad \times \big\| Z_{\beta_N,N}(t_1^- N,\star,t_1 N,z_1) \mathcal{Z}_{\beta_N,N}(t_1 N, z_1 \mid N, z_2) \big\|_{L^2(\mathbb{P})}.
\end{aligned}$$

  The second term on the right-hand side is uniformly bounded by Lemma 4.2.4(ii). The first term, on the other hand, can be estimated us-

ing Hölder's inequality and the trivial estimate $\|a-b\| \leq \|a\| + \|b\|$:

$$\begin{aligned}
&\left\| \frac{\widehat{Z}_{\beta_N,N}(0,0,t_0^+N,\star) - \widehat{Z}_{\beta_N,N}(0,0,N,\star)}{Z_{\beta_N,N}(0,0,N,\star)} \right\|_{L^2(\mathbb{P})} \\
&\leq \|Z_{\beta_N,N}(0,0,N,\star)^{-1}\|_{L^{2+\delta^{-1}}(\mathbb{P})} \\
&\qquad \times (\|\widehat{Z}_{\beta_N,N}(0,0,t_0^+N,\star)\|_{L^{2+\delta}(\mathbb{P})} + \|\widehat{Z}_{\beta_N,N}(0,0,N,\star)\|_{L^{2+\delta}(\mathbb{P})}),
\end{aligned} \tag{4.52}$$

where $\delta > 0$ is choosen sufficently small. Because $\|\widehat{Z}_{\beta_N,N}(0,0,N,\star)\|_{L^{2+\delta}(\mathbb{P})}$ and $\|\widehat{Z}_{\beta_N,N}(0,0,t_0^+N,\star)\|_{L^{2+\delta}(\mathbb{P})}$ converge to zero by Lemma 4.2.4(i), while the first term is uniformly bounded, this concludes convergence to zero of the expression in (4.50).

Having proven that (4.50) vanishes, it is only left to replace $\mathfrak{Z}_{\beta_N,N}(t_1N, z_1 \mid N, z_2)$ by $Z_{\beta_N,N}(t_1N, z_1, t_1^+N, \star)Z_{\beta_N,N}(t_2^-N, \star, N, z_2)$, which holds again by means of Proposition 4.2.1. This concludes the proof of (4.48).

Overall, (4.48) implies

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}[\mathbf{P}^{\omega}_{\beta_N,N}(S_{t_1N} \in \sqrt{N}B_1, S_N \in \sqrt{N}B_2)] \\
&= \lim_{N\to\infty} \mathbb{E}[Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)] \\
&= \lim_{N\to\infty} \sum_{\substack{z_i \in \sqrt{N}B_i \\ 1\leq i\leq 2}} \prod_{j=1}^{2} q_{(t_j-t_{j-1})N}(z_j - z_{j-1}),
\end{aligned}$$

where we used the fact that

$$\mathbb{E}[Z_{\beta_N,N}(t_1^-N, \star, t_1N, z_1)Z_{\beta_N,N}(t_1N, z_1, t_1^+N, \star)Z_{\beta_N,N}(t_2^-N, \star, N, z_2)] = 1\,.$$

It is well known that the simple random walk in $d = 2$ converges in law to Brownian motion with diffusion matrix $\frac{1}{\sqrt{2}}I_2$, see for example [LL10, Theorem 7.6.1]. □

Note that Lemma 4.4.2 is enough to deduce weak convergence of finite-dimensional distributions. In order to lift the convergence result to the annealed polymer measures, we require tightness.

Lemma 4.4.3. *The family of annealed polymer measures $(\mu_{\beta_N,N})_N$ is tight in $\mathcal{M}_1(C[0,1])$.*

Proof. We prove the tightness of $(\mu_{\beta_N,N})_N$ by following [Kal02, Theorem 16.5]

which states that it suffices to show that for every $\varepsilon > 0$

$$\lim_{\delta\to 0}\limsup_{N\to\infty}\mu_{\beta_N,N}(m_\delta(X)\geq\varepsilon)=\lim_{\delta\to 0}\limsup_{N\to\infty}\mathbb{E}[\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(m_\delta(X)\geq\varepsilon)]=0, \tag{4.53}$$

where $m_\delta(\varphi)$ denotes the modulus of continuity on the Wiener space, i.e.

$$m_\delta(\varphi):=\sup_{\substack{0\leq s,t\leq 1\\ |t-s|<\delta}}|\varphi_t-\varphi_s| \qquad \forall\ \varphi\in C[0,1].$$

Next, we use the same trick as stated in [CY06, Lemma 4.2] for random polymers in $d \geq 3$. See also [AKQ14a, Lemma 4.2] for a similar application in the case of the continuum random polymer in $d = 1$. For every $\mathbf{P}_N$-integrable Y, we have

$$\mathbb{E}[Z_{\beta_N,N}(0,0,N,\star)\mathbf{E}^{\omega}_{\beta_N,N}[Y]]=\mathbf{E}_N[Y].$$

Thus, for arbitrary $\varepsilon > 0$, we can write

$$\mathbb{E}[Z_{\beta_N,N}(0,0,N,\star)\,\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(m_\delta(X)\geq\varepsilon)]=\overline{\pi}_{\#}\mathbf{P}_N(m_\delta(X)\geq\varepsilon). \tag{4.54}$$

However, the sequence of rescaled, interpolated simple random walks is known to be tight in the Wiener space. Hence, the right hand side of (4.54) vanishes, when taking first the limit superior $N \to \infty$ and then the limit $\delta \to 0$. Notably, this implies that for any $\lambda > 0$

$$\lambda\mathbb{E}[\mathbb{1}_{Z_{\beta_N,N}(0,0,N,\star)\geq\lambda}\,\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(m_\delta(X)\geq\varepsilon)]\to 0,$$

as $N \to \infty$, $\delta \to 0$. On the other hand,

$$\mathbb{E}[\mathbb{1}_{Z_{\beta_N,N}(0,0,N,\star)<\lambda}\,\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(m_\delta(X)\geq\varepsilon)]\leq\lambda\mathbb{E}[Z_{\beta_N,N}(0,0,N,\star)^{-1}],$$

where we first dropped the inner probability before applying Markov's inequality. Once again, we make use of the fact that

$$\sup_{N\in\mathbf{N}}\mathbb{E}[Z_{\beta_N,N}(0,0,N,\star)^{-1}]\leqslant\widehat{C},$$

for some $\widehat{C} = \widehat{C}_{\widehat{\beta}} > 0$, using Lemma 4.2.4(ii). Overall, we can write

$$\begin{aligned}\mu_{\beta_N,N}(m_\delta(X) \geq \varepsilon) &= \mathbb{E}[\overline{\pi}_\# \mathbf{P}^\omega_{\beta_N,N}(m_\delta(X) \geq \varepsilon)] \\ &\leq \mathbb{E}[\mathbb{1}_{Z_{\beta_N,N}(0,0,N,\star)\geq\lambda}\overline{\pi}_\# \mathbf{P}^\omega_{\beta_N,N}(m_\delta(X) \geq \varepsilon)] + \lambda\widehat{C}\,.\end{aligned}$$

Taking first the limits $N \to \infty$, $\delta \to 0$ and lastly the limit $\lambda \to 0$, finally yields (4.53). Thus, tightness of $(\mu_{\beta_N,N})_N$ follows. □

Combining both Lemma 4.4.2 and Lemma 4.4.3, we can show that annealed polymer paths converge in distribution to the ones of Brownian motion.

Proof of Proposition 4.4.1. We first note that Lemma 4.4.2 implies convergence of the finite-dimensional marginals of the annealed polymer measure. Let $0 < t_1 < \ldots < t_k \leq 1$ be arbitrary. Tightness of the marginals follows from the fact that, for every $\varepsilon > 0$, we can choose bounded continuity sets $B_i \subset \mathbf{R}^2$ such that

$$\mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1} \in B_1, \ldots, \tfrac{1}{\sqrt{2}}W_{t_k} \in B_k) > 1 - \varepsilon\,.$$

Thus, using Lemma 4.4.2, the corresponding polymer marginals satisfy

$$\inf_{N>\overline{N}} \mathbb{E}\big[\mathbf{P}^\omega_{\beta_N,N}\big(\tfrac{1}{\sqrt{N}}S_{\lfloor t_1 N\rfloor} \in B_1, \ldots, \tfrac{1}{\sqrt{N}}S_{\lfloor t_k N\rfloor} \in B_k\big)\big] > 1 - \varepsilon\,,$$

for some $\overline{N} = \overline{N}_\varepsilon \in \mathbf{N}$. This can be lifted to all $N \in \mathbf{N}$ by extending the sets B_i. Therefore, the laws of $\frac{1}{\sqrt{N}}(S_{\lfloor t_1 N\rfloor}, \ldots, S_{\lfloor t_k N\rfloor})$ are tight and converge weakly to the law of $\frac{1}{\sqrt{2}}(W_{t_1}, \ldots, W_{t_k})$, since bounded continuity sets suffice to identify the limiting measure uniquely. The same holds for $(X_{t_1}, \ldots, X_{t_k})$, see (4.44) and the discussion below, thus,

$$\mu_{\beta_N,N}(X_{t_1} \in \cdot, \ldots, X_{t_k} \in \cdot) \xrightarrow{d} \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1} \in \cdot, \ldots, \tfrac{1}{\sqrt{2}}W_{t_k} \in \cdot).$$

Recall now from Lemma 4.4.3 that the annealed measures $(\mu_{\beta_N,N})_N$ are tight. Hence, convergence of the annealed polymer measures is immediate, because weak accumulation points of $(\mu_{\beta_N,N})_N$ are determined by their finite-dimensional distributions, see for example [Bil99, Theorem 7.1]. □

4.5 The quenched polymer measure

After proving the invariance principle for the annealed polymer measures in the previous section, we can now proceed with proving the chapter's main results for the quenched polymer measures. In the first part of the section, we show convergence of finite-dimensional distributions of the rescaled, interpolated polymer paths (Proposition 4.1.1). Afterwards, we prove Proposition 4.5.7, which is a functional central limit theorem for the polymer paths, when tested against individual test functions in $C_b(C[0,1])$. We then explain how the functional central limit theorem can be lifted to an invariance principle (Proposition 4.5.9), using a countable convergence determining class of functions. Lastly, we prove the local limit theorem (Proposition 4.0.2) for the polymer marginals on microscopic scales.

4.5.1 Finite-dimensional distributions

Self-averaging of the random polymer

We begin by showing convergence of the quenched polymer marginals in $L^1(\mathbb{P})$, evaluated on fixed bounded continuity sets B.

Proposition 4.5.1. *For $\widehat{\beta} \in (0,1)$ and $0 < t_1 < \ldots < t_k \leq 1$, we have*

$$\begin{aligned} \big\| \mathbf{P}^{\omega}_{\beta_N,N}\big(\tfrac{1}{\sqrt{N}} S_{t_1 N} \in B_1, \ldots, \tfrac{1}{\sqrt{N}} S_{t_k N} \in B_k\big) \\ - \mathbf{P}\big(\tfrac{1}{\sqrt{2}} W_{t_1} \in B_1, \ldots, \tfrac{1}{\sqrt{2}} W_{t_k} \in B_k\big) \big\|_{L^1(\mathbb{P})} \to 0, \end{aligned} \tag{4.55}$$

for any choice of bounded measurable sets $(B_i)_{i=1}^k \in (\mathbf{R}^2)^k$ satisfying $\lambda(\partial B_i) = 0$, where λ denotes the Lebesgue measure.

Recall, from Lemma 4.4.2, that marginals of $\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}$ can be approximated in terms of $Q^{\omega}_{\beta_N,N}$. To lift Lemma 4.4.2 to the convergence of quenched marginals, it is enough to show convergence of $Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^k)$ to the corresponding transition probabilities of Brownian motion. We will do so first, and state the proof of Proposition 4.5.1 at the end of the section.

In the remainder of this section, we restrict ourselves to the case of $k = 2$ with $t_2 = 1$. The general statement of Proposition 4.5.1 follows along the same lines.

Lemma 4.5.2. *Let $t_1 \in (0,1)$ and $B_1, B_2 \subset \mathbf{R}^2$ be arbitrary bounded continuity sets, then*

$$\lim_{N\to\infty} \|Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)\|^2_{L^2(\mathbb{P})} = \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1} \in B_1, \tfrac{1}{\sqrt{2}}W_1 \in B_2)^2.$$

For the definition of $Q^{\omega}_{\beta_N,N}$ recall (4.47).

We defer the proof of Lemma 4.5.2 until after Lemma 4.5.3. First, we note that the second moment $\|Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)\|^2_{L^2(\mathbb{P})}$ agrees with

$$\begin{aligned}\sum_{\substack{y_i\in\sqrt{N}B_i\\1\le i\le 2}} \sum_{\substack{z_i\in\sqrt{N}B_i\\1\le i\le 2}} &\mathbb{E}[Z_{\beta_N,N}(t_1^- N,\star,t_1 N,y_1)Z_{\beta_N,N}(t_1^- N,\star,t_1 N,z_1)]\\ &\times \mathbb{E}[Z_{\beta_N,N}(t_1 N,y_1,t_1^+ N,\star)Z_{\beta_N,N}(t_1 N,z_1,t_1^+ N,\star)]\\ &\times \mathbb{E}[Z_{\beta_N,N}(t_2^- N,\star,N,y_2)Z_{\beta_N,N}(t_2^- N,\star,N,z_2)]\\ &\times \prod_{j=1}^{2} q_{(t_j-t_{j-1})N}(y_j-y_{j-1})q_{(t_j-t_{j-1})N}(z_j-z_{j-1}),\end{aligned} \tag{4.56}$$

due to independence of the occurring partition functions. Hence, to get a sharp bound and conclude Lemma 4.5.2, we need precise estimates of mixed moments of the form

$$\mathbb{E}[Z_{\beta_N,N}(sN,y,tN,\star)Z_{\beta_N,N}(sN,z,tN,\star)],$$

locally uniformly in space, as N tends to infinity.

Recall from (4.14) that partition functions starting at macroscopically separated points are asymptotically independent. Therefore, one expects a law-of-large-number-like behaviour when ignoring the contribution of starting points y_i and z_i lying "too close" to each other. This idea encourages us to divide the points $(z_i)_{1\le i\le 2}$ into two groups, while keeping $(y_i)_{1\le i\le 2}$ fixed. We either have $|z_i - y_i| > 2N^{1/2-a_N/4}$ for every $1 \le i \le 2$ or there exists an index $1 \le j \le 2$ such that $|z_j - y_j| \le 2N^{1/2-a_N/4}$.

Before stating the full proof of Lemma 4.5.2, we conclude the following covariance estimate of two point-to-plane partition functions.

Lemma 4.5.3. *Let $s,t \in [0,1]$ such that $s < t$, $x \in \mathbf{R}^2$ and $r > 0$. Then*

$$\lim_{N\to\infty} \sup_{\substack{y,z\in\sqrt{N}B(x,r)\\|y-z|>2N^{\frac{1}{2}-\frac{a_N}{4}}}} \mathbb{E}[Z_{\beta_N,N}(sN,y,tN,\star)Z_{\beta_N,N}(sN,z,tN,\star)] = 1.$$

The same statement also holds for plane-to-point partition functions.

Proof. Let $y, z \in \sqrt{N}B(x,r)$ such that $|y-z| > 2N^{\frac{1}{2}-\frac{a_N}{4}}$. First, we expand the expectation in the statement

$$\begin{aligned}
&\mathbb{E}[Z_{\beta_N,N}(sN,y,tN,\star)Z_{\beta_N,N}(sN,z,tN,\star)] \\
&= \mathbb{E}[Z^A_{\beta_N,N}(sN,y,tN,\star)Z^A_{\beta_N,N}(sN,z,tN,\star)] \\
&\quad + \mathbb{E}[\widehat{Z}_{\beta_N,N}(sN,y,tN,\star)Z^A_{\beta_N,N}(sN,z,tN,\star)] \\
&\quad + \mathbb{E}[Z^A_{\beta_N,N}(sN,y,tN,\star)\widehat{Z}_{\beta_N,N}(sN,z,tN,\star)] \\
&\quad + \mathbb{E}[\widehat{Z}_{\beta_N,N}(sN,y,tN,\star)\widehat{Z}_{\beta_N,N}(sN,z,tN,\star)].
\end{aligned}$$

Now, using the Cauchy-Schwarz inequality and Lemma 4.2.4(i), the last three terms on the right-hand side vanish uniformly over $y, z \in \sqrt{N}B(x,r)$ in the large N limit. For the remaining term, we use the fact that $|y-z| > 2N^{\frac{1}{2}-\frac{a_N}{4}}$ implies $A^+_N(sN,y) \cap A^+_N(sN,z) = \emptyset$, whence

$$\begin{aligned}
&\mathbb{E}[Z^A_{\beta_N,N}(sN,y,tN,\star)Z^A_{\beta_N,N}(sN,z,tN,\star)] \\
&= \mathbb{E}[Z^A_{\beta_N,N}(sN,y,tN,\star)]\mathbb{E}[Z^A_{\beta_N,N}(sN,z,tN,\star)] = 1.
\end{aligned}$$

This concludes the proof. □

Proof of Lemma 4.5.2. As in the proof of Lemma 4.4.2, we assume that the sets B_i are covered by open balls $B(x_i,r_i)$. Recall that the second moment of $Q^{\omega}_{\beta_N,N}((t_i,B_i)^2_{i=1})$ is given by the expression (4.56).

First, we show that points y_i and z_i lying "too close" to each other are negligible. Thereafter, we prove the partition functions' self-averaging effect for the remaining points. Overall, this yields a sharp enough estimate for (4.56). Note that in the case where B_i's have empty interior, self-averaging does not take place necessarily. However, in this case we can follow the same steps as in the first bullet below to deduce that (4.55) still holds, since both the random walk and polymer marginals evaluated on such sets converge to zero.

- We begin by analysing points lying "too close", i.e. at least one of the z_i's lies in a $2N^{1/2-a_N/4}$ neighbourhood around y_i. Without loss of generality, we assume that this is the case for z_1. Note that all expectation arising in (4.56) are uniformly bounded in N and $y_i, z_i \in \sqrt{N}B(x_i,r_i)$, say by a constant $C > 0$. This follows by an application of the Cauchy-Schwarz

inequality and Lemma 4.2.4. Thus, when considering the right-hand side of (4.56) with the second sum restricted to those $z_1 \in \sqrt{N}B_1$ satisfying $|y_1 - z_1| \leq 2N^{1/2-a_N/4}$, we can upper bound the expression by

$$C^3 \sum_{\substack{y_i \in \sqrt{N}B_i \\ 1 \leq i \leq 2}} \sum_{\substack{z_1 \in \sqrt{N}B_1 \\ |y_1 - z_1| \leq 2N^{1/2-a_N/4} \\ z_2 \in \sqrt{N}B_2}} \prod_{j=1}^{2} q_{(t_j - t_{j-1})N}(y_j - y_{j-1}) q_{(t_j - t_{j-1})N}(z_j - z_{j-1})$$
$$\leq C^3 \sum_{y_1 \in \sqrt{N}B_1} q_{t_1 N}(y_1) \sum_{\substack{z_1 \in \sqrt{N}B_1 \\ |y_1 - z_1| \leq 2N^{1/2-a_N/4}}} q_{t_1 N}(z_1),$$

where we summed over the probability kernels which don't depend on z_1 or y_1 in the last step. However, $N^{1/2-a_N/4}$ vanishes in the macroscopic limit and therefore the right hand side converges to zero as $N \to \infty$, because

$$\sup_{y_1 \in \sqrt{N}B_1} \sum_{\substack{z_1 \in \sqrt{N}B_1 \\ |y_1 - z_1| \leq 2N^{1/2-a_N/4}}} q_{t_1 N}(z_1)$$
$$\sim \sup_{y_1 \in \sqrt{N}B_1} \frac{1}{N} \sum_{\substack{z_1 \in \sqrt{N}B_1 \\ |y_1 - z_1| \leq 2N^{1/2-a_N/4}}} 2\, p_{\frac{t_1}{2}}\left(\tfrac{z_1}{\sqrt{N}}\right) \to 0,$$

by the local limit theorem for the simple random walk. Consequently, whenever the space points lie "too close" to each other, in the sense that there exists an index j such that $|z_j - y_j| < 2N^{1/2-a_N/4}$, they do not contribute to the limiting marginal distribution of the polymer path.

- It only remains to estimate the second part of the decomposition of (4.56), where we restrict the sum over z_i's such that every z_i has at least distance $2N^{1/2-a_N/4}$ from y_i, i.e.

$$\sum_{\substack{y_i \in \sqrt{N}B_i \\ 1 \leq i \leq 2}} \sum_{\substack{z_i \in \sqrt{N}B_i \\ |y_i - z_i| > 2N^{1/2-a_N/4} \\ 1 \leq i \leq 2}} \mathbb{E}[Z_{\beta_N,N}(t_1^- N, \star, t_1 N, y_1) Z_{\beta_N,N}(t_1^- N, \star, t_1 N, z_1)]$$
$$\times \mathbb{E}[Z_{\beta_N,N}(t_1 N, y_1, t_1^+ N, \star) Z_{\beta_N,N}(t_1 N, z_1, t_1^+ N, \star)] \tag{4.57}$$
$$\times \mathbb{E}[Z_{\beta_N,N}(t_2^- N, \star, N, y_2) Z_{\beta_N,N}(t_2^- N, \star, N, z_2)]$$
$$\times \prod_{j=1}^{2} q_{(t_j - t_{j-1})N}(y_j - y_{j-1}) q_{(t_j - t_{j-1})N}(z_j - z_{j-1}).$$

Using Lemma 4.5.3, each occurring expectation in the expression above converges (locally) uniformly to 1, when taking the large N limit.

Coming back to the term we wanted to estimate originally, we can now write (after taking $\lim_{N\to\infty}$ on both sides)

$$\lim_{N\to\infty} \|Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)\|^2_{L^2(\mathbb{P})} = \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1}\in B_1, \tfrac{1}{\sqrt{2}}W_1\in B_2)^2,$$

where we started with (4.56), neglected the space-points lying "too close" to each other and proved convergence of the remaining sum (4.57) using Lemma 4.5.3. □

Having just proved convergence of the second moment of $Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)$ to its squared mean, it is an immediate consequence of Lemma 4.5.2 that $Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)$ converges to its mean in $L^2(\mathbb{P})$:

Corollary 4.5.4. *Let $t_1\in(0,1)$ and $B_1,B_2\subset\mathbf{R}^2$ be arbitrary bounded continuity sets, then*

$$\lim_{N\to\infty} \left\|Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1}\in B_1, \tfrac{1}{\sqrt{2}}W_1\in B_2)\right\|_{L^2(\mathbb{P})} = 0.$$

Finally, we can summarise the results of above lemmas in the proof of Proposition 4.5.1:

Proof of Proposition 4.5.1. Application of the triangle inequality yields

$$\begin{aligned}\|\mathbf{P}^{\omega}_{\beta_N,N}(S_{t_1N}\in\sqrt{N}B_1, S_N\in\sqrt{N}B_2) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1}\in B_1, \tfrac{1}{\sqrt{2}}W_1\in B_2)\|_{L^1(\mathbb{P})}\\ \le \|\mathbf{P}^{\omega}_{\beta_N,N}(S_{t_1N}\in\sqrt{N}B_1, S_N\in\sqrt{N}B_2) - Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2)\|_{L^1(\mathbb{P})}\\ + \|Q^{\omega}_{\beta_N,N}((t_i,B_i)_{i=1}^2) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_{t_1}\in B_1, \tfrac{1}{\sqrt{2}}W_1\in B_2)\|_{L^1(\mathbb{P})}.\end{aligned}$$

The first term on the right-hand side vanishes by Lemma 4.4.2, whereas the second term vanishes due to Corollary 4.5.4. □

Using standard estimates, one can show that the transition probabilities of the interpolated polymer path are well approximated by the corner points of the discrete path. Thus, Proposition 4.5.1 also holds for the rescaled and interpolated polymer paths under $\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}$:

Corollary 4.5.5. *For $\widehat{\beta} \in (0,1)$ and $0 < t_1 < \ldots < t_k \leq 1$, we have*

$$\lim_{N\to\infty} \big\| \overline{\pi}_{\#} \mathbf{P}^{\omega}_{\beta_N,N}\big(X_{t_1} \in B_1, \ldots, X_{t_k} \in B_k\big) \\ - \mathbf{P}\big(\tfrac{1}{\sqrt{2}} W_{t_1} \in B_1, \ldots, \tfrac{1}{\sqrt{2}} W_{t_k} \in B_k\big) \big\|_{L^1(\mathbb{P})} = 0,$$

for every choice of bounded measurable sets $(B_i)_{i=1}^k \in (\mathbf{R}^2)^k$ satisfying $\lambda(\partial B_i) = 0$.

Tightness and uniqueness of the limit

In Corollary 4.5.5 we showed convergence of polymer marginals evaluated on bounded, factorised, continuity sets in $L^1(\mathbb{P})$. This can be lifted to unbounded measurable sets $U \subset (\mathbf{R}^2)^k$ satisfying the same properties. However, we want to stress that this does not imply weak convergence of the polymer marginals yet, since exceptional points of the disorder can depend on the choice of sets U. Nevertheless, we are able to show weak convergence of quenched finite-dimensional distributions in probability, since probability measures on $(\mathbf{R}^2)^k$ are uniquely identified by evaluation on a countable family of sets.

We begin by recalling a standard result: A sequence of random variables on a metric space converges in probability if and only if every subsequence has a further subsequence, which converges almost surely, see for example [Kal02, Lemma 4.2]. Thus, in order to show convergence in probability of the marginal distributions, it suffices to prove tightness along sufficiently many subsequences and identify the limit points using a π-system. This step is motivated by the recent article [Jun23], where Junk showed convergence of the polymer-endpoint distribution in bond disorder for $d \geq 3$, using a very similar approach.

For the π-system, on which we will identify the limiting finite-dimensional distributions, we choose half-open cylinders:

$$\mathcal{E}^k := \{[a_1, b_1) \times \cdots \times [a_{2k}, b_{2k}) \subset (\mathbf{R}^2)^k \,:\, a_i, b_i \in \mathbb{Q} \text{ and } a_i < b_i\,,\, 1 \leq i \leq 2k\},$$

which generate the Borel-sigma-algebra on $(\mathbf{R}^2)^k$. Note that $\mathcal{E}^k$ has countably many elements, and let $\{E_i\}_{i=1}^{\infty}$ be an arbitrary enumeration of them. We omitted the dependency of E_i's on k for the sake of a lighter notation.

We start by showing that there exist sufficiently many subsequences along which finite-dimensional distributions (evaluated on sets in $\mathcal{E}^k$) converge almost surely.

Lemma 4.5.6. *Let $0 \leq t_1 < \ldots < t_k \leq 1$, then for every sequence $(N_j)_{j\in\mathbf{N}}$ in $\mathbf{N}$ there exists a subsequence $(N_{j_m})_{m\in\mathbf{N}}$ and $\Omega_1 = \Omega_1((N_j)_{j\in\mathbf{N}}, (t_i)_{1\leq i\leq k}) \subset \Omega$ with $\mathbb{P}(\Omega_1) = 1$, such that for every $\omega \in \Omega_1$*

$$\lim_{m\to\infty} \overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}((X_{t_1},\ldots,X_{t_k}) \in E) = \mathbf{P}(\tfrac{1}{\sqrt{2}}(W_{t_1},\ldots,W_{t_k}) \in E) \quad \forall E \in \mathcal{E}^k.$$

Proof. Let $(N_j)_{j\in\mathbf{N}}$ be an arbitrary sequence in $\mathbf{N}$. We prove the lemma only for a single marginal $t \in [0,1]$, the multi-marginal case follows along the same lines.

First, Corollary 4.5.5 implies that for every $\varepsilon > 0$ and $E_i \in \mathcal{E}$

$$\lim_{j\to\infty} \mathbb{P}(|\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_j},N_j}(X_t \in E_i) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_t \in E_i)| > \varepsilon) = 0.$$

In particular, for every $i, m \in \mathbf{N}$ there exists a $M_{i,m} \in \mathbf{N}$ such that

$$\mathbb{P}(|\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_j},N_j}(X_t \in E_i) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_t \in E_i)| > \tfrac{1}{m}) \leq m^{-1}2^{-m} \quad \forall j \geq M_{i,m}.$$

We define a subsequence of $(N_j)_j$ using $j_m := j_{m-1} \vee \max_{1\leq i\leq m} M_{i,m}$, then for every $m \in \mathbf{N}$

$$\mathbb{P}(\exists i \leq m \text{ with } |\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}(X_t \in E_i) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_t \in E_i)| > \tfrac{1}{m}) \leq 2^{-m},$$

which is summable in m. The Borel-Cantelli lemma then yields

$$\mathbb{P}\big(\sup_{i\leq m} |\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}(X_t \in E_i) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_t \in E_i)| > \tfrac{1}{m} \text{ infinitely often}\big) = 0,$$

which implies that

$$\Omega_1 := \big\{\omega \in \Omega : \lim_{m\to\infty} \sup_{i\leq m} |\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}(X_t \in E_i) - \mathbf{P}(\tfrac{1}{\sqrt{2}}W_t \in E_i)| = 0\big\}$$

has full mass, i.e. $\mathbb{P}(\Omega_1) = 1$. This concludes the proof. □

Finally, we can prove convergence of the quenched polymer marginals.

Proof of Proposition 4.1.1. We fix $0 \leq t_1 < \ldots < t_k \leq 1$ and let $(N_j)_{j\in\mathbf{N}}$ be a sequence in $\mathbf{N}$. In Lemma 4.5.6 we proved the existence of a subsequence $(N_{j_m})_{m\in\mathbf{N}}$ and disorders Ω_1, with $\mathbb{P}(\Omega_1) = 1$, such that

$$\lim_{m\to\infty} \overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}((X_{t_1},\ldots,X_{t_k}) \in E) = \mathbf{P}(\tfrac{1}{\sqrt{2}}(W_{t_1},\ldots,W_{t_k}) \in E)\,, \qquad (4.58)$$

for all $\omega \in \Omega_1$ and $E \in \mathcal{E}^k$. Tightness of the sequence

$$\left(\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_{N_{j_m}},N_{j_m}}((X_{t_1},\dots,X_{t_k}) \in \cdot)\right)_{m\in\mathbf{N}}$$

is an immediate consequence of (4.58), see the proof of Proposition 4.4.1, and the limiting probability measure is uniquely determined by the π-system $\mathcal{E}^k$.

Overall, we showed that for every sequence $(N_j)_{j\in\mathbf{N}}$ in $\mathbf{N}$ there exists a subsequence $(N_{j_m})_{m\in\mathbf{N}}$, along which the finite-dimensional distributions converge almost surely to the ones of Brownian motion with diffusion matrix $\frac{1}{\sqrt{2}}I_2$. This is equivalent to weak convergence of the polymer marginals in $\mathbb{P}$-probability. □

4.5.2 An invariance principle for polymer paths

We are finally ready to prove Theorem 4.0.1. Using tightness of the annealed polymer paths (Lemma 4.4.3) and convergence of the finite-dimensional distributions (Proposition 4.1.1), we can prove the desired result. The steps resemble very much the ones when proving the invariance principle for the simple random walk. However, due to the double randomness of paths and the environment, cf. Remark 4.1.4, we require an additional argument to conclude the full invariance principle.

We begin by proving a functional central limit theorem:

Proposition 4.5.7. *Let $\widehat{\beta} \in (0,1)$ and β_N as in (4.3). Then for every $F \in C_b(C[0,1])$*

$$\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)] \to \mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)], \qquad \text{as } N\to\infty, \quad \text{in } \mathbb{P}\text{-probability},$$

where $\mathbf{E}$ is the expectation with respect to the Wiener measure on $C[0,1]$.

Remark 4.5.8. *Convergence in the functional central limit theorem above also holds in $L^1(\mathbb{P})$, since the random variables $(\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)])_N$ are uniformly bounded by $\|F\|_\infty$. In particular, this implies convergence of expectations which is equivalent to the annealed invariance principle, cf. Proposition 4.4.1.*

Proof. Let $F \in C_b(C[0,1])$. The statement in the proposition is equivalent to

$$\lim_{N\to\infty} \mathbb{P}(|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)] - \mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)]| > \varepsilon) = 0 \quad \forall \varepsilon > 0. \tag{4.59}$$

In the following, we fix $\varepsilon > 0$ and choose $\delta > 0$ arbitrary. Let then $K = K(\delta, \varepsilon, F) \subset C[0,1]$ be a compact set such that

$$\mathbf{P}(\tfrac{1}{\sqrt{2}} W \in K^c) \leq \frac{\varepsilon}{8\|F\|_\infty} \tag{4.60}$$

and

$$\sup_{N\in\mathbf{N}} \mu_{\beta_N,N}(K^c) = \sup_{N\in\mathbf{N}} \mathbb{E}[\overline{\pi}_\# \mathbf{P}_{\beta_N,N}(K^c)] \leq \frac{\delta\,\varepsilon}{8\|F\|_\infty}. \tag{4.61}$$

A set K with these properties exists due to tightness of the Wiener measure and tightness of the annealed polymer measure, see Lemma 4.4.3.

Throughout the remainder of the proof, we will denote by $\Pi_{t_1,\ldots,t_k} : C[0,1] \mapsto (\mathbf{R}^2)^k$ the projection of a path onto previously chosen coordinates $0 \leq t_1 < \ldots < t_k \leq 1$, i.e.

$$\Pi_{t_1,\ldots,t_k}(X) = (X_{t_1}, \ldots, X_{t_k}) \in (\mathbf{R}^2)^k.$$

Restricting the domain of F to the compact set K, we can approximate F uniformly by cylinder functions on $C(K, \mathbf{R})$, i.e. functions that only depend on finitely many coordinates of the path, using the Stone-Weierstrass theorem [Lax02, Theorem 13.4]. More precisely, there exist $0 \leq t_1 < \ldots < t_k \leq 1$ and a continuous $f : \Pi_{t_1,\ldots,t_k}(K) \mapsto \mathbf{R}$ such that $\|f \circ \Pi_{t_1,\ldots,t_k}\|_{K,\infty} \leq \|F\|_\infty$ and

$$\|F - f \circ \Pi_{t_1,\ldots,t_k}\|_{K,\infty} := \sup_{\varphi\in K} |F(\varphi) - f(\varphi_{t_1}, \ldots, \varphi_{t_k})| < \frac{\varepsilon}{4}. \tag{4.62}$$

Using the Tietze extension theorem [Rud86, Theorem 20.4], we can extend f continuously from $\Pi_{t_1,\ldots,t_k}(K)$ to $(\mathbf{R}^2)^k$. At the same time, this yields a continuous extension of $f \circ \Pi_{t_1,\ldots,t_k}$ on $C[0,1]$. The extension can be chosen in such a way that $\|f \circ \Pi_{t_1,\ldots,t_k}\|_\infty \leq \|f \circ \Pi_{t_1,\ldots,t_k}\|_{K,\infty} \leq \|F\|_\infty$. Estimating now $(F - f \circ \Pi_{t_1,\ldots,t_k})$ on K and K^c respectively, after applying the triangle inequality, yields

$$\begin{aligned}
&|\overline{\pi}_\# \mathbf{E}^\omega_{\beta_N,N}[(\mathbb{1}_K + \mathbb{1}_{K^c})(F - f \circ \Pi_{t_1,\ldots,t_k})(X)] \\
&\qquad - \mathbf{E}[(\mathbb{1}_K + \mathbb{1}_{K^c})(F - f \circ \Pi_{t_1,\ldots,t_k})(\tfrac{1}{\sqrt{2}} W)]| \\
&\leq 2\|F - f \circ \Pi_{t_1,\ldots,t_k}\|_{K,\infty} + 2\|F\|_\infty \big|\overline{\pi}_\# \mathbf{P}^\omega_{\beta_N,N}(K^c) + \mathbf{P}(\tfrac{1}{\sqrt{2}} W \in K^c)\big|.
\end{aligned}$$

Thus, due to (4.62) and the choice of K in (4.60), the previous display implies

$$\begin{aligned}&|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[(F-f\circ\Pi_{t_1,\dots,t_k})(X)]-\mathbf{E}[(F-f\circ\Pi_{t_1,\dots,t_k})(\tfrac{1}{\sqrt{2}}W)]|\\&\leq\frac{3\varepsilon}{4}+2\|F\|_{\infty}\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(K^c).\end{aligned}\tag{4.63}$$

All together, we can upper bound the term of interest

$$\begin{aligned}&|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)]-\mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)]|\\&\leq|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[f(X_{t_1},\dots,X_{t_k})]-\mathbf{E}[f(\tfrac{1}{\sqrt{2}}W_{t_1},\dots,\tfrac{1}{\sqrt{2}}W_{t_k})]|\\&\quad+|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[(F-f\circ\Pi_{t_1,\dots,t_k})(X)]-\mathbf{E}[(F-f\circ\Pi_{t_1,\dots,t_k})(\tfrac{1}{\sqrt{2}}W)]|,\end{aligned}$$

which implies that

$$\begin{aligned}&\mathbb{P}(|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[F(X)]-\mathbf{E}[F(\tfrac{1}{\sqrt{2}}W)]|>2\varepsilon)\\&\leq\mathbb{P}(|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[f(X_{t_1},\dots,X_{t_k})]-\mathbf{E}[f(\tfrac{1}{\sqrt{2}}W_{t_1},\dots,\tfrac{1}{\sqrt{2}}W_{t_k})]|>\varepsilon)\\&\quad+\mathbb{P}(|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[(F-f\circ\Pi_{t_1,\dots,t_k})(X)]-\mathbf{E}[(F-f\circ\Pi_{t_1,\dots,t_k})(\tfrac{1}{\sqrt{2}}W)]|>\varepsilon).\end{aligned}$$

The first term on the right vanishes, as $N\to\infty$, due to Proposition 4.1.1. On the other hand, the second term can be further upper bounded using (4.63), such that

$$\begin{aligned}&\sup_{N\in\mathbf{N}}\mathbb{P}(|\overline{\pi}_{\#}\mathbf{E}^{\omega}_{\beta_N,N}[(F-f\circ\Pi_{t_1,\dots,t_k})(X)]-\mathbf{E}[(F-f\circ\Pi_{t_1,\dots,t_k})(\tfrac{1}{\sqrt{2}}W)]|>\varepsilon)\\&\leq\sup_{N\in\mathbf{N}}\mathbb{P}\big(\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(K^c)>\tfrac{\varepsilon}{8\|F\|_{\infty}}\big)\leq\frac{8\|F\|_{\infty}}{\varepsilon}\sup_{N\in\mathbf{N}}\mathbb{E}[\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N}(K^c)]\leq\delta,\end{aligned}$$

where we applied Markov's inequality, before using again (4.61). Because $\delta>0$ can be chosen arbitrarily small after taking the large N limit, (4.59) follows. This concludes the proof. □

It is only left to lift the functional central limit theorem (in $\mathbb{P}$-probability) to an invariance principle as stated in Theorem 4.0.1. In fact, we can show more generally the equivalence of functional central limit theorem and invariance principle for random probability measures:

Proposition 4.5.9. *Let (S,d) be a separable, complete metric space and $(\Omega,\mathcal{G},\mathbb{P})$ a probability space. Moreover, let $(\mathrm{P}^{\omega}_N)_{N\in\mathbb{N}}$ be random probability measures and P a deterministic probability measure on $(S,\mathcal{F})$, with $\mathcal{F}$ denoting the Borel-σ-algebra. Then the following two statements are equivalent*

(i) For every $F \in C_b(S)$, $\mathrm{E}_N^\omega[F] \to \mathrm{E}[F]$ *in* $\mathbb{P}$*-probability,*

(ii) $\mathrm{P}_N^\omega \xrightarrow{d} \mathrm{P}$ *in* $\mathbb{P}$*-probability,*

where E_N^ω *and* E *denote the expectations with respect to* P_N^ω *and* P*, respectively. The statement remains true, when replacing convergence in probability with almost sure convergence.*

First, we remind the reader that a set of functions $\mathcal{A} \subset C_b(S)$, where S Polish, is called **weak convergence determining**, if for $\nu_n, \nu \in \mathcal{M}_1(S)$

$$\lim_{n\to\infty} \int F \, \mathrm{d}\nu_n = \int F \, \mathrm{d}\nu \quad \forall F \in \mathcal{A}\,,$$

implies $\nu_n \xrightarrow{d} \nu$.

The following lemma, which is a Corollary of [BK10, Lemma 2], states that we can always find such a family of functions which is countable.

Lemma 4.5.10. *Let* (S, d) *be a separable and complete metric space, then there exists a countable algebra* $\mathcal{A} = \{F_i\}_{i\in\mathbf{N}} \subset C_b(S)$ *that is weak convergence determining.*

Proof. First, note that the family $\overline{\mathcal{A}} \subset C_b(S)$, of uniformly continuous functions with bounded support, strongly separates points. Then [BK10, Lemma 2] yields existence of a countable subfamily $\mathcal{A} := \{F_i\}_{i\in\mathbf{N}} \subset \overline{\mathcal{A}}$ that strongly separates points and remains closed under multiplication. Lastly, applying [EK86, Theorem 3.4.5], which states that an algebra in $C_b(S)$ that strongly separates points is convergence determining, finishes the proof. □

Proof of Proposition 4.5.9. We begin by proving the direction (i) to (ii). Let $\{F_i\}_{i\in\mathbf{N}} \subset C_b(S)$ be a countable family of functions that is weak convergence determining, whose existence is guaranteed by Lemma 4.5.10. For every $i \in \mathbf{N}$, we have

$$\lim_{N\to\infty} \mathrm{E}_N^\omega[F_i] = \mathrm{E}[F_i], \quad \text{in } \mathbb{P}\text{-probability.} \tag{4.64}$$

Now a diagonal argument, as we performed it in the proof of Lemma 4.5.6, yields that for every sequence $(N_j)_j \subset \mathbf{N}$ there exists a further subsequence $(N_{j_m})_m$ and a set $\overline{\Omega} \subset \Omega$, with $\mathbb{P}(\overline{\Omega}) = 1$, such that

$$\lim_{m\to\infty} \overline{\pi}_\# \mathrm{E}_{N_{j_m}}^\omega[F_i] = \mathrm{E}[F_i] \quad \forall i \in \mathbf{N}\,,\ \omega \in \overline{\Omega}\,.$$

Because $\{F_i\}_{i\in\mathbf{N}}$ is weak convergence determining, this implies $\mathrm{P}^{\omega}_{N_{j_m}} \xrightarrow{d} \mathrm{P}$, as $m \to \infty$, for every $\omega \in \overline{\Omega}$. This is equivalent to weak convergence in $\mathbb{P}$-probability as stated in (ii).

The reverse direction, from (ii) to (i), is immediate. Moreover, replacing convergence in $\mathbb{P}$-probability with $\mathbb{P}$-almost-sure convergence, we can simply use that for every F_i there exists a set $\Omega_i \subset \Omega$ with $\mathbb{P}(\Omega_i) = 1$ such that (4.64) holds pointwise for every $\omega \in \Omega_i$. Taking now the countable intersection over all such Ω_i's, we conclude $\mathrm{P}^{\omega}_{N} \xrightarrow{d} \mathrm{P}$ $\mathbb{P}$-almost surely. □

Proof of Theorem 4.0.1. The invariance principle for the polymer measures follows now directly from the functional CLT in Proposition 4.5.7 and Proposition 4.5.9. □

Instead of taking the detour via the functional central limit theorem first, we could have also argued that $\{\mathrm{Law}(\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\beta_N,N})\}_N \subset \mathcal{M}_1(\mathcal{M}_1(C[0,1]))$ is tight. Together with convergence of finite-dimensional distributions, Corollary 4.5.5, this yields a direct argument for the invariance principle. However, we want to put emphasis on the (non-trivial) equivalence of the functional CLT and the invariance principle in the case of random path measures whenever the limit is deterministic.

Lastly, we note that Proposition 4.5.9 also concludes the invariance principle from the functional CLT in higher dimensions [CY06, Theorem 1.2], which was – to the author's best knowledge – not yet mentioned in the literature.

Corollary 4.5.11 (Invariance principle for $d \geq 3$, weak disorder)**.** *Let $\widehat{\beta} \geq 0$ such that weak disorder holds, i.e. $\lim_{N\to\infty} Z_{\widehat{\beta},N}(0,0,N,\star) > 0$, then*

$$\overline{\pi}_{\#}\mathbf{P}^{\omega}_{\widehat{\beta},N} \xrightarrow{d} \mathbf{P}\big(\tfrac{1}{\sqrt{d}}W \in \cdot\,\big), \qquad \text{as } N\to\infty, \quad \text{in } \mathbb{P}\text{-probability},$$

with $\mathbf{P}$ being the d-dimensional Wiener measure. The statement holds in particular for all $\widehat{\beta} \in [0, \beta_c(d))$, with $\beta_c(d)$ defined in (1.8).

4.5.3 Local limit theorem for the polymer marginals

We want to close this chapter by proving the local limit theorem for the marginals of the polymer measure, Proposition 4.0.2. Recall from (4.9), that finite-dimensional distributions of the discrete polymer measure can be writ-

ten as

$$\begin{aligned}&\mathbf{P}^{\omega}_{\beta_N,N}(S_{m_1}=z_1,\dots,S_{m_k}=z_k) \\ &=\frac{1}{Z_{\beta_N,N}(0,0,N,\star)}\prod_{j=1}^{k+1}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\,q_{m_j-m_{j-1}}(z_j-z_{j-1})\,,\end{aligned} \tag{4.65}$$

where $m_0=z_0=0$, $m_{k+1}=N$ and $z_{k+1}=\star$. Together with Proposition 4.2.1 and [CSZ17b, Theorem 2.12], this suffices to deduce Proposition 4.0.2.

Proof of Proposition 4.0.2. By the local limit theorem, we know that $\frac{N}{2}\,q_{n_i-n_{i-1}}(z_i-z_{i-1})$ converges to $p_{(t_i-t_{i-1})/2}(x_i-x_{i-1})$. It is only left to show convergence in distribution of the partition functions in (4.65). We will prove that, as N diverges,

$$\begin{aligned}\Big\|\frac{1}{Z_{\beta_N,N}(0,0,N,\star)}\prod_{j=1}^{k+1}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\\ -\prod_{j=1}^{k}Z_{\beta_N,N}(m_j^-,\star,m_j,z_j)Z_{\beta_N,N}(m_j,z_j,m_j^+,\star)\Big\|_{L^1(\mathbb{P})}\to 0,\end{aligned} \tag{4.66}$$

where $(m_j^{\pm})_{j=1}^k$ are non-negative integers such that

$$0\le m_j^-<m_j<m_j^+<m_{j+1}^-<N\quad\text{and}\quad\lim_{N\to\infty}\frac{|m_j^{\pm}-m_j|}{N}>0.$$

For example, we can choose $m_1^-=0$ and $m_j^{\pm}=\lfloor m_j\pm\frac{1}{3}|m_{j\pm1}-m_j|\rfloor$ for the remaining variables.

We will justify the convergence in (4.66), using a chain of triangle inequalities. First, we note that

$$\begin{aligned}\Big\|\frac{1}{Z_{\beta_N,N}(0,0,N,\star)}\prod_{j=1}^{k+1}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\\ -Z_{\beta_N,N}(0,\star,m_1,z_1)\prod_{j=2}^{k+1}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j)\Big\|_{L^1(\mathbb{P})}\end{aligned}$$

vanishes, as N tends to infinity. The proof follows the same lines as the one of (4.50), in Lemma 4.4.2. Next, we replace the remaining point-to-point partition functions $(\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j))_{j=2}^{k+1}$ with their point-to-plane counterparts by Proposition 4.2.1. For the sake of brevity, we restrict ourselves to $k=2$ for the remainder of this proof, the general case follows using a

telescopic sum argument in the subsequent step. We have

$$\begin{aligned}
&\Big\| Z_{\beta_N,N}(0,\star,m_1,z_1)\Big(\prod_{j=2}^{3}\mathfrak{Z}_{\beta_N,N}(m_{j-1},z_{j-1}\mid m_j,z_j) \\
&\qquad - \big(Z_{\beta_N,N}(m_1,z_1,m_2^-,\star)Z_{\beta_N,N}(m_1^+,\star,m_2,z_2)\big)Z_{\beta_N,N}(m_2,z_2,N,\star)\Big)\Big\|_{L^1(\mathbb{P})} \\
&= \|Z_{\beta_N,N}(0,\star,m_1,z_1)\|_{L^1(\mathbb{P})} \\
&\qquad \times \|\mathfrak{Z}_{\beta_N,N}(m_1,z_1\mid m_2,z_2) - Z_{\beta_N,N}(m_1,z_1,m_2^-,\star)Z_{\beta_N,N}(m_1^+,\star,m_2,z_2)\|_{L^1(\mathbb{P})} \\
&\qquad \times \|Z_{\beta_N,N}(m_2,z_2,N,\star)\|_{L^1(\mathbb{P})}.
\end{aligned}$$

Here, we made use of the disorder's independence on the disjoint time intervals $(0,m_1]$, $(m_1,m_2]$ and $(m_2,N]$, and applied the fact that $\mathfrak{Z}_{\beta_N,N}(m_k,z_k\mid N,\star) = Z_{\beta_N,N}(m_k,z_k,N,\star)$. The middle term on the right-hand side vanishes due to Proposition 4.2.1, whereas the remaining terms are all equal to one. This finally yields (4.66), by adding and subtracting the above introduced intermediate terms and applying the triangle inequality.

The last step consists of determining the limiting distribution of the partition functions in

$$\prod_{j=1}^{k} Z_{\beta_N,N}(m_j^-,\star,m_j,z_j)Z_{\beta_N,N}(m_j,z_j,m_j^+,\star).$$

Using [CSZ17b, Theorem 2.12], see (4.14) and the discussion thereafter, we know the (joint) limit of point-to-plane partition functions is given by independent log-normal random variables:

$$(Z_{\beta_N,N}(m_j^-,\star,m_j,z_j),Z_{\beta_N,N}(m_j,z_j,m_j^+,\star))_{j=1}^{k} \xrightarrow{d} (:e^{Y^-(t_j,x_j)}:,:e^{Y^+(t_j,x_j)}:)_{j=1}^{k},$$

where $Y^{\pm}(t_j,x_j)$ are independent centred Gaussian random variables with variance $\log(1-\widehat{\beta}^2)^{-1}$. □

Chapter 5

Size-biased diffusion limits and the inclusion process

We consider the inclusion process $(\eta^{(L,N)}(t))_{t\geq 0}$ on the complete graph of size L with N particles. Let us recall from the introduction that the inclusion process is characterised by the infinitesimal generator

$$\mathfrak{L}_{L,N} f(\eta) := \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(\mathtt{d}+\eta_y)[f(\eta^{x,y}) - f(\eta)]\,, \tag{5.1}$$

with $\eta \in \Omega_{L,N} := \{\eta \in \mathbf{N}_0^L : \sum_{x=1}^L \eta_x = N\}$ and $\eta^{x,y}$ denoting the configuration η where one particle moved from x to y, i.e. $\eta^{x,y} = \eta + e^y - e^x$, with $(e^x)_z := \mathbb{1}_{x=z}$, provided $\eta_x > 0$.

The scope of this chapter, based on the paper [CGG23], is to determine the dynamics of the condensate, cf. Definition 1.1.1, of the inclusion process (5.1) in the thermodynamic limit $N/L \to \rho$, i.e. $N, L \to \infty$ such that $N/L \to \rho$. The presented approach covers all scaling regimes of the system parameters and does not rely on previous results on condensation. In fact, the clustering of particles on diverging scales is an implicit consequence of the derived scaling limits.

In the case where the diffusivity $\mathtt{d}$ vanishes like the inverse of the system size, i.e. $\mathtt{d}(L)L = O(1)$, the derived scaling limit is equivalent to the well known Poisson-Dirichlet diffusion [EK81], offering an alternative viewpoint on these well-established dynamics. In particular, the presented approach yields a natural extension of the Poisson-Dirichlet diffusion to infinite mutation rate.

In the introduction, we saw that condensation is a macroscopic phenomenon and difficult to describe in terms of the microscopic state space $\Omega_{L,N}$. Instead a suitable **macroscopic state space** has to be considered. We begin the chapter by introducing some notation.

Preliminaries on size-biased sampling

First, we will write $\hat{\eta} \in \Omega_{L,N}$, when representing entries of $\eta \in \Omega_{L,N}$ in decreasing order, i.e.

$$\hat{\eta}_1 \geqslant \hat{\eta}_2 \geqslant \cdots \geqslant \hat{\eta}_L \geqslant 0\,.$$

Then, by appropriate rescaling, $\frac{1}{N}\hat{\eta}$ lies in the so called Kingman simplex

$$\overline{\nabla} := \left\{ p \in [0,1]^{\mathbf{N}} \,:\, p_1 \geq p_2 \geq \ldots\,,\, \sum_{i=1}^{\infty} p_i \leq 1 \right\}\,, \tag{5.2}$$

which is a compact space with respect to the product topology, and our first candidate for a macroscopic state space. Due to its structure, the Kingman simplex is a common state space in population genetics. In particular, it supports a one-parameter family of probability measures: the Poisson-Dirichlet distributions.

The **Poisson-Dirichlet distribution** was first introduced by Kingman [Kin75] as a natural limit of Dirichlet distributions. Another, more intuitive construction, uses a stick-breaking procedure: Let $U_1, U_2, \ldots$ be independent Beta$(1,\theta)$-distributed random variables (supported on $[0,1]$) and define

$$V_1 := U_1\,,\;\; V_2 := (1-U_1)U_2\,,\;\; V_3 := (1-U_1)(1-U_2)U_3\,,\ldots,$$

i.e. we start with a stick of unit length and continue by breaking a random fraction of U_1 apart. Then, we iteratively do the same with the remaining part of the stick. The resulting random vector $V = (V_i)_{i\geq 1}$ is said to be GEM(θ)-distributed. Reordering the entries of V in decreasing order yields $\widehat{V}$, which is known to be Poisson-Dirichlet distributed with parameter θ, see e.g. [Fen10]. We will write PD(θ) for the Poisson-Dirichlet distribution with parameter θ. Note that PD(θ) gives full mass to the set

$$\nabla := \left\{ p \in \overline{\nabla} \,:\, \sum_{i=1}^{\infty} p_i = 1 \right\} \subset \overline{\nabla}\,.$$

Later, the Poisson-Dirichlet was identified as the unique stationary measure

of the split-merge dynamics [ZZMWD04, Sch05] and the Poisson-Dirichlet diffusion [EK81]. Despite it being introduced in the field of population genetics, the Poisson-Dirichlet distribution also appeared in statistical mechanics [GUW11, BU11, IT20] and recently in the study of IPS [JCG19, CGG22].

A last, and very important, ingredient for this chapter is **size-biased sampling**. Let us introduce the *size-biased sampled configuration* $\tilde{\eta}$ with respect to $\eta \in \Omega_{L,N}$. We start by constructing a random permutation σ iteratively as follows: First,

(i) $\sigma(1) = x$ with probability $\frac{\eta_x}{N}$, $x \in \{1, \dots, L\}$,

and for any following index $k = 1, \dots, L$

(ii) $\sigma(k) = x$ with probability $\frac{\eta_x}{N - \sum_{j=1}^{k-1} \eta_{\sigma(j)}}$,
for all $x \in \{1, \dots, L\} \setminus \{\sigma(1), \dots, \sigma(k-1)\}$.

The size-biased sampled distribution $\tilde{\eta}$ is then given by $\tilde{\eta}_k = \eta_{\sigma(k)}$. In other words, in the k^{th} iteration of the construction, we pick (uniformly at random) one of the (remaining) particles and assign the total number of particles, found at the same location to $\tilde{\eta}_k$. The procedure terminates, once no more particles can be sampled. For convenience, we fill all remaining entries of $\tilde{\eta}$ with zeros, to reach length L.

The reader may think of size-biased sampling in terms of labelling all particles, according to their location, and uniformly at random picking a particle. Once a particle, thus a location, is picked, all particles having the same label (location) are removed and the procedure repeats. We refer to [JCG19, Definition 2] and [CGG22, Section 2.2] for details.

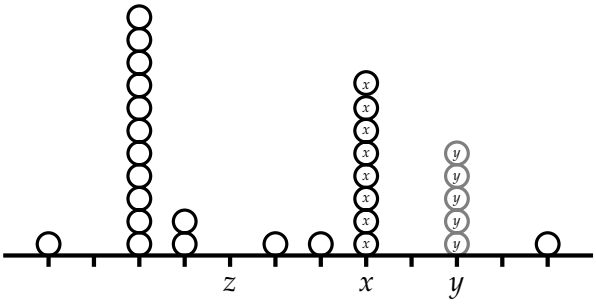


Figure 5.1: Size-biased sampling of a particle configuration ($N = 30$), where location y was picked in the first iteraton. In the second iteration, x will be sampled with probability $\frac{\eta_x}{N - \eta_y} = \frac{8}{25}$. Note that empty sites, such as z, will not be sampled.

Condensation in the inclusion process

For the inclusion process, condensation was first studied in [GRV11], however for spatially inhomogeneous systems on a finite lattice. The finite lattice case has been subject to further studies in [GRV13, BDG17, KS21]. The present chapter concerns the study of the inclusion process in the thermodynamic limit $N/L \to \rho \geq 0$. In this case, condensation was studied heuristically in [CCG14]. They considered a one-dimensional periodic lattice with totally asymmetric dynamics and strongly vanishing diffusion rate. A modified model with stronger particle interactions, leading to instantaneous condensation, has been considered in [WE12, CCG15]. On a rigorous level, the thermodynamic limit of stationary distributions $(\pi_{L,N})_{L,N}$ (assosciated to the generators $(\mathfrak{L}_{L,N})_{L,N}$) was subject of [JCG19]. They observed that the condensation regime $\mathtt{d} = \mathtt{d}(L) \to 0$, as $L \to \infty$, is subdivided into three parts:

- If $\mathtt{d}L \to 0$, then the condensate is given by a single cluster, and if in addition $\mathtt{d}L \log L \to 0$, this cluster is holding all particles.
- If $\mathtt{d}L \to \theta \in (0, \infty)$, then the condensate concentrates on macroscopic scales and is distributed according to a Poisson-Dirichlet distribution $\mathrm{PD}(\theta)$ in the sense
$$\frac{1}{N}\hat{\eta} \xrightarrow{d} \mathrm{PD}(\theta)\,,$$
where $\hat{\eta}$ denotes the ordered configuration $\eta \sim \pi_{L,N}$.
- On the other hand, if $\mathtt{d}L \to \infty$, the condensate is located on mesoscopic scales and the clusters are independent. More precisely,
$$\mathtt{d}(\tilde{\eta}_1, \dots, \tilde{\eta}_n) \xrightarrow{d} \mathrm{Exp}(\tfrac{1}{\rho})^{\otimes n}\,, \tag{5.3}$$
where $\tilde{\eta}$ denotes a size-biased sample of $\eta \sim \pi_{L,N}$.

In fact, the above classification holds for any irreducible and spatially homogeneous dynamics on diverging finite graphs, where the inclusion process has stationary product measures. [JCG19] entirely characterises the stationary condensates for the inclusion process, since condensation only occurs if $\mathtt{d} \to 0$. It was proven in [CGG22] that perturbations of the transition rates (in the case $\theta \leq 1$) still give rise to a Poisson-Dirichlet distributed condensate.

From the summary of condensation results for the inclusion process above, we can see the importance of *size-biased sampling*: In the case $\mathtt{d}(L)L \to \infty$, the "strong ordering" $\frac{1}{N}\hat{\eta}$ does not pick up information of the condensate

on mesoscopic scales, while also disregarding the bulk phase of the system. Instead, size-biased sampling allows for the analysis of both the bulk and condensed phase, see also [CGG22].

5.1 Results and comparison to the literature

The chapter's results are twofold. On the one hand, we describe convergence of the inclusion process towards measure-valued diffusions, which are parametrised by $\theta := \lim_{N/L\to\rho} \mathrm{d}L \in [0,\infty]$. In the case $\theta < \infty$, the determined scaling limit is equivalent to the well known infinitely-many-neutral-alleles diffusion model [EK81], also known as Poisson-Dirichlet diffusion. On the other hand, the new description of the infinitely-many-neutral-alleles diffusion model, allows us to construct a natural extension of the process, when $\theta = \infty$.

5.1.1 Scaling limits of the inclusion process

In the beginning of this chapter, we introduced the Kingman simplex as a suitable macroscopic state space. However, the basis of this chapter is to consider a different choice, allowing for a finer analysis of the underlying dynamics. In the following, we distinguish between the cases $\theta < \infty$ and $\theta = \infty$.

When $\theta < \infty$, we embed particle configurations into $\mathcal{M}_1([0,1])$, with the topology induced by weak convergence of measures, using the maps

$$\mu^{(\cdot)}_{L,N} : \Omega_{L,N} \to \mathcal{M}_1([0,1]) \quad \text{of the form} \quad \mu^{(\eta)}_{L,N} := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{\eta_x}{N}} \,. \tag{5.4}$$

Note that this corresponds to a size-biased empirical measure on the space of mass fractions $[0,1]$. Not every measure in $\mathcal{M}_1([0,1])$ can be approximated by particle configurations using $\mu^{(\cdot)}_{L,N}$, e.g. every point mass $\alpha\delta_z$ in (5.4) satisfies $\alpha \geq z$. Instead, we restrict ourselves to the closed subspace

$$E := \mu^{(\overline{\nabla})} \subset \mathcal{M}_1([0,1]) \,, \tag{5.5}$$

defined as the range of the function

$$\mu^{(\cdot)} : \overline{\nabla} \to \mathcal{M}_1([0,1]) \quad \text{with} \quad \mu^{(p)} := (1 - \|p\|_1)\delta_0 + \sum_{i=1}^{\infty} p_i \delta_{p_i} \,, \tag{5.6}$$

where $\|p\|_1 := \sum_{i=1}^{\infty} p_i$ and $\overline{\nabla}$ denotes the Kingman simplex (5.2). We may often drop the subscripts in (5.4) and simply write $\mu^{(\cdot)}$, but the meaning will be clear from the context. Moreover, we will sometimes write $\mu_{\#}\eta$ instead of $\mu^{(\eta)}$ in order to avoid overloaded notation. The above mappings $\mu^{(\cdot)}$ and $\mu^{(\cdot)}_{L,N}$ do not preserve spatial information of particle configurations, but this also does not enter the dynamics on the complete graph. Note that the map (5.6) was already mentioned in [EK81], however, only to prove denseness of the domain of functions considered there.

In order to describe the limiting dynamics, we consider the domain of functions

$$\begin{aligned}\mathcal{D}(\mathcal{L}_\theta) =&\{\text{sub-algebra of } C(E) \text{ generated} \\ &\text{by functions } \mu \mapsto \mu(h)\,,\ h \in C^3([0,1])\}\,.\end{aligned} \tag{5.7}$$

We associate a pre-generator of a superprocess, which acts on functions $H(\mu) = \mu(h_1)\cdots\mu(h_n) \in \mathcal{D}(\mathcal{L}_\theta)$ via

$$\begin{aligned}\mathcal{L}_\theta H(\mu) :=& 2\sum_{1\le k<l\le n}\big(\mu(Bh_kBh_l) - \mu(Bh_k)\mu(Bh_l)\big)\prod_{j\neq k,l}\mu(h_j) \\ &+ \sum_{1\le k\le n}\mu(A_\theta h_k)\prod_{j\neq k}\mu(h_j)\,,\end{aligned} \tag{5.8}$$

with $Bh(z) := h(z) + zh'(z) = (zh(z))'$. Here, the first part is usually referred to as interaction term and A_θ denotes the single-particle operator of the form

$$\begin{aligned}A_\theta h(z) :=&(1-z)(Bh)'(z) + \theta(Bh(0) - Bh(z)) \\ =&z(1-z)h''(z) + (2 - z(2+\theta))h'(z) + \theta(h(0) - h(z))\,.\end{aligned} \tag{5.9}$$

The operator $Bh(z)$ should be thought of as a "size-biased derivative", which appears due to our choice of embedding (5.4). For example, a single site containing a mass fraction $z \in [0,1]$ will be represented by a point-mass $z\delta_z$. Thus, changes in z will result both in a change of the amount of mass and its position.

Our first result identifies the process described by $\mathcal{L}_\theta$ as the correct scaling limit of the inclusion process.

Theorem 5.1.1 ([CGG23]). *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L)$ such that $\mathtt{d}L \to \theta \in [0,\infty)$. If $\eta^{(L,N)}(0)$ such that $\mu_\# \eta^{(L,N)}(0) \xrightarrow{d} \mu_0 \in E$, then as $\frac{N}{L} \to \rho$*

$$\left(\mu_\# \eta^{(L,N)}(t)\right)_{t\geq 0} \xrightarrow{d} (\mu_t)_{t\geq 0}\,, \quad \text{in } D([0,\infty), E)\,. \tag{5.10}$$

Here $(\mu_t)_{t\geq 0}$ denotes the measure-valued process on E (5.5) generated by $\mathcal{L}_\theta$ (5.8), with initial value μ_0.

In Proposition 5.2.6 we prove that the closure of $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ is indeed the generator of a Feller process on the state space E. An immediate consequence of Theorem 5.1.1 is that the inclusion process exhibits condensation for arbitrary times $t > 0$, independent of the initial condition.

For the case $\theta = \infty$, i.e. $\mathtt{d}L$ diverging, we expect clusters on the scale $1/\mathtt{d}$, cf. (5.3). In this case, we consider the embedding

$$\hat{\mu}^{(\cdot)} = \hat{\mu}^{(\cdot)}_{L,N} : \Omega_{L,N} \to \mathcal{M}_1(\mathbf{R}_+) \quad \text{with} \quad \hat{\mu}^{(\eta)}_{L,N} := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\mathtt{d}L\frac{\eta_x}{N}}\,, \tag{5.11}$$

mapping particle configurations into the space of probability measures $\mathcal{M}_1(\mathbf{R}_+)$, again with the topology induced by weak convergence of measures. In contrast to $\theta < \infty$, any measure in $\mathcal{M}_1(\mathbf{R}_+)$ can be approximated by particle configurations using (5.11), cf. Lemma 5.2.7 below, which is why we do not have to restrict ourselves to a strict subset of probability measures as above.

The lack of compactness of $\mathbf{R}_+$ allows for diverging rescaled masses of particle configurations, thus when $\mathtt{d}L \to \infty$, we expect the scaling limit to be a measure-valued process on $\mathcal{M}_1(\overline{\mathbf{R}}_+)$, with $\overline{\mathbf{R}}_+ = [0,\infty]$. We include ∞ to describe mass on larger scales than $1/\mathtt{d}$. Indeed the correct limit turns out to be a process on $\mathcal{M}_1(\overline{\mathbf{R}}_+)$ without interaction and single-particle operator

$$\begin{aligned} \hat{A}h(z) :=& (Bh)'(z) + (Bh(0) - Bh(z)) \\ =& zh''(z) + (2-z)h'(z) + (h(0) - h(z))\,, \end{aligned} \tag{5.12}$$

acting on h in the domain

$$\begin{aligned} \mathcal{D}(\hat{A}) := \{h \,:\, h(\infty) = 0 \text{ and } h|_{\mathbf{R}_+} \in C^3_c(\mathbf{R}_+)\} \\ \cup \{\text{constant functions}\} \subset C(\overline{\mathbf{R}}_+)\,. \end{aligned} \tag{5.13}$$

Slowing down the evolution of the inclusion process appropriately, we have:

Theorem 5.1.2 ([CGG23]). *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L) \to 0$ such that $\mathtt{d}L \to \infty$. If $\hat{\mu}_\# \eta^{(L,N)}(0) \xrightarrow{d} \hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$, then as $\frac{N}{L} \to \rho$*

$$\left(\hat{\mu}_\# \eta^{(L,N)}\left(\tfrac{t}{\mathtt{d}L}\right)\right)_{t\geq 0} \xrightarrow{d} (\hat{\mu}_t)_{t\geq 0}\,, \quad \text{in } D([0,\infty), \mathcal{M}_1(\overline{\mathbf{R}}_+))\,. \tag{5.14}$$

Here $(\hat{\mu}_t)_{t\geq 0}$ denotes the measure-valued process on $\mathcal{M}(\overline{\mathbf{R}}_+)$ with initial value $\hat{\mu}_0$ generated by

$$\hat{\mathcal{L}} H(\mu) = \sum_{1\leq k\leq n} \mu(\hat{A}h_k) \prod_{\substack{m=1\\ m\neq k}}^{n} \mu(h_m)\,, \tag{5.15}$$

with $H(\mu) = \mu(h_1)\cdots\mu(h_n)$, $h_k \in \mathcal{D}(\hat{A})$.

The operator $\hat{\mathcal{L}}$ may be interpreted as a Fleming-Viot process without interaction. This is in contrast to the generator $\mathcal{L}_\theta$, which does not have Fleming-Viot interactions of the form

$$2 \sum_{1\leq k<l\leq n} \big(\mu(h_k h_l) - \mu(h_k)\mu(h_l)\big) \prod_{j\neq k,l} \mu(h_j)\,.$$

We will see that the limiting dynamics $\hat{\mathcal{L}}$ are deterministic with absorbing state $\hat{\mu} = \mathrm{Exp}(1) \in \mathcal{M}_1(\mathbf{R}_+)$. In fact, the statement of Theorem 5.1.2 can be reformulated into a hydrodynamic limit (Proposition 5.3.6). Moreover, if $\hat{\mu}_0(\mathbf{R}_+) = 1$, then $\hat{\mu}_t(\mathbf{R}_+) = 1$ for every $t \geq 0$, and mass does not escape to larger scales (Corollary 5.3.8).

Again, an immediate consequence of the theorem above is condensation of the inclusion process, in the regime $\mathtt{d}L \to \infty$, for any positive time and initial condition. The fact that the dynamics were slowed down appropriately, implies rapid mixing in the inclusion process. Overall, Theorems 5.1.1 and 5.1.2 fully determine the dynamics of the inclusion process, with vanishing diffusivity, on complete graphs in the thermodynamic limit with density $\rho \in (0,\infty)$. For a discussion of the boundary cases $\rho \in \{0,\infty\}$, we refer to Section 5.4.

5.1.2 A size-biased viewpoint on the PD-diffusion

In their seminal work [EK81] Ethier and Kurtz introduced the **infinitely-many-neutral-alleles diffusion model**, as a natural scaling limit of finite-dimensional Wrigh–Fisher diffusion models. This process is also commonly referred to as **Poisson-Dirichlet diffusion**, which we will use throughout the chapter. In this section, we will see that the measure-valued process described in Theorem 5.1.1 provides an alternative description of the Poisson-Dirichlet diffusion. The latter is an infinite-dimensional Feller process on $\overline{\nabla}$ (5.2) with pre-generator[1]

$$\mathcal{G}_\theta f = \sum_{i,j=1}^{\infty} p_i(\mathbb{1}_{i=j} - p_j)\partial^2_{p_i p_j} f - \theta \sum_{i=1}^{\infty} p_i \partial_{p_i} f\,, \tag{5.16}$$

acting on functions in the domain

$$\mathcal{D}_{mon}(\mathcal{G}_\theta) := \{\text{sub-algebra of } C(\overline{\nabla}) \text{ generated by } 1, \varphi_2, \varphi_3, \ldots\}\,, \tag{5.17}$$

where $\varphi_m(p) := \sum_{i=1}^{\infty} p_i^m$ for $m \geq 2$. $\mathcal{G}_\theta$ acts on such test functions with the convention that occurring sums on the right-hand side of (5.16) are *evaluated on ∇, and extended to $\overline{\nabla}$ by continuity*. The name Poisson-Dirichlet diffusion is adequate, since its unique invariant distribution is the Poisson-Dirichlet distribution PD(θ).

Naturally, one can consider the mapping of the Poisson-Dirichlet diffusion under $\mu^{(\cdot)}$ (5.6), which yields a process in $E \subset \mathcal{M}_1([0,1])$. As already indicated above, this push-forward process indeed agrees with the process derived in Theorem 5.1.1.

Proposition 5.1.3. *Let $(\mu_t)_{t\geq 0}$ be the measure-valued process generated by $\mathcal{L}_\theta$ (5.8) with initial data $\mu_0 \in E$, then*

$$(\mu_t)_{t\geq 0} \overset{d}{=} \left(\mu^{(X_t)}\right)_{t\geq 0}\,, \tag{5.18}$$

where $(X_t)_{t\geq 0}$ denotes the corresponding Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$ (5.16), with $\mu^{(X_0)} = \mu_0$. The following properties translate immediately from $(X_t)_{t\geq 0}$ to $(\mu_t)_{t\geq 0}$:

(i) The process $(\mu_t)_{t\geq 0}$ has a unique stationary distribution, which is reversible.

[1] The original formulation of the pre-generator in [EK81] includes a multiplicative factor of $\frac{1}{2}$ which we omitted here.

It is given by $\mathbf{P} = \mu_{\#}\mathrm{PD}(\theta)$*, i.e. the the law of*

$$\mu^{(X)} = \sum_{i=1}^{\infty} X_i \delta_{X_i}\,, \quad X \sim \mathrm{PD}(\theta)\,. \tag{5.19}$$

(ii) The process $(\mu_t)_{t\geq 0}$ *has continuous sample paths in* E*.*

(iii) For any initial value $\mu_0 \in E$*, we have*

$$\mathbb{P}(\mu_t(\{0\}) = 0\ \forall t > 0) = 1\,. \tag{5.20}$$

Together with Theorem 5.1.1, this yields the following corollary.

Corollary 5.1.4. *Let* $\rho \in (0,\infty)$ *and* $\mathtt{d} = \mathtt{d}(L)$ *such that* $\mathtt{d}L \to \theta \in [0,\infty)$*. If* $\eta^{(L,N)}(0)$ *such that* $\frac{1}{N}\hat{\eta}^{(L,N)}(0) \xrightarrow{d} X_0 \in \overline{\nabla}$*, then*

$$\frac{1}{N}\left(\hat{\eta}^{(L,N)}(t)\right)_{t\geq 0} \xrightarrow{d} (X_t)_{t\geq 0}\,, \quad \text{as} \quad \frac{N}{L} \to \rho\,. \tag{5.21}$$

Here $(X_t)_{t\geq 0}$ *denotes the Poisson-Dirichlet diffusion on* $\overline{\nabla}$ *with parameter* θ *and initial value* X_0*, generated by* $\mathcal{G}_\theta$ (5.16)*, and* $\hat{\eta}$ *denotes the ordered particle configuration.*

The measure valued process generated by $\hat{\mathcal{L}}$ is the natural extension of the process $\mathcal{L}_\theta$ (and thus to the Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$) when $\theta \to \infty$. A first indication for this relationship can already be observed on the level of stationary distributions. From Proposition 5.1.3(i), we recall that the stationary distribution with respect to $\mathcal{L}_\theta$ is given by the size-biased sample of $\mathrm{PD}(\theta)$. Consider $X^{(\theta)} \sim \mathrm{PD}(\theta)$ and sample an index $I \in \mathbf{N}$ such that

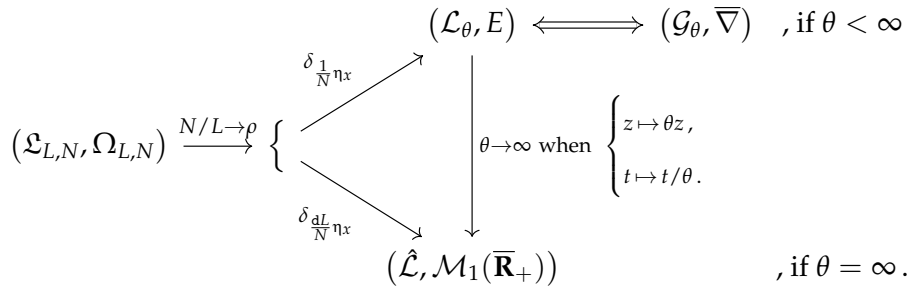
$$I = i \quad \text{with probability } X_i^{(\theta)}\,, \tag{5.22}$$

i.e. we pick the index I with size-bias. It is well known [Fen10, Theorem 2.7] that $X_I^{(\theta)} \sim \mathrm{Beta}(1,\theta)$. Now, the following connection between a Beta and an Exponential distribution holds:

$$\theta\,\mathrm{Beta}(1,\theta) \xrightarrow{d} \mathrm{Exp}(1)\,, \quad \text{as } \theta \to \infty\,. \tag{5.23}$$

Hence, as $\theta \to \infty$, the rescaled size-biased sample $\theta\, X_I^{(\theta)}$ converges weakly to an Exp(1) random variable, which is the absorbing state of the deterministic dynamics induced by $\hat{\mathcal{L}}$ (Corollary 5.3.10).

This relationship can also be made sense of on the level of processes, summarised in the following diagram:



We analyse the inclusion process $(\mathfrak{L}_{L,N}, \Omega_{L,N})$ and consider the two cases $\theta < \infty$ (Theorem 5.1.1) and $\theta = \infty$ (Theorem 5.1.2), with appropriate embeddings of configurations in the space of probability measures. In the case $\theta < \infty$ the limiting process is equivalent to the Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$ (Proposition 5.1.3). Furthermore, our size-biased approach allows for a meaningful limit when $\theta = \infty$, identifying a natural extension for models with Poisson-Dirichlet diffusion limit. For this matter, we introduce the scaling operator $S_\theta : E \to \mathcal{M}_1(\mathbf{R}_+)$, which linearly scales measures on the unit interval to measures on the interval $[0, \theta]$:

$$S_\theta : \mu(\,\mathrm{d}z\,) \mapsto \mu(\,\mathrm{d}\tfrac{z}{\theta}\,)\,. \tag{5.24}$$

Theorem 5.1.5 ([CGG23]). *Let $(\mu_t^{(\theta)})_{t\geq 0}$ be the process generated by $\mathcal{L}_\theta$ and $(\hat{\mu}_t)_{t\geq 0}$ be the process generated by $\hat{\mathcal{L}}$. If $S_\theta \mu_0^{(\theta)} \xrightarrow{d} \hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$, then*

$$\left(S_\theta \mu_{t/\theta}^{(\theta)}\right)_{t\geq 0} \xrightarrow{d} (\hat{\mu}_t)_{t\geq 0}\,, \quad \text{in } C([0,\infty), \mathcal{M}_1(\overline{\mathbf{R}}_+))\,, \quad \text{as } \theta \to \infty\,, \tag{5.25}$$

where we consider the topology induced by weak convergence on $\mathcal{M}_1(\overline{\mathbf{R}}_+)$.

Remark 5.1.6. *In the above theorem, we saw that scaling of space is necessary to observe a meaningful limit as $\theta \to \infty$. Similarly, one could scale the Kingman simplex to $\theta\overline{\nabla}$. However, in the limit we lack the property of distinguishing between separate scales. Consider for example $\theta = n^2$, $n \in \mathbf{N}$, and the sequence*

$$p^{(\theta)} = \tfrac{1}{2}\big(\underbrace{\tfrac{1}{\sqrt{\theta}}, \dots, \tfrac{1}{\sqrt{\theta}}}_{\sqrt{\theta}\text{ times}}, \underbrace{\tfrac{1}{\theta}, \dots, \tfrac{1}{\theta}}_{\theta\text{ times}}, 0, \dots\big) \in \nabla\,. \tag{5.26}$$

Then $\theta\, p_i^{(\theta)} \to \infty$, *for every* $i \in \mathbf{N}$. *On the other hand,*

$$S_\theta \mu^{(p^\theta)} = \tfrac{1}{2}\delta_{\frac{1}{2}} + \tfrac{1}{2}\delta_{\frac{1}{2}\sqrt{\theta}} \xrightarrow{d} \tfrac{1}{2}\delta_{\frac{1}{2}} + \tfrac{1}{2}\delta_{\infty} \in \mathcal{M}_1(\overline{\mathbf{R}}_+)\,, \tag{5.27}$$

which captures both the amount of diverging mass and information on scales $\frac{1}{\theta}$. *This highlights the fact that considering the space E, instead of* $\overline{\nabla}$, *is essential for a detailed analysis of the Poisson-Dirichlet diffusion, in the boundary case* $\theta \to \infty$.

Our results link the inclusion process directly to the Poisson-Dirichlet diffusion (5.16), which allows for an enhanced understanding of the latter. In the literature, the precise dynamics within the Poisson-Dirichlet diffusion remain difficult to grasp, due to the restriction of test functions in $\mathcal{D}_{mon}(\mathcal{G}_\theta)$ (despite being a convenient choice). In [EK81], also an enlarged domain of test-functions of the form

$$p \mapsto \sum_{i=1}^{\infty} h(p_i)\,, \quad h \in C^2([0,1]) \text{ with } h(0) = h'(0) = 0\,, \tag{5.28}$$

was considered. However, this does not improve the understanding of the dynamics on an intuitive level, which is particularly due to the fact that "sums are evaluated on ∇ and extended to $\overline{\nabla}$ by continuity". Therefore, leaving for example the question open, if the immediate clustering in the Poisson-Dirichlet diffusion (Proposition 5.1.3(iii)) is a consequence of this convention or the dynamics. In the present work, we instead consider test functions of the form

$$p \mapsto h(0) + \sum_{i=1}^{\infty} p_i(h(p_i) - h(0))\,, \quad h \in C^2([0,1])\,. \tag{5.29}$$

For a fixed $p \in \overline{\nabla}$, the right-hand side is the expectation with respect to the probability measure $\mu^{(p)}$, recall (5.6). Thus, functions can be evaluated directly on $\overline{\nabla}$ (without continuous extension). Notably, the functions φ_m in (5.17) are of the form (5.29) with $h(p) = p^{m-1}$. This size-biased viewpoint allows to represent observables in the Poisson-Dirichlet diffusion in terms of one-dimensional diffusions, thanks to a duality result, see Section 5.2.3 below.

The Poisson-Dirichlet diffusion was treated previously as a measure-valued process in $\mathcal{M}_1([0,1])$ in [EG87, EK87]. There it was considered as

a Fleming-Viot process, characterised by the generator

$$\mathcal{L}_{FV}H(\nu) = 2\sum_{1\le k<l\le n}\big(\nu(h_kh_l)-\nu(h_k)\nu(h_l)\big)\prod_{j\neq k,l}\nu(h_j) + \sum_{1\le k\le n}\nu(A_{FV}h_k)\prod_{j\neq k}\nu(h_j)\,. \tag{5.30}$$

when applied to a cylindrical test function of the form $H(\nu) = \nu(h_1)\cdots\nu(h_n)$, with mutation operator

$$A_{FV}h(u) := \theta\int_0^1 [h(w)-h(u)]\,\mathrm{d}w\,. \tag{5.31}$$

Here, elements in $[0,1]$ are interpreted as types, and uniform jumps at rate θ in the mutation operator correspond to mutation events. In the thermodynamic limit, these dynamics can be derived directly from the inclusion process on a complete graph with $\mathtt{d}L \to \theta \in (0,\infty)$, using the embedding

$$\eta \in \Omega_{L,N} \mapsto \sum_{x=1}^{L}\frac{\eta_x}{N}\delta_{\frac{x}{L}} \in \mathcal{M}([0,1])\,. \tag{5.32}$$

This is presented briefly in Section 5.2.4. The corresponding process describes the spatial distribution of mass on the rescaled lattice. This is different from the approach in this chapter, where we ignore spatial information and only keep track of the mass distribution, cf. (5.6). On the other hand, our approach is more robust and allows for an extended analysis of the model when $\theta = \infty$. The embedding (5.32) was also considered in [CT16] for the inclusion process on a complete graph of fixed size. They study convergence to equilibrium in the long-time limit, with diverging mass $N \to \infty$.

The usual approach in the literature to construct the Poisson-Dirichlet diffusion, is to take the large L-limit of an L-dimensional diffusion model. Alternatively, discrete models have been considered, then convergence to the L-dimensional diffusion model, when $N \to \infty$, is proven first. See for example [EK81, CBE+17, RW09]. To our best knowledge, the present work [CGG23] is the first to consider a thermodynamic limit. This makes sense, both from a physical and population genetics perspective.

Applying our approach to other geometries may be possible for dense random graphs along the lines of [BHM+22], which have diverging degrees leading to a self-averaging effect similar to the complete graph. In general, spatial models are difficult to treat since the inclusion process, after the em-

bedding (5.6), is not Markovian in general. Consider for example nearest-neighbour dynamics on a regular lattice, then it is known that the random-walk and the inclusion part of the dynamics have two different time scales [ACR21], and more sophisticated methods are necessary to treat this case.

5.1.3 The Trotter–Kurtz approximation theorem

In this section, we present our main tool which we shall use throughout the chapter, namely the Trotter–Kurtz approximation theorem. It states that convergence of generators is equivalent to the convergence of semigroups. In the special case of Markov generators, this suffices to conclude convergence of the corresponding processes.

In the following, let $(L_n)_{n\in\mathbf{N}}$, L be Banach spaces and $\hat{\pi}_n : L \to L_n$ bounded linear transformations.

Proposition 5.1.7 (Trotter–Kurtz). *For $n \in \mathbf{N}$, let $(\mathcal{T}_t^{(n)})_{t\geqslant 0}$ and $(\mathcal{T}_t)_{t\geqslant 0}$ be strongly continuous contraction semigroups on L_n and L, with generators $\mathcal{A}_n$ and $\mathcal{A}$. Furthermore, let $\mathcal{D}$ be a core for $\mathcal{A}$. Then the following statements are equivalent:*

(i) For every $f \in L$ and $T > 0$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|\mathcal{T}_t^{(n)}\hat{\pi}_n f - \hat{\pi}_n \mathcal{T}_t f\| = 0\,.$$

(ii) For every $f \in L$ and $t \geqslant 0$,

$$\lim_{n\to\infty} \|\mathcal{T}_t^{(n)}\hat{\pi}_n f - \hat{\pi}_n \mathcal{T}_t f\| = 0\,.$$

(iii) For every $f \in \mathcal{D}$, there exist $f_n \in \mathcal{D}(\mathcal{A}_n)$ such that

$$\lim_{n\to\infty} \|f_n - \hat{\pi}_n f\| = 0 \quad \textit{and} \quad \lim_{n\to\infty} \|\mathcal{A}_n f_n - \hat{\pi}_n \mathcal{A} f\| = 0\,.$$

We refer the reader to the original works [Tro58, Kur69], for a reference of the above proposition see [EK86, Theorem 1.6.1]. The result holds for a wide class of semigroups, in particular, it can be applied to generators of Feller processes which yields the following convergence result: For $n \in \mathbf{N}$, let S_n and S be metric spaces, with S compact. Moreover, consider measurable embeddings $\iota_n : S_n \to S$, for all $n \in \mathbf{N}$.

Proposition 5.1.8. *Let $(\mathcal{T}_t)_{t\geqslant 0}$ be a Feller semigroup on $C(S)$, and for $n \in \mathbf{N}$, let $(\mathcal{T}_t^{(n)})_{t\geqslant 0}$ be a semigroup on $\mathcal{B}(S_n)$ (bounded measurable functions) given by a transition function. If*

(i) $X^{(n)} := \iota_n \circ Y^{(n)}$ has sample paths in $D([0,\infty), S)$, where $Y^{(n)} = (Y_t^{(n)})_{t\geqslant 0}$ is the Markov process corresponding to $(\mathcal{T}_t^{(n)})_{t\geqslant 0}$,

(ii) for every $f \in C(S)$ and $t \geqslant 0$

$$\lim_{n\to\infty} \|\mathcal{T}_t^{(n)} \hat{\pi}_n f - \hat{\pi}_n \mathcal{T}_t f\|_\infty = 0\,,$$

where $\hat{\pi}_n f := f \circ \iota_n$,

(iii) and $X_0^{(n)} \xrightarrow{d} \nu \in \mathcal{M}_1(S)$.

Then, as $n \to \infty$,

$$(X_t^{(n)})_{t\geqslant 0} \xrightarrow{d} (X_t)_{t\geqslant 0}\,, \qquad \text{in } D([0,\infty), S)\,,$$

with $(X_t)_{t\geqslant 0}$ being the Feller process associated to $(\mathcal{T}_t)_{t\geqslant 0}$, with initial law ν.

The proposition holds in greater generality for locally compact, separable metric spaces S, see [EK86, Theorem 4.2.11]. We also refer the reader to [Kal02, Theorem 19.25].

It is remarkable that convergence of the semigroups in Proposition 5.1.8 implies tightness of the corresponding processes, which is worth a brief explanation. To conclude tightness of $(X_t^{(n)})_{t\geqslant 0}$, it suffices to show that $(f(X_t^{(n)}))_{t\geqslant 0}$ is tight for all $f \in C(E)$ [EK86, Corollary 3.9.3]. To see this, first note that convergence of the semigroups implies convergence of the generators in the sense of Proposition 5.1.7(iii) above. Thus, for every $f \in \mathcal{D}(\mathcal{A})$ there exists a sequence $f_n \in \mathcal{D}(\mathcal{A}_n)$ such that

$$t \mapsto f_n(X_t^{(n)}) - \int_0^t \mathcal{A}_n f_n(X_s^{(n)})\,\mathrm{d}s$$

is a martingale, while $\|f_n - \hat{\pi}_n f\|_{\infty,S_n} \to 0$. Here $\mathcal{A}_n$ and $\mathcal{A}$ denote the generators of the semigroups in Proposition 5.1.8. The martingale structure allows to uniformly control (in time and n) a modulus of continuity of $(f(X_t^{(n)}))_{t\geqslant 0}$ on $D([0,\infty), \mathbf{R})$, implying tightness. This can be extended to all $f \in C(S)$. For details, we refer the reader to [EK86, Theorem 3.9.4]. On the other hand, convergence of finite-dimensional distributions of $(X_t^{(n)})_{t\geqslant 0}$ is an immediate

consequence of the convergence of semigroups, which concludes the desired result.

Structure of the chapter

In Section 5.2 we show that $\mathcal{L}_\theta$ generates a Feller process and prove Theorem 5.1.1. We make use of explicit approximations of the inclusion process generator and the Trotter-Kurtz approximation theorem. Moreover, we prove the equivalence of the Poisson-Dirichlet diffusion and our scaling limit in Section 5.2.2. Lastly, we discuss the advantages of considering size-biased dynamics in Section 5.2.3 and prove convergence of the inclusion process to a Fleming-Viot process in Section 5.2.4. In Section 5.3 we determine the scaling limit when $\theta = \infty$, following a similar approach as in the case $\theta < \infty$. We finish the section by proving the convergence $\frac{1}{\theta}\mathcal{L}_\theta \to \hat{\mathcal{L}}$ stated in Theorem 5.1.5. Lastly, we discuss boundary cases $\rho \in \{0, \infty\}$, fluctuations and open problems in Section 5.4.

5.2 Scaling limits in the case $\mathrm{d}L \to \theta < \infty$

5.2.1 The measure-valued process

In this section, we will prove that the measure-valued process generated by $\mathcal{L}_\theta$ (5.8) is a Feller process on the state space E (5.5). Furthermore, we deduce weak convergence for the inclusion process (on the path space) embedded in the space of probability measures on the unit interval.

At this point, the space of size-biased probability measures E might still seem abstract. However, as we see in the following lemma, it is in fact isomorphic to the Kingman simplex.

Lemma 5.2.1. *The map $\mu^{(\cdot)} : \overline{\nabla} \to E$, cf. (5.6), is an isomorphism.*

Proof. First note that surjectivity is trivial due to the definition of E. Now, consider $p, q \in \overline{\nabla}$ such that $p \neq q$. Then there exists an index $i \in \mathbf{N}$ such that $p_i \neq q_i$ and $p_j = q_j$ for all $j < i$, without loss of generality assume $p_i > q_i$. Then

$$\mu^{(p)}([p_i, 1]) \geq \sum_{j=1}^{i} p_j > \sum_{j=1}^{i-1} q_j = \mu^{(q)}([p_i, 1])\,. \tag{5.33}$$

thus, $\mu^{(p)} \neq \mu^{(q)}$.

In order to show that the map $\mu^{(\cdot)}$ is continuous, consider a sequence of partitions $(p^{(n)})_{n\in\mathbf{N}}$ converging to p in $\overline{\nabla}$. Then for every $h \in C([0,1])$ (uniformly in n)

$$\Big| \sum_{i=M}^{\infty} p_i^{(n)} (h(p_i^{(n)}) - h(0)) \Big| \leq \sup_{0\leq z\leq \frac{1}{M}} |h(z) - h(0)| \to 0\,, \quad \text{as } M \to \infty\,, \tag{5.34}$$

where we used the fact that $p_i \leq \frac{1}{i}$ for any $i \in \mathbf{N}$. This implies in particular $\mu^{(p^{(n)})} \xrightarrow{d} \mu^{(p)}$, recalling that $\mu^{(p)}(h) = h(0) + \sum_{i=1}^{\infty} p_i(h(p_i) - h(0))$.

Continuity of the inverse is now immediate: Let $(\mu_n)_{n\in\mathbf{N}}$ be a sequence in E, weakly converging to $\mu \in \mathcal{M}_1([0,1])$, then we can identify each μ_n with a unique $p^{(n)}$ satisfying $\mu^{(p^{(n)})} = \mu_n$. Due to compactness of $\overline{\nabla}$, it suffices to consider convergent subsequences, say $(p^{(n_j)})_{j\in\mathbf{N}}$ with limit p. By assumption and continuity of $\mu^{(\cdot)}$

$$\mu_{\#} p^{(n_j)} \xrightarrow{d} \mu^{(p)} = \mu\,, \tag{5.35}$$

which particularly implies that $\mu \in E$. This concludes that each accumulation point must agree with $(\mu^{(\cdot)})^{-1}(\mu) = p$. □

Approximation of infinitesimal dynamics

The key result of this section is the following convergence result on the level of pre-generators.

Proposition 5.2.2. *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L)$ such that $\mathtt{d}L \to \theta \in [0,\infty)$. For every $H \in \mathcal{D}(\mathcal{L}_\theta)$, cf. (5.7), we have with $\mathfrak{L}_{L,N}$ defined in (5.1)*

$$\lim_{N/L\to\rho} \sup_{\eta\in\Omega_{L,N}} \big|\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) - (\mathcal{L}_\theta H)(\mu^{(\eta)})\big| = 0\,. \tag{5.36}$$

We split the proof of Proposition 5.2.2 into two parts. First, we only consider test functions of elementary form $H(\mu) = \mu(h)$, which corresponds to measuring a single observable $h \in C^3([0,1])$. We then extend the convergence result to arbitrary test functions in the domain, which requires to understand correlations between several observables. As usual, it turns out that only pairwise correlations contribute to leading order.

Lemma 5.2.3. *Let $\rho \in (0,\infty)$ and $\mathtt{d} = \mathtt{d}(L)$ such that $\mathtt{d}L \to \theta \in [0,\infty)$. Consider $H \in \mathcal{D}(\mathcal{L}_\theta)$ of the elementary form $H(\mu) = \mu(h)$, for some $h \in C^3([0,1])$. Then*

$$\lim_{N/L\to\rho} \sup_{\eta\in\Omega_{L,N}} \left|\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) - \mu^{(\eta)}(A_\theta h)\right| = 0\,, \tag{5.37}$$

where A_θ is the single-particle generator (5.9).

Proof. Let $h \in C^3([0,1])$ and define $H(\mu) := \mu(h)$. For the sake of convenience we introduce the notation $\tilde{h}(z) := z\,h(z)$. Thus,

$$H(\mu^{(\eta)}) = H(\mu_\#\eta) = \mu^{(\eta)}(h) = \sum_{x=1}^{L} \frac{\eta_x}{N} h\left(\tfrac{\eta_x}{N}\right) = \sum_{x=1}^{L} \tilde{h}\left(\tfrac{\eta_x}{N}\right), \tag{5.38}$$

which allows us to write

$$\begin{aligned} H(\mu_\#\eta^{x,y}) - H(\mu_\#\eta) &= \tilde{h}\left(\tfrac{\eta_y+1}{N}\right) - \tilde{h}\left(\tfrac{\eta_y}{N}\right) + \tilde{h}\left(\tfrac{\eta_x-1}{N}\right) - \tilde{h}\left(\tfrac{\eta_x}{N}\right) \\ &= \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_y}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_y}{N}\right) - \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_x}{N}\right) \\ &\quad + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_x}{N}\right) + \frac{1}{6}\frac{1}{N^3}\tilde{h}'''(\xi)\,, \end{aligned} \tag{5.39}$$

using a second-order Taylor approximation of $\tilde{h}$, with $\xi \in [0,1]$. Therefore, we have uniformly over configurations $\eta \in \Omega_{L,N}$

$$\begin{aligned} \mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \sum_{\substack{x,y\in\Lambda \\ x\neq y}} \eta_x(\mathtt{d}+\eta_y)\Big[&\frac{1}{N}\tilde{h}'\left(\tfrac{\eta_y}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_y}{N}\right) \\ &- \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_x}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_x}{N}\right)\Big] + o(1)\,, \end{aligned}$$

where we used that $\tilde{h}''' \in C([0,1])$. We split the sum into two parts, analysing terms with coefficients $\mathtt{d}\,\eta_x$ and $\eta_x\eta_y$ separately. We begin with the latter:

- The contribution of inclusion rates $\eta_x\eta_y$ is limited to

$$\sum_{\substack{x,y\in\Lambda \\ x\neq y}} \frac{\eta_x}{N}\frac{\eta_y}{N}\tilde{h}''\left(\tfrac{\eta_x}{N}\right) = \sum_{x\in\Lambda} \frac{\eta_x}{N}\Big(1 - \frac{\eta_x}{N}\Big)\tilde{h}''\left(\tfrac{\eta_x}{N}\right), \tag{5.40}$$

  due to exact cancellation of the first-order terms $\tilde{h}'$.

- On the other hand, contributions of the random-walk dynamics, induced

by rates $\mathtt{d}\,\eta_x$, are given by

$$\mathtt{d} \sum_{\substack{x,y\in\Lambda\\ x\neq y}} \frac{\eta_x}{N}\left[\widetilde{h}'\big(\tfrac{\eta_y}{N}\big) + \frac{1}{2}\frac{1}{N}\widetilde{h}''\big(\tfrac{\eta_y}{N}\big) - \widetilde{h}'\big(\tfrac{\eta_x}{N}\big) + \frac{1}{2}\frac{1}{N}\widetilde{h}''\big(\tfrac{\eta_x}{N}\big)\right]$$
$$= \mathtt{d}L \sum_{x\in\Lambda} \frac{\eta_x}{N}\left[\frac{1}{L}\sum_{y\neq x}\widetilde{h}'\big(\tfrac{\eta_y}{N}\big) - \frac{L-1}{L}\widetilde{h}'\big(\tfrac{\eta_x}{N}\big)\right] + o(1)\,, \tag{5.41}$$

because second-order terms $\tilde{h}''$ vanish in the thermodynamic limit due to

$$\left|\mathtt{d} \sum_{\substack{x,y\in\Lambda\\ x\neq y}} \frac{\eta_x}{2N^2}\Big(\tilde{h}''\big(\tfrac{\eta_x}{N}\big) + \tilde{h}''\big(\tfrac{\eta_y}{N}\big)\Big)\right| \leq \mathtt{d}L \sum_{x\in\Lambda} \frac{\eta_x}{N^2}\|\tilde{h}''\|_\infty \leq \frac{\mathtt{d}L}{N}\|\tilde{h}''\|_\infty \to 0\,.$$

Furthermore, we can absorb errors arising from replacing $\frac{L-1}{L}\tilde{h}'$ with $\tilde{h}'$, into $o(1)$.

Now, combining (5.40) and (5.41) yields

$$\begin{aligned}&\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) \\ &= \sum_{x\in\Lambda}\frac{\eta_x}{N}\Big(1-\frac{\eta_x}{N}\Big)\tilde{h}''\big(\tfrac{\eta_x}{N}\big) + \mathtt{d}L\left[\tilde{h}'(0) - \sum_{x\in\Lambda}\frac{\eta_x}{N}\tilde{h}'\big(\tfrac{\eta_x}{N}\big)\right] + o(1)\,,\end{aligned} \tag{5.42}$$

where we additionally used Lemma 5.2.5 below to conclude the uniform approximation $\frac{1}{L}\sum_{y\in\Lambda, y\neq x}\tilde{h}'\big(\frac{\eta_y}{N}\big) = \tilde{h}'(0) + o(1)$. Rewriting (5.42) in terms of $\mu^{(\eta)}$, we have

$$\begin{aligned}&\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) \\ &= \mu^{(\eta)}\big(Z(1-Z)h''(Z) + 2(1-Z)h'(Z) + \mathtt{d}L(h(0) - h(Z) - Zh'(Z))\big) + o(1)\end{aligned} \tag{5.43}$$

where $Z \sim \mu^{(\eta)}$ and we used that $\tilde{h}'(z) = h(z) + zh'(z) = Bh(z)$ and $\tilde{h}''(z) = 2h'(z) + zh''(z) = (Bh)'(z)$. Lastly, since $\|Bh\|_\infty < \infty$ and $\mathtt{d}L \to \theta$, we indeed have

$$\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \mu^{(\eta)}(A_\theta h) + o(1)\,, \tag{5.44}$$

uniformly over all $\eta \in \Omega_{L,N}$. This concludes the proof. □

Remark 5.2.4. *The equivalence to the Poisson-Dirichlet diffusion can already be observed in* (5.42) *when considering h to be of the form $h(z) = z^{m-1}$, $m \geq 2$. In this*

case $H(\mu^{(\eta)}) = \mu^{(\eta)}(h) = \sum_{x=1}^{L} \tilde{h}\big(\frac{\eta_x}{N}\big) = \varphi_m\big(\frac{\eta}{N}\big)$, *cf.* (5.17), *and*

$$\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \mathcal{G}_\theta \varphi_m\big(\tfrac{\hat{\eta}}{N}\big) + o(1)\,. \tag{5.45}$$

In the previous proof, we made use of the fact that a Riemann approximation over rescaled particle configurations is approximately given by evaluation against a point-mass at zero. The following lemma provides a rigorous proof of this fact.

Lemma 5.2.5. *Let* $h \in C(\mathbf{R}_+)$ *and* $\rho \in [0,\infty)$. *Then for any* $\zeta_L \to 0$

$$\lim_{N/L\to\rho} \sup_{\eta\in\Omega_{L,N}} \Big| \frac{1}{L} \sum_{x=1}^{L} h(\zeta_L \eta_x) - h(0) \Big| = 0\,. \tag{5.46}$$

Proof. Let $\varepsilon > 0$ and $\eta \in \Omega_{L,N}$, then

$$\begin{aligned}
&\Big| \frac{1}{L} \sum_{x=1}^{L} h(\zeta_L \eta_x) - h(0) \Big| \\
&\le \Big| \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x > \varepsilon} \big(h(\zeta_L \eta_x) - h(0) \big) \Big| + \Big| \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x \le \varepsilon} \big(h(\zeta_L \eta_x) - h(0) \big) \Big| \\
&\le 2\|h\|_\infty \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x > \varepsilon} + \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x \le \varepsilon} \big| h(\zeta_L \eta_x) - h(0) \big| \\
&\le 2\|h\|_\infty \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x > \varepsilon} + \sup_{0\le v\le\varepsilon} \big| h(v) - h(0) \big| \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L\eta_x \le \varepsilon}\,.
\end{aligned}$$

The first term on the right-hand side vanishes, because the number of sites satisfying $\zeta_L \eta_x > \varepsilon$ is upper bounded by $\zeta_L\, N\, \varepsilon^{-1}$ (otherwise the total mass exceeds N). Thus,

$$\frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\eta_x > \varepsilon\zeta_L^{-1}} \le \frac{\zeta_L\, N}{\varepsilon L}\,. \tag{5.47}$$

The second term, on the other hand, is upper bounded by $\sup_{0\le v\le\varepsilon} \big| h(v) - h(0) \big|$, which vanishes in the small ε-limit. Note that both upper bounds are uniform in $\Omega_{L,N}$. Now, taking first the thermodynamic limit $N/L \to \rho$ before taking $\varepsilon \to 0$, finishes the proof. □

After having proved the statement of Proposition 5.2.2 for specific functions, we can proceed with the proof of the full statement.

Proof of Proposition 5.2.2. Let $H \in \mathcal{D}(\mathcal{L}_\theta)$. We assume H has the form

$$H(\mu) = \mu(h_1) \cdots \mu(h_n)\,, \quad h_k \in C^3([0,1])\,, \quad 1 \le k \le n\,, \tag{5.48}$$

since linear combinations of such functions can be treated by linearity of the operators and the triangle inequality. Thus, considering $\eta \in \Omega_{L,N}$ and the configuration after one particle jumped from x to y, we have

$$\begin{aligned} H\left(\mu_\# \eta^{x,y}\right) &= \prod_{k=1}^{n} \left(\mu_\# \eta^{x,y}\right)(h_k) \\ &= \prod_{k=1}^{n} \left[\widetilde{h}_k\left(\tfrac{\eta_y+1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_y}{N}\right) + \widetilde{h}_k\left(\tfrac{\eta_x-1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_x}{N}\right) + \mu^{(\eta)}(h_k)\right]. \end{aligned}$$

Now, expanding the product yields

$$\begin{aligned} &H\left(\mu_\# \eta^{x,y}\right) \\ &= H\left(\mu_\# \eta\right) + \sum_{k=1}^{n} \left[\widetilde{h}_k\left(\tfrac{\eta_y+1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_y}{N}\right) + \widetilde{h}_k\left(\tfrac{\eta_x-1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_x}{N}\right)\right] \prod_{\substack{l=1 \\ l\neq k}}^{n} \mu^{(\eta)}(h_l) \\ &\quad + \sum_{1\le k<l\le n} \left[\widetilde{h}_k\left(\tfrac{\eta_y+1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_y}{N}\right) + \widetilde{h}_k\left(\tfrac{\eta_x-1}{N}\right) - \widetilde{h}_k\left(\tfrac{\eta_x}{N}\right)\right] \\ &\qquad \times \left[\widetilde{h}_l\left(\tfrac{\eta_y+1}{N}\right) - \widetilde{h}_l\left(\tfrac{\eta_y}{N}\right) + \widetilde{h}_l\left(\tfrac{\eta_x-1}{N}\right) - \widetilde{h}_l\left(\tfrac{\eta_x}{N}\right)\right] \prod_{\substack{j=1 \\ j\neq k,l}}^{n} \mu^{(\eta)}(h_j) + r(\eta)\,, \end{aligned} \tag{5.49}$$

with r denoting the remainder. This expansion allows to split

$$\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(\mathtt{d}+\eta_y)\left[H(\mu_\#\eta^{x,y}) - H(\mu_\#\eta)\right] \tag{5.50}$$

into three parts:

- First, we make use of Lemma 5.2.3 which yields

$$\begin{aligned} &\sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(\mathtt{d}+\eta_y) \sum_{k=1}^{n} [\widetilde{h}_k(\tfrac{\eta_y+1}{N}) - \widetilde{h}_k(\tfrac{\eta_y}{N}) + \widetilde{h}_k(\tfrac{\eta_x-1}{N}) - \widetilde{h}_k(\tfrac{\eta_x}{N})] \prod_{\substack{l=1 \\ l\neq k}}^{n} \mu^{(\eta)}(h_l) \\ &\quad = \sum_{k=1}^{n} \mathfrak{L}_{L,N}\left(\mu^{(\cdot)}(h_k)\right)(\eta) \prod_{\substack{l=1 \\ l\neq k}}^{n} \mu^{(\eta)}(h_l) \end{aligned}$$

$$= \sum_{k=1}^{n} \mu^{(\eta)}(A_\theta h_k) \prod_{\substack{l=1 \\ l\neq k}}^{n} \mu^{(\eta)}(h_l) + o(1)\,.$$

- Next, we prove that the remainder r has no contribution. More precisely, for any non-negative sequence a_N satisfying $N^2\, a_N \to 0$, i.e. a_N lies in $o(\frac{1}{N^2})$, we have

$$a_N \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(\mathtt{d}+\eta_y) \leq a_N\,(\mathtt{d}L+N)N \to 0\,. \tag{5.51}$$

  This includes, in particular, the remainder $r(\eta)$ because each summand lies in $O(\frac{1}{N^3})$, recall that each square bracket in (5.49) vanishes uniformly like N^{-1}, cf. (5.39).

- Lastly, we derive the interaction part where two observables are affected by the transition of a particle. Again, we perform a Taylor approximation for each of the two square brackets appearing in (5.49). Due to (5.51), together with (5.39), it suffices to consider only products of first-order terms $\tilde{h}'$. Therefore, we are left with

$$\frac{1}{N^2} \sum_{1\leq k<l\leq n} \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(\mathtt{d}+\eta_y) \left[\tilde{h}_k'(\tfrac{\eta_y}{N}) - \tilde{h}_k'(\tfrac{\eta_x}{N})\right] \left[\tilde{h}_l'(\tfrac{\eta_y}{N}) - \tilde{h}_l'(\tfrac{\eta_x}{N})\right] \prod_{\substack{j=1 \\ j\neq k,l}}^{n} \mu^{(\eta)}(h_j)$$

  plus some vanishing term, which we will denote by $o(1)$. For the same reason, we include the random-walk interactions coming from $\mathtt{d}\,\eta_x$ in $o(1)$, and finally arrive at

$$\begin{aligned}
&\sum_{1\leq k<l\leq n} \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \frac{\eta_x}{N}\frac{\eta_y}{N} \left[\tilde{h}_k'(\tfrac{\eta_y}{N}) - \tilde{h}_k'(\tfrac{\eta_x}{N})\right] \left[\tilde{h}_l'(\tfrac{\eta_y}{N}) - \tilde{h}_l'(\tfrac{\eta_x}{N})\right] \prod_{\substack{j=1 \\ j\neq k,l}}^{n} \mu^{(\eta)}(h_j) + o(1) \\
&\quad = 2 \sum_{1\leq k<l\leq n} \left(\mu^{(\eta)}(\tilde{h}_k'\tilde{h}_l') - \mu^{(\eta)}(\tilde{h}_k')\mu^{(\eta)}(\tilde{h}_l')\right) \prod_{\substack{j=1 \\ j\neq k,l}}^{n} \mu^{(\eta)}(h_j) + o(1)\,,
\end{aligned}$$

  where we expanded the product of square brackets and added the (non-contributing) diagonal $x = y$, before writing the expression in terms of $\mu^{(\eta)}$. Also, recall that $\tilde{h}' = Bh$.

Overall, combining the three bullets above, we conclude

$$\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \mathcal{L}_\theta H(\mu^{(\eta)}) + o(1)\,, \tag{5.52}$$

uniformly in $\eta \in \Omega_{L,N}$. This finishes the proof. □

Convergence to the measure-valued process

The measure-valued process takes values in the space $E = \mu^{(\overline{\nabla})} \subset \mathcal{M}_1([0,1])$, cf. (5.5). Due to Lemma 5.2.1, E itself is closed, thus, compact, when considering the topology induced by weak convergence. Because this topology coincides with the subspace topology, the Hausdorff property of E is inherited from $\mathcal{M}_1([0,1])$.

In this section, we show that the dynamics described by $\mathcal{L}_\theta$ give rise to a Feller process and prove Theorem 5.1.1, which states that the process arises naturally as the scaling limit of the inclusion process.

Proposition 5.2.6. *For $\theta \in [0,\infty)$ the linear operator $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ is closable and its closure generates a Feller process on the state space $E \subset \mathcal{M}_1([0,1])$.*

The proof follows along the lines of [EK81, Theorem 2.5], where existence of the Poisson-Dirichlet diffusion is proven.

Proof. Throughout the proof, we will make use of the sub-domain

$$\mathcal{D}_{mon}(\mathcal{L}_\theta) := \Big\{ \text{sub-algebra of } C(E) \text{ generated by functions } \mu \mapsto \mu(h) \text{ with } h(z) = z^m\,, m \in \mathbf{N}_0 \Big\} \subset \mathcal{D}(\mathcal{L}_\theta)\,. \tag{5.53}$$

First note that, due to the Stone-Weierstrass theorem, $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ (and therefore $\mathcal{D}(\mathcal{L}_\theta)$) is dense in $C(E)$, since it separates points: Consider $\mu, \sigma \in E$ such that $\mu \neq \sigma$, then $\mu(z^m) \neq \sigma(z^m)$ for some $m \in \mathbf{N}$ since otherwise all moments, and hence μ and σ, agree.

Next, dissipativity of $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ follows from $(\mathfrak{L}_{L,N})_{L,N}$, since for any $H \in \mathcal{D}(\mathcal{L}_\theta)$ we have

$$\|(\lambda - \mathfrak{L}_{L,N}) H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty} \geq \lambda \|H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty} \quad \forall \lambda > 0\,, \tag{5.54}$$

with $\|\cdot\|_{\Omega_{L,N},\infty}$ denoting the supremum norm on $C(\Omega_{L,N})$. The left-hand side

is upper bounded by

$$\|(\lambda - \mathcal{L}_\theta)H\|_{E,\infty} + \|\mathcal{L}_\theta H(\mu^{(\cdot)}) - \mathfrak{L}_{L,N}H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty}\,, \tag{5.55}$$

since $\{\mu^{(\eta)} : \eta \in \Omega_{L,N}\} \subset E$, with $\|\cdot\|_{E,\infty}$ denoting the supremum norm on $C(E)$. The second term vanishes due to Proposition 5.2.2. On the other hand, using Lemma 5.2.7 below, we have

$$\|H\|_{\Omega_{L,N},\infty} = \sup_{\eta\in\Omega_{L,N}} |H(\mu^{(\eta)})| \to \sup_{p\in\overline{\nabla}} |H(\mu^{(p)})| = \|H\|_{E,\infty}\,. \tag{5.56}$$

In the remainder of the proof, we first conclude that $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ is a core for $\mathcal{L}_\theta$, using the fact that $\mathcal{L}_\theta$ is triangulisable. Then, the full statement follows immediately by an extension argument. For that purpose, we define subspaces

$$D_n(\mathcal{L}_\theta) := \{H \in \mathcal{D}_{mon}(\mathcal{L}_\theta) : deg(H) \leq n\}, \tag{5.57}$$

where $deg(H) = m_1 + \cdots + m_k$ if H is of the form $\mu(z^{m_1}) \cdots \mu(z^{m_k})$, $m_j \in \mathbf{N}$ for $1 \leq j \leq k$. When H is given by linear combinations of such products, the degree denotes the maximum degree of the products. Note that $\big(D_n(\mathcal{L}_\theta)\big)_{n\geq 1}$ defines an increasing sequence with limit $\mathcal{D}_{mon}(\mathcal{L}_\theta)$. It is only left to show that $\mathcal{L}_\theta$ maps elements of $D_n(\mathcal{L}_\theta)$ back into itself. This is, however, immediate since both parts of the generator $\mathcal{L}_\theta$ (5.8) map polynomials of a certain degree back into polynomials of the same degree, i.e. $\mathcal{L}_\theta$ is triangulisable. By [EK86, Proposition 1.3.5] this is enough to conclude that $(\mathcal{L}_\theta, \mathcal{D}_{mon}(\mathcal{L}_\theta))$ is indeed closable, and gives rise to a strongly continuous contraction semigroup $(T_t)_{t\geq 0}$ on $C(E)$.

We can verify that $\mathcal{D}(\mathcal{L}_\theta)$ is a core, by noting that

$$\mathcal{R}ange\big(\lambda - \mathcal{L}_\theta|_{\mathcal{D}(\mathcal{L}_\theta)}\big) \supset \mathcal{R}ange\big(\lambda - \mathcal{L}_\theta|_{\mathcal{D}_{mon}(\mathcal{L}_\theta)}\big) \tag{5.58}$$

is dense in $C(E)$ for some $\lambda > 0$, [EK86, Proposition 1.3.1]. Moreover, generators are maximal dissipative, hence, we know that the closures of both cores $\mathcal{D}(\mathcal{L}_\theta)$ and $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ must agree, see e.g. [EK86, Proposition 1.4.1], and gives rise to the same semigroup $(T_t)_{t\geq 0}$. It is only left to show that the semigroup is positive and conservative, in particular E is invariant under the dynamics $\mathcal{L}_\theta$. In order to see this, we apply Proposition 5.1.7, which concludes that convergence of generators in Proposition 5.2.2 implies for all $H \in C(E)$ and

$t \geq 0$

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \mathfrak{T}_t^{(L,N)} H(\mu^{(\cdot)})(\eta) - (T_t H)(\mu^{(\eta)}) \right| = 0\,, \tag{5.59}$$

where $\mathfrak{T}^{(L,N)}$ denotes the semigroup generated by $\mathfrak{L}_{L,N}$. Now, both positivity and conservation follow from $(\mathfrak{T}^{(L,N)})_{L,N}$. This concludes the proof. □

It is natural to ask why going through the inconveniences of extending the core from $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ to $\mathcal{D}(\mathcal{L}_\theta)$ is necessary. We will see in the next section, that the extended core allows for a better interpretation of the underlying dynamics in the Poisson-Dirichlet diffusion.

The missing component in the proof above is that any element in $\overline{\nabla}$ can be approximated by particle configurations, independently of the thermodynamic limit considered.

Lemma 5.2.7. *Let $\frac{N}{L} \to \rho \in [0, \infty]$, then for any $p \in \overline{\nabla}$ there exist $\eta^{(L,N)} \in \Omega_{L,N}$ such that*

$$\frac{1}{N}\hat{\eta}^{(L,N)} = \frac{1}{N}\eta^{(L,N)} \to p\,, \quad \text{in } \overline{\nabla}\,.$$

Proof. Consider $p \in \overline{\nabla}$ with $\|p\|_1 = \sum_{i=1}^{\infty} p_i = 1 - \gamma$, we then define $\bar{\eta}_i^{(L,N)} := \lfloor p_i N \rfloor$, for $i \in \{1, \dots, L\}$. Hence, there are

$$M_{L,N}(p) := N - \sum_{x=1}^{L} \bar{\eta}_x^{(L,N)} = \gamma N + \sum_{i=1}^{L} (p_i N - \lfloor p_i N \rfloor) \leq \gamma N + L$$

particles to spare. Thus, defining $\eta^{(L,N)} \in \Omega_{L,N}$ via

$$\eta_x^{(L,N)} := \bar{\eta}_x^{(L,N)} + \left\lfloor \frac{M_{L,N}(p)}{L} \right\rfloor + \mathbb{1}_{x \leq (M_{L,N}(p) \mod L)}\,, \quad x \in \{1, \dots, L\}\,,$$

yields the desired approximation, since $\frac{M_{L,N}(p)}{NL} \to 0$, as $N, L \to \infty$. □

The proof of convergence of the inclusion process (when embedded in the space of probability measures) to the measure-valued process characterised by $\mathcal{L}_\theta$, is now an immediate consequence of a classical convergence theorem.

Proof of Theorem 5.1.1. We apply Proposition 5.1.8, which states that convergence of the semigroups in (5.59) implies convergence of the corresponding processes. This immediately concludes the desired convergence result. □

We can now use Theorem 5.1.1, to prove convergence of the inclusion process to the Poisson-Dirichlet diffusion.

Proof of Corollary 5.1.4. Every function φ_m can be written in terms of an expectation

$$p \mapsto \mu^{(p)}(h_m) = \varphi_m(p)\,,$$

with $h_m(z) := z^{m-1}$. Thus, we have

$$\Big(\varphi_m\big(\tfrac{1}{N}\hat{\eta}^{(L,N)}(t)\big)\Big)_{t\geq 0} = \Big((\mu_\# \eta^{(L,N)}(t))(h_m)\Big)_{t\geq 0} \xrightarrow{d} \big(\mu_t(h_m)\big)_{t\geq 0}\,, \tag{5.60}$$

using Theorem 5.1.1. This convergence can be extended to arbitrary elements in $\mathcal{D}_{mon}(\mathcal{G}_\theta)$, in particular such sequences are tight. Thus, the sequence $\big(\big(\frac{1}{N}\hat{\eta}^{(L,N)}(t)\big)_{t\geq 0}\big)_{L,N}$ is tight and has subsequential limits, see e.g. [EK86, Theorem 3.9.1]. As convergence of finite dimensional marginals follows from (5.60), we conclude the statement together with Proposition 5.1.3, which is proved in the next subsection. □

5.2.2 Equivalence of the superprocess with PD-diffusion

In this section, we prove Proposition 5.1.3 and investigate the equivalence of the measure-valued process generated by $\mathcal{L}_\theta$ (5.8) and the Poisson-Dirichlet diffusion on the simplex $\overline{\nabla}$, generated by $\mathcal{G}_\theta$ (5.16). We already saw in the proof of Lemma 5.2.3, cf. (5.42), the similarity of dynamics $\mathcal{L}_\theta$ and $\mathcal{G}_\theta$. Indeed, a simple calculation shows that the two can be linked: Using the embedding (5.6), we get for all $p \in \overline{\nabla}$ and $H(\mu) = \mu(h)$, with $h \in \mathcal{D}(A)$,

$$\begin{aligned}
&\mathcal{L}_\theta H(\mu^{(p)}) = \mu^{(p)}(A_\theta h)\\
&= \big(1 - \|p\|_1\big) A_\theta h(0)\\
&\quad + \sum_{i=1}^{\infty} p_i \Big(p_i(1-p_i)h''(p_i) + \big(2(1-p_i) - \theta p_i\big)h'(p_i) + \theta(h(0) - h(p_i))\Big)\,.
\end{aligned}$$

Defining now $f(p) := \mu^{(p)}(h)$, we have

$$\mathcal{L}_\theta H(\mu^{(p)}) = \mathcal{G}_\theta f(p) + 2h'(0)(1 - \|p\|_1)\,, \tag{5.61}$$

where we used that $\partial_{p_i} f(p) = -h(0) + p_i h'(p_i) + h(p_i)$ and $\partial_{p_i p_j} f(p) = \mathbb{1}_{i=j}\big(2h'(p_i) + p_i h''(p_i)\big)$.

Remark 5.2.8. *In [EK81], the domain of $\mathcal{G}_\theta$ was extended from $\mathcal{D}(\mathcal{G}_\theta)$ to the subalgebra of $C(\overline{\nabla})$ generated by functions of the form $p \mapsto \sum_{i=1}^{\infty} g(p_i)$, with $g \in C^2([0,1])$ such that $g(0) = g'(0) = 0$. This yields a similar expression as (5.61), cf. [EK81, Display (2.17)]. However, the expression again only made sense with the convention that sums are evaluated on ∇ and extended by continuity, in which case the second summand in (5.61) disappears.*

Proof of Proposition 5.1.3. In order to show the equivalence of the two processes, it suffices to restrict ourselves to the domains generated by monomials, as defined in (5.17) and (5.53).

First note that every function $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ can be mapped naturally to a $f_H \in \mathcal{D}_{mon}(\mathcal{G}_\theta)$ (and vise versa): Let $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ be of the form $H(\mu) = \mu(h_1) \cdots \mu(h_n)$, with $h_k(z) := z^{m_k - 1}$, then $f_H = \varphi_{m_1} \cdots \varphi_{m_n}$ where we recall $\varphi_m(p) = \sum_{i=1}^{\infty} p_i^m$. Then

$$\begin{aligned}(\mathcal{L}_\theta H)(\mu^{(p)}) = 2 \sum_{1 \le k < l \le n} m_k m_l \big(\mu^{(p)}(h_k h_l) - \mu^{(p)}(h_k)\mu^{(p)}(h_l)\big) \prod_{j \neq k,l} \mu^{(p)}(h_j) \\ + \sum_{1 \le k \le n} \mu^{(p)}(A_\theta h_k) \prod_{j \neq k} \mu^{(p)}(h_j)\,, \end{aligned} \tag{5.62}$$

where we used that $Bh_k = m_k h_k$. Rewriting the right-hand side in terms of φ's, we have

$$\begin{aligned}(\mathcal{L}_\theta H)(\mu^{(\cdot)}) = 2 \sum_{1 \le k < l \le n} m_k m_l \big(\varphi_{m_k + m_l - 1} - \varphi_{m_k}\varphi_{m_l}\big) \prod_{j \neq k,l} \varphi_{m_j} \\ + \sum_{1 \le k \le n} \mathcal{G}_\theta \varphi_{m_k} \prod_{j \neq k} \varphi_{m_j}\,, \end{aligned}$$

where we used $\mu^{(p)}(A_\theta h_k) = \mathcal{G}_\theta \varphi_{m_k}$ from (5.61). Thus, $(\mathcal{L}_\theta H)(\mu^{(\cdot)})$ agrees with $\mathcal{G}_\theta f_H$ on $\mathcal{D}_{mon}(\mathcal{G}_\theta)$, cf. [EK81, Display (2.13)]. Let $(X_t)_{t \ge 0}$ be the Poisson-Dirichlet diffusion, then for every $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$

$$H(\mu^{(X_t)}) - \int_0^t \mathcal{L}_\theta H(\mu^{(X_s)})\, \mathrm{d}s = f_H(X_t) - \int_0^t \mathcal{G}_\theta f_H(X_s)\, \mathrm{d}s \tag{5.63}$$

defines a martingale in t. Thus, $(\mu^{(X_t)})_{t \ge 0}$ solves the martingale problem for $(\mathcal{L}_\theta, \mathcal{D}_{mon}(\mathcal{L}_\theta))$.

Now, it is almost immediate that properties (i) – (iii) in Proposition 5.1.3 hold for the measure valued process $(\mu_t)_{t \ge 0}$. First, let $G, H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ and choose corresponding $f_G, f_H \in \mathcal{D}_{mon}(\mathcal{G}_\theta)$ as above. We know that $(\mathcal{L}_\theta G)(\mu^{(p)}) = \mathcal{G}_\theta f_G(p)$. Writing $\nu = \mathrm{PD}(\theta)$ for simplicity, we have for

$\mathbf{P} := \mu_{\#}\nu$

$$\mathbf{P}(H\mathcal{L}_\theta G) = \nu\big(H(\mu^{(p)})(\mathcal{L}_\theta G)(\mu^{(p)})\big) = \nu(f_H \mathcal{G}_\theta f_G)\,. \tag{5.64}$$

It is known that $\nu = \mathrm{PD}(\theta)$ is the unique invariant distribution with respect to $\mathcal{G}_\theta$ [EK81, Theorem 4.3]; which is also reversible. Together with the above display, this yields $\mathbf{P}(H\mathcal{L}_\theta G) = \mathbf{P}(G\mathcal{L}_\theta H)$.

Continuity of the trajectories in (ii) follows from the diffusion property of $(X_t(\omega))_{t\geq 0}$ and continuity of the map $\mu^{(\cdot)}$.

Lastly, (iii) follows because

$$\begin{aligned}\mathbb{P}\left(\mu_t(\{0\}) = 0 \quad \forall t > 0\right) &= \mathbb{P}\left(\mu^{(X_t)}(\{0\}) = 0 \quad \forall t > 0\right)\\ &= \mathbb{P}\left(X_t \in \nabla \quad \forall t > 0\right) = 1\,,\end{aligned}$$

where the last equality is a consequence of [EK81, Theorem 2.6]. □

Remark 5.2.9. *Naturally, we could have proven convergence of the inclusion process to the Poisson-Dirichlet diffusion directly and then defined the measure-valued dynamics using the embedding via $\mu^{(\cdot)}$. This would have slightly shortened the exposition in the present section, since it would not have been necessary to verify existence of the limiting dynamics. We refrained from doing so, for a better understanding of the underlying dynamics of the measure-valued process, in particular on the extended domain $\mathcal{D}(\mathcal{L}_\theta)$.*

We conclude this section with a brief study of the single-particle dynamics (5.9):

$$A_\theta h(z) := z(1-z)h''(z) + (2(1-z) - \theta z)h'(z) + \theta(h(0) - h(z))\,,$$

which characterises a Feller process on the unit interval. The process evolves according to a diffusion with an additional renewal mechanism, due to jumps to zero.

Lemma 5.2.10. *The Beta distribution* $\mathrm{Beta}(1,\theta)$ *is the unique invariant distribution with respect to A_θ.*

Proof. For $\theta = 0$, we interpret the degenerate distribution $\mathrm{Beta}(1,0)$ as the Dirac point mass δ_1. The statement is then clear since $A_\theta h(1) = 0$, in this case the point mass is even reversible.

Now, let $\theta > 0$ and consider $H(\mu) := \mu(h)$, then by Proposition 5.1.3(i)

$$0 = \mathbf{P}(\mathcal{L}_\theta H) = \int \mu(A_\theta h)\, \mathbf{P}(\,\mathrm{d}\mu) = \mathbb{E}\Big[\sum_{i=1}^{\infty} X_i A_\theta h(X_i)\Big] = \mathbb{E}[A_\theta h(\tilde{X}_1)]\,,$$

where $X \sim \mathrm{PD}(\theta)$. It is well known that the first size-biased marginal $\tilde{X}_1$ is $\mathrm{Beta}(1,\theta)$-distributed [Fen10, Theorem 2.7].

Uniqueness of the invariant distribution is due to Harris recurrence of the process, see e.g. [MT93]. For the case $\theta > 0$, the resetting mechanism guarantees that the process returns to zero infinitely often almost surely. On the other hand for $\theta = 0$, A_θ agrees with a Jacobi diffusion, see (5.66) below, and the corresponding process runs into the absorbing state $z = 1$ in finite time, independent of the initial condition. □

Due to the jumps to zero, one does not expect that $\mathrm{Beta}(1,\theta)$, $\theta > 0$, is reversible with respect to A_θ. Indeed, this can be verified easily by considering the example $h(x) = x$ and $g(x) = x^2$, in which case for every $\theta > 0$

$$\mathrm{Beta}(1,\theta)(gA_\theta h) = \frac{8\theta\Gamma(\theta+1)}{\Gamma(\theta+4)} \neq \frac{6\theta\Gamma(\theta+1)}{\Gamma(\theta+4)} = \mathrm{Beta}(1,\theta)(hA_\theta g)\,.$$

5.2.3 The advantage of a size-biased evolution

Throughout the previous sections, we have seen two viewpoints of the same dynamics. The classical Poisson-Dirichlet diffusion considers a ranked configuration space. However, this obscures the dynamics on microscopic scales, which results in defining the right-hand side of the generator $\mathcal{G}_\theta$ (5.16) to be "evaluated on ∇ and extended to $\overline{\nabla}$ by continuity". Alternatively, one can consider unordered dynamics, i.e. observing the evolution from a fixed position, or a size-biased viewpoint.

The Poisson-Dirichlet diffusion concentrates immediately on configurations with macroscopic-sized fragments (condensation occurs), which can only be supported on a vanishing fraction of the volume. Hence, an unordered state space can only describe dynamics up to a certain point, at which the mass at the observed positions disappears. The goal of this section is to emphasise that a size-biased viewpoint allows for both, a complete description of the macroscopic dynamics, while observing interaction with the microscopic scale.

Time evolution on fixed sites

First, we look at arbitrary finite positions and observe the evolution of mass on them. As the inclusion process is spatially homogeneous, we may choose for simplicity $\eta \mapsto (\eta_1, \dots, \eta_n)$.

We start by only considering the evolution on the first site. Performing similar approximations as in Section 5.2, we can see for an arbitrary function $h \in C^3([0,1])$ that

$$\mathfrak{L}_{L,N} h(\tfrac{(\cdot)_1}{N})(\eta) = A_{Jac(\theta)} h(\tfrac{\eta_1}{N}) + o(1)\,, \tag{5.65}$$

where

$$A_{Jac(\theta)} h(z) := z(1-z)h''(z) - \theta z\, h'(z)\,. \tag{5.66}$$

In fact, $A_{Jac(\theta)}$ emerges from (5.42) when fixing a position x. The operator $A_{Jac(\theta)}$ is the generator of a Jacobi-diffusion, cf. [FPRW23], and describes the evolution of a single chunk of mass located at a given position.

To describe the evolution on the first n positions we introduce for $i = 0, 1, \dots, n$

$$\xi_i = \xi_i(\eta) := \begin{cases} \eta_i & \text{if } 1 \le i \le n \\ N - \sum_{j=1}^n \xi_j & \text{if } i = 0\,, \end{cases} \tag{5.67}$$

where ξ_0 denotes the remaining mass in the system outside locations $1, \dots, n$. Thus, the rescaled vector $\frac{1}{N}\xi$ lies in $\Delta_{n+1} := \{p \in [0,1]^{n+1} : \sum_{i=0}^n p_i = 1\}$. Again, by approximation of the generators, one can show that

$$\tfrac{1}{N}(\xi(t))_{t\ge 0} \xrightarrow{d} \mathrm{WF}_{n+1}(\theta, 0, \dots, 0)\,. \tag{5.68}$$

Here, $\mathrm{WF}_{n+1}(\theta, 0, \dots, 0)$ denotes the Wright-Fisher diffusion on Δ_{n+1} which is characterised by the generator

$$\begin{aligned} A_{WF_{n+1}(\theta,\mathbf{0})} h(z_{0,\dots,n}) = &\sum_{i,j=0}^n z_i(\mathbb{1}_{i=j} - z_j)\partial^2_{z_i z_j} h(z_{0,\dots,n}) \\ &+ \theta \sum_{i=1}^n z_i(\partial_{z_0} h - \partial_{z_i} h)(z_{0,\dots,n})\,, \end{aligned} \tag{5.69}$$

acting on those h that have an extension to $\mathbf{R}^{n+1}$ which is twice continuously differentiable.

For a single observable, the Jacobi-diffusion has an absorbing state at $z = 0$, which it runs into in finite time almost surely [Dur08, Section 7.10]. Similarly, the Wright-Fisher diffusion will be absorbed at $(1, 0, \dots 0) \in \Delta_{n+1}$, after which the process does not capture the dynamics of the infinite-dimensional process anymore as all the mass has moved away from the first n sites.

For the Poisson-Dirichlet diffusion, the relationship to the Jacobi and Wright-Fisher diffusion has been studied in the two-parameter setting [FPRW23] in greater generality. There a Fleming-Viot construction of the process, cf. (5.32), is used. Because they start from the Poisson-Dirichlet diffusion on $\overline{\nabla}$, there is no underlying graph structure and instead of placing mass at a fixed position, they choose a uniform random variable on $[0, 1]$ which determines the position of the point mass.

It is interesting to note that the boundary behaviour of the Jacobi diffusion agrees with the one of the PD-diffusion. More precisely, in the case of the Jacobi diffusion the state 1 can be reached if and only if $\theta < 1$, see e.g. [Shi81, Theorem 4.1] or [Dur08, Section 7.10]. Similarly, the PD-diffusion $(X_t)_{t\geq 0}$ hits the finite dimensional sub-simplices $\nabla \cap \{\sum_{i=1}^{n} p_i = 1\}$, for any $n \geq 1$, if and only if $\theta < 1$ [Sch91].

Duality and size-biased time evolution

For a measure valued process with generator $\mathcal{L}_\theta$ (5.8), the evolution with respect to a simple test function $H(\mu) = \mu(h)$ is given by Dynkin's formula

$$\mathrm{d}\mu_t(h) = \mu_t(A_\theta h)\ \mathrm{d}t + \mathrm{d}M_t^{(h)}\,, \quad \forall h \in \mathcal{D}(A)\,. \tag{5.70}$$

Here, $(M_t^{(h)})_{t\geqslant 0}$ is a martingale. Continuity of the process $(\mu_t)_{t\geq 0}$, and thus of the martingale $(M_t^{(h)})_{t\geq 0}$, follows from the equivalence with the Poisson-Dirichlet diffusion in Proposition 5.1.3. Now, taking expectations in (5.70), we see that $\bar{\mu}_t := \mathbb{E}_{\mu_0}[\mu_t]$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mu}_t(h) = \bar{\mu}_t(A_\theta h)\,, \quad \forall h \in \mathcal{D}(A)\,,$$

which is solved by the time evolution of $(\mathbf{E}_{\mu_0}[h(Z_t)])_{t\geqslant 0}$, for a Feller process $(Z_t)_{t\geq 0}$ on $[0, 1]$ with generator A_θ and initial distribution μ_0. As a consequence, we have the following duality

$$\mathbb{E}_{\mu_0}[\mu_t(h)] = \mathbf{E}_{\mu_0}[h(Z_t)]\,, \quad \forall h \in \mathcal{D}(A)\,, \tag{5.71}$$

where $(Z_t)_{t\geq 0}$ is the one-dimensional diffusion with resetting to 0, generated by A_θ. The identity can be extended to all $h \in C([0,1])$ by standard arguments, see [EK93, Section 6] or the proof of Proposition 5.3.5 below.

In analogy to known duality properties of the microscopic particle system [GKR07, CGR21], this can be used to get closed evolution equations for moments. For example, $h(z) = z$ describes the expected second moment of the mass distribution and we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_{\mu_0}[\mu_t(z)] &= \mathbf{E}_{\mu_0}[A_\theta \mathrm{Id}(Z_t)] \\ &= \mathbf{E}_{\mu_0}\Big[2\big(1-(1+\theta)Z_t\big)\Big] = 2-2(1+\theta)\mathbb{E}_{\mu_0}[\mu_t(Z)]\,, \end{aligned} \tag{5.72}$$

which has an exponential solution that converges to the stationary point $\frac{1}{1+\theta}$, the expected second moment of the GEM(θ) distribution, in accordance with Proposition 5.1.3(i). The duality in (5.71) is also interesting from a computational point of view, as it allows to continuously track the expected behaviour of the infinite-dimensional process. For this, only a finite-dimensional diffusion is necessary, which does not run into any absorbing states (in contrast when observing a fixed set of lattice sites), e.g. (5.72).

Dualities for measure-valued processes were previously considered in [DH82] and [DK99], see also [EK93, Section 6] for a summary. We stress once more the difference in point of view: In previous works, the dual particles encode the position of clusters on the underlying lattice. On the other hand, in our size-biased approach the state of dual particles characterises the fragmentation of mass in a given configuration/partition. In particular, dual particles take into account microscopic states, due to the resetting to 0.

In Proposition 5.3.5, we will see that for $\mathtt{d}L \to \infty$ such dualities extend directly to nonlinear test functions $H(\mu)$, due to the absence of an interaction part in the generator $\hat{\mathcal{L}}$ (5.15).

5.2.4 Convergence to a Fleming-Viot process

This section is an extended remark on how to prove convergence of the inclusion process to the Fleming-Viot process (5.30), with mutation operator (5.31)

$$A_{FV}h(u) = \theta \int_0^1 [h(w) - h(u)]\ \mathrm{d}w\,.$$

Instead of considering the generator $\mathfrak{L}_{L,N}$ (5.1) on $\Omega_{L,N}$, it is more convenient to define the inclusion process on a state space that keeps track of particle positions. More precisely, we define

$$S_{L,N} := \Lambda_L^N\,, \quad \text{with} \quad \Lambda_L := \{1,\dots,L\}\,.$$

Now, the (labelled) inclusion process is described by the infinitesimal generator

$$\mathfrak{G}_{L,N}g(\sigma) = \sum_{i,j=1}^{N}[g(\sigma^{i\to\sigma_j}) - g(\sigma)] + \mathtt{d}\sum_{i=1}^{N}\sum_{x\in\Lambda}[g(\sigma^{i\to x}) - g(\sigma)]\,, \tag{5.73}$$

where

$$\sigma_j^{i\to x} := \begin{cases} x & \text{if } i = j\,,\\ \sigma_j & \text{if } i \neq j\,, \end{cases} \tag{5.74}$$

denotes the updated position after the i-th particle jumped onto site x. The generator $\mathfrak{G}_{L,N}$ characterises a Markov process on $S_{L,N}$, which we denote by $(\sigma^{(L,N)}(t))_{t\geqslant 0}$. We can recover the corresponding unlabelled particle configuration using

$$\iota : S_{L,N} \to \Omega_{L,N} \quad \text{with} \quad \iota(\sigma)_x := \sum_{i=1}^{N}\mathbb{1}_{\sigma_i = x}\,,$$

in particular we have $\mathfrak{L}_{L,N}f(\iota(\sigma)) = \mathfrak{G}_{L,N}f(\iota(\cdot))(\sigma)$, for all $f \in C(\Omega_{L,N})$.

Once more, we interpret particle configurations as probability measures on $[0,1]$. However, now we consider the embedding

$$\nu^{(\cdot)} : \sigma \mapsto \frac{1}{N}\sum_{i=1}^{N}\delta_{\frac{\sigma_i}{L}} \in \mathcal{M}_1([0,1])\,, \tag{5.75}$$

where rescaled particle locations are encoded on the "type space" $[0,1]$. The convergence of processes, under the embedding (5.75), follows from the approximation of generators in analogy to our main result. Again, we start with test functions of the form

$$\nu^{(\sigma)}(h) = \frac{1}{N}\sum_{i=1}^{N}h\left(\tfrac{\sigma_i}{L}\right)\,, \quad h \in C^3([0,1])\,,$$

in which case the action of $\mathfrak{G}_{L,N}$ reads

$$\mathfrak{G}_{L,N}\nu^{(\sigma)}(h) = \sum_{i,j=1}^{N} \frac{1}{N}\big[h(\tfrac{\sigma_j}{L}) - h(\tfrac{\sigma_j}{L})\big] + \mathtt{d}\sum_{i=1}^{N}\sum_{x\in\Lambda}\frac{1}{N}\big[h(\tfrac{x}{L}) - h(\tfrac{\sigma_i}{L})\big]\,.$$

Because the first sum on the right-hand side vanishes by symmetry, we are only left with

$$\mathfrak{G}_{L,N}\nu^{(\sigma)}(h) = \mathtt{d}\sum_{i=1}^{N}\sum_{x\in\Lambda}\frac{1}{N}\big[h(\tfrac{x}{L}) - h(\tfrac{\sigma_i}{L})\big] = \mathtt{d}L\cdot\nu^{(\sigma)}\left(\frac{1}{L}\sum_{x\in\Lambda}h(\tfrac{x}{L}) - h\right),$$

which implies the uniform convergence

$$\lim_{N/L\to\rho}\sup_{\sigma\in\Lambda^N}\Big|\mathfrak{G}_{L,N}(\nu^{(\cdot)}(h))(\sigma) - \nu^{(\sigma)}(A_{FV}h)\Big| = 0\,.$$

By considering cylindrical test-functions, this convergence can be extended to a core of the Fleming-Viot process with generator $\mathcal{L}_{FV}$ (5.30), in full analogy to our main results. We leave out further details.

Overall, this yields convergence of the (labelled) inclusion process in the following sense: If $\nu_\#\sigma^{(L,N)}(0) \xrightarrow{d} \nu_0$ then

$$\Big(\nu_\#\sigma^{(L,N)}(t)\Big)_{t\geqslant 0} \xrightarrow{d} (\nu_t)_{t\geqslant 0}\,, \quad \text{in } D([0,\infty),\mathcal{M}_1([0,1]))\,,$$

where $(\nu_t)_{t\geqslant 0}$ denotes the Fleming-Viot process generated by (5.30) with initial condition ν_0.

5.3 The diffusion limit in the case $\mathtt{d}L \to \infty$

This section treats the scaling limit of the inclusion process in the regime $\mathtt{d}L \to \infty$, where condensation on mesoscopic scales is expected. This case may also be considered as an interpretation of the Poisson-Dirichlet diffusion with infinite mutation rate $\theta = \infty$. Clearly, this corresponds to an infinite drift towards zero in the single particle operator A_θ, cf. (5.9). Thus, in order to see non-trivial dynamics, we have to rescale time appropriately. Recalling (5.43), we see that

$$\frac{1}{\mathtt{d}L}\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \mu^{(\eta)}\Big((h(0) - h(z) - zh'(z))\Big) + o(1)\,, \tag{5.76}$$

where $H(\mu) = \mu(h)$, $h \in C^3([0,1])$. The time-change also eradicates the interaction term in the corresponding measure-valued process. Indeed, we are left with a process that pushes mass (deterministically) from the interval $(0,1]$ onto zero. Hence, mass will not accumulate on the macroscopic scale. Instead, we need to consider an appropriate mesoscopic scale to see the actual dynamics of the fast mixing mechanism.

In [JCG19] it was proven that, at stationarity, mass condenses on the mesoscopic scale of order $\mathtt{d}^{-1}$, when $\rho \in (0,\infty)$, cf. (5.3). Thus the embedding of particle configurations into $\mathcal{M}_1(\mathbf{R}_+)$ via (5.11) with $\hat{\mu}^{(\eta)} = \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{\mathtt{d}L}{N}\eta_x}$ is an appropriate a-priori choice.[2] In order to take mass into account which lies on larger scales, we will consider probability measures $\mathcal{M}_1(\overline{\mathbf{R}}_+)$ on the one-point compactification $\overline{\mathbf{R}}_+ = [0,\infty]$. We equip $\mathcal{M}_1(\overline{\mathbf{R}}_+)$ with the topology induced by weak convergence, thus, $\mathcal{M}_1(\overline{\mathbf{R}}_+)$ is compact.

5.3.1 Deriving the diffusion limit

Once more, we rely on the Trotter-Kurtz approximation to conclude the scaling limit in Theorem 5.1.2. We follow the same steps as in Section 5.2, carried out below for completeness.

Proposition 5.3.1. *Let $H \in \mathcal{D}(\hat{\mathcal{L}})$, then*

$$\lim_{N/L\to\rho} \sup_{\eta\in\Omega_{L,N}} \left| \frac{1}{\mathtt{d}L} \mathfrak{L}_{L,N} H(\hat{\mu}^{(\cdot)})(\eta) - \hat{\mathcal{L}} H(\hat{\mu}^{(\eta)}) \right| = 0\,. \tag{5.77}$$

Below, we will show that the interaction term of the limiting measure-valued process indeed vanishes. First, we only consider test functions of the form $\mu \mapsto \mu(h)$.

Lemma 5.3.2. *Let $\rho \in (0,\infty)$ and $H(\mu) = \mu(h)$, with $h \in \mathcal{D}(\hat{A})$ (5.13). Then*

$$\lim_{N/L\to\rho} \sup_{\eta\in\Omega_{L,N}} \left| \frac{1}{\mathtt{d}L} \mathfrak{L}_{L,N} H(\hat{\mu}^{(\cdot)})(\eta) - \hat{\mu}^{(\eta)}(\hat{A}h) \right| = 0\,, \tag{5.78}$$

where $\hat{A}$ is the single-particle generator defined in (5.12).

Proof. We assume that $h \in C_c^3(\mathbf{R}_+)$, since the claim is trivially true for constant functions. Moreover, we will write $p_x := \mathtt{d}L\frac{\eta_x}{N}$.

[2] For the case $\rho \in (0,\infty)$, also the choice $\delta_{\mathtt{d}\eta_x}$ is appropriate and leads to a ρ dependent limit. However, for the boundary cases $\rho \in \{0,\infty\}$, the given choice turns out to be the correct one.

Following the same steps as in the proof of Lemma 5.2.3, we have

$$\frac{1}{\mathrm{d}L}\mathfrak{L}_{L,N}H(\hat{\mu}^{(\cdot)})(\eta) = \frac{1}{2N^2}\sum_{\substack{x,y=1\\x\neq y}}^{L}\eta_x\eta_y\big(\widetilde{h}''(p_x)+\widetilde{h}''(p_y)\big) + \frac{1}{LN}\sum_{\substack{x,y=1\\x\neq y}}^{L}\eta_x\big(\widetilde{h}'(p_y)-\widetilde{h}'(p_x)\big)+o(1)\,, \tag{5.79}$$

where we gained an additional factor $(\mathrm{d}L)^{-1}$ by rewriting $\frac{\eta_x}{N}h(\mathrm{d}\eta_x) = \frac{1}{\mathrm{d}L}\tilde{h}(p_x)$. Here we used again the fact that second-order terms in the second sum have a vanishing contribution, and first-order terms in the first sum cancel exactly, cf. proof of Lemma 5.2.3. Hence, (5.79) can be written as

$$\frac{1}{\mathrm{d}L}\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \sum_{x\in\Lambda}\frac{\eta_x}{N}\Big(1-\frac{p_x}{\mathrm{d}L}\Big)\tilde{h}''(p_x) + \left[\frac{1}{L}\sum_{y\in\Lambda}\tilde{h}'(p_y)-\sum_{x\in\Lambda}\frac{\eta_x}{N}\tilde{h}'(p_x)\right]+o(1)\,. \tag{5.80}$$

Using again Lemma 5.2.5, we have $\frac{1}{L}\sum_{y\in\Lambda}\tilde{h}'(p_y) = \tilde{h}'(0)+o(1)$. Hence,

$$\begin{aligned}\frac{1}{\mathrm{d}L}\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) &= \hat{\mu}^{(\eta)}\left(\tilde{h}''(p)+[\tilde{h}'(0)-\tilde{h}'(p)]\right)+o(1)\\ &= \hat{\mu}^{(\eta)}\big(\hat{A}h\big)+o(1)\,,\end{aligned}$$

where we additionally used the fact that $\tilde{h}''$ is bounded and of compact support, thus, $\frac{p}{\mathrm{d}L}\tilde{h}''(p)$ vanishes in the thermodynamic limit because $\mathrm{d}L\to\infty$. □

Remark 5.3.3. *Considering the state space $\mathcal{M}_{\leq 1}(\mathbf{R}_+)$ instead of $\mathcal{M}_1(\overline{\mathbf{R}}_+)$, (5.80) suggests that $\hat{A}$ should act via*

$$H\mapsto\big(\mu\mapsto\mu(\hat{A}h)+(1-\mu(1))h(0)\big)\,, \tag{5.81}$$

on functions $H:\mathcal{M}_{\leq 1}(\mathbf{R}_+)\mapsto\mathbf{R}$ of the form $H(\mu)=\mu(h)$. The extra term takes into account the transfer of mass from larger scales (above $\frac{N}{\mathrm{d}L}$), which is pushed onto microscopic scales, cf. Corollary 5.3.8. However, $\mu\mapsto\mu(1)$ is not a continuous function on $\mathcal{M}_{\leq 1}(\mathbf{R}_+)$. Instead, the mass transport from larger scales is implicit in the generator $\hat{\mathcal{L}}$, as we will see below.

Proof of Proposition 5.3.1. It suffices to consider functions $H\in\mathcal{D}(\hat{\mathcal{L}})$ of the

form $H(\mu) = \mu(h_1) \cdots \mu(h_n)$, with $h_k \in C_c^3(\mathbf{R})$. The interaction term of the operator $\hat{\mathcal{L}}$ is again given by the second-order term of the following expansion

$$
\begin{aligned}
&H(\hat{\mu}_{\#}\eta^{x,y}) \\
&= H(\hat{\mu}^{(\eta)}) \\
&\quad + \frac{1}{\mathtt{d}L} \sum_{k=1}^{n} [\widetilde{h}_k(p_y + \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_k(p_y) + \widetilde{h}_k(p_x - \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_k(p_x)] \prod_{\substack{l=1 \\ l \neq k}}^{n} \hat{\mu}^{(\eta)}(h_l) \\
&\quad + \frac{1}{(\mathtt{d}L)^2} \sum_{1 \le k < l \le n} [\widetilde{h}_k(p_y + \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_k(p_y) + \widetilde{h}_k(p_x - \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_k(p_x)] \\
&\qquad \times [\widetilde{h}_l(p_y + \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_l(p_y) + \widetilde{h}_l(p_x - \tfrac{\mathtt{d}L}{N}) - \widetilde{h}_l(p_x)] \prod_{\substack{j=1 \\ j \neq k,l}}^{n} \hat{\mu}^{(\eta)}(h_j) \\
&\quad + \frac{1}{(\mathtt{d}L)^3} r(\eta)\,.
\end{aligned}
\tag{5.82}
$$

In the first order term, each summand can be treated individually using Lemma 5.3.2. It only remains to check that both second-order term and remainder have no contribution.

A first-order Taylor expansion yields the following bound

$$
\begin{aligned}
&\left| \widetilde{h}(p_y + \tfrac{\mathtt{d}L}{N}) - \widetilde{h}(p_y) + \widetilde{h}(p_x - \tfrac{\mathtt{d}L}{N}) - \widetilde{h}(p_x) \right| \\
&\qquad = \tfrac{\mathtt{d}L}{N} \left| \tilde{h}'(p_y) - \tilde{h}'(p_x) + \frac{1}{2} \tfrac{\mathtt{d}L}{N} (\tilde{h}''(\xi_y) + \tilde{h}''(\xi_x)) \right| \\
&\qquad \le 2 \tfrac{\mathtt{d}L}{N} (\|\tilde{h}'\|_\infty + \|\tilde{h}''\|_\infty),
\end{aligned}
$$

where $\xi_x, \xi_y \in [0, \mathtt{d}N]$ correspond to the associated remainder term. Hence, the second-order term in (5.82), after applying $\frac{1}{\mathtt{d}L}\mathfrak{L}_{L,N}$, is upper bounded (up to a constant) by

$$
\frac{1}{dL} \sum_{1 \le k < l \le n} \frac{1}{N^2} \sum_{\substack{x,y=1 \\ x \neq y}}^{L} \eta_x(\mathtt{d} + \eta_y)(\|h_k'\|_\infty + \|h_k''\|_\infty)(\|h_l'\|_\infty + \|h_l''\|_\infty) \prod_{\substack{j=1 \\ j \neq k,l}}^{n} \|h_j\|_\infty\,,
$$

which vanishes as $\mathtt{d}L \to \infty$. For the same reason, higher order terms in the expansion (5.82) have no contribution either. □

The closure of $(\hat{A}, \mathcal{D}(\hat{A}))$ generates a Feller semigroup on $\overline{\mathbf{R}}_+$. Thus, the closure of $(\hat{\mathcal{L}}, \mathcal{D}(\hat{\mathcal{L}}))$ generates a Fleming-Viot process on the compact space

$\mathcal{M}_1(\overline{\mathbf{R}}_+)$, with trajectories in $C([0,\infty), \mathcal{M}_1(\overline{\mathbf{R}}_+))$ [EK87, Theorem 2.3], in the absence of interaction. We now have everything at hand, to prove Theorem 5.1.2.

Proof of Theorem 5.1.2. Once more, convergence of generators (thus semigroups, cf. Proposition 5.1.7) in Proposition 5.3.1 suffices to conclude convergence of the processes in $D([0,\infty), \mathcal{M}_1(\overline{\mathbf{R}}_+))$, by Proposition 5.1.8. Note that any fixed initial condition $\mu \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$ can be approximated by particle configurations using the embedding $\hat{\mu}^{(\cdot)}$, cf. Lemma 5.3.4 below. This completes the proof. □

Lemma 5.3.4. *Let $\rho \in [0,\infty]$ and $\mathtt{d} = \mathtt{d}(L)$ such that*

$$\frac{N}{L} \to \rho\,, \quad \mathtt{d}L \to \infty \quad \text{and} \quad \frac{\mathtt{d}L}{N} \to 0\,.$$

Then for any $\mu \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$, there exist $\eta^{(L,N)} \in \Omega_{L,N}$ such that

$$\hat{\mu}_{\#}\eta^{(L,N)} \xrightarrow{d} \mu\,. \tag{5.83}$$

Proof. We will see that it suffices to approximate discrete measures of the form

$$\alpha_0\delta_0 + \alpha_\infty\delta_\infty + \sum_{i=1}^{m} \alpha_i\delta_{p_i} \in \mathcal{M}_1(\overline{\mathbf{R}}_+)\,, \tag{5.84}$$

with $p_i \in (0,\infty)$, $1 \le i \le m$. Let ν be such a probability measure.

We explicitly construct configurations in $\Omega_{L,N}$ that converge to ν under the map (5.11)

$$\hat{\mu}^{(\eta)} = \hat{\mu}^{(\eta)}_{L,N} = \sum_{x=1}^{L} \frac{\eta_x}{N}\delta_{\mathtt{d}L\frac{\eta_x}{N}}\,,$$

when considering the thermodynamic limit $N/L \to \rho \in [0,\infty]$.

First, we consider the point masses lying in $(0,\infty)$. For convenience, let us introduce

$$k_i := \left\lfloor \frac{N}{\mathtt{d}L} p_i \right\rfloor \quad \text{and} \quad \#_i := \left\lfloor \frac{\alpha_i N}{k_i} \right\rfloor\,.$$

Note that $k_i \to \infty$, as $N/L \to \rho$, by assumption.

Now, let η' be the vector given by gluing together vectors $(k_i, \dots, k_i) \in$

$\mathbf{N}^{\#_i}$, $1 \leq i \leq m$. Recalling that $p_i > 0$, we then have

$$\frac{1}{L}\sum_{i=1}^{m} \#_i \leqslant \sum_{i=1}^{m} \frac{\alpha_i N}{\frac{1}{2}k_i L} \leqslant 2\mathtt{d}\sum_{i=1}^{m} \frac{\alpha_i}{p_i}\,,$$

and the right-hand side vanishes, because the sum is finite. We will assume that the resulting vector η' lies in $\mathbf{N}^L$, since we can append zeros until η' has length L. In particular, the relative number of empty sites converges to one.

It only remains to distribute the remaining particles, to create the point masses at zero and infinity. Thus far, only $\#_\Sigma = \sum_{i=1}^{m} \#_i$ sites are occupied in η'. We start by adding $k_\infty := \lfloor \alpha_\infty N \rfloor$ particles onto the $(\#_\Sigma + 1)$-th position of η' (which is empty). This corresponds to the point mass at infinity.

The number of particles allocated to η' is upper bounded by

$$k_\infty + \sum_{i=1}^{m} k_i \#_i \leqslant \alpha_\infty N + \sum_{i=1}^{m} \alpha_i N = (1-\alpha_0)N\,.$$

We distribute the remaining $k_0 := N - k_\infty - \sum_{i=1}^{m} k_i \#_i$ particles as uniform as possible, among all empty sites of η' (there are $\#_0 := L - \#_\Sigma - 1$ many). This yields the configuration

$$\eta_x := \begin{cases} \eta'_x & \text{if } 1 \leq x \leq \#_\Sigma + 1, \\ \left\lfloor \frac{k_0}{\#_0} \right\rfloor + \mathbb{1}_{x \in \{\#_\Sigma + 2, \ldots, \#_\Sigma + 1 + (k_0 \mod \#_0)\}} & \text{otherwise.} \end{cases}$$

Because the number of non-empty sites in η' is relatively vanishing, all particles distributed on previously empty sites will correspond to the point mass at zero. Note that $\eta \in \Omega_{L,N}$ by construction.

Indeed, the constructed particle configuration η approximates the discrete measure arbitrarily well in the thermodynamic limit, as for every $f \in C_b(\overline{\mathbf{R}}_+)$ we have

$$\hat{\mu}^{(\eta)}(f) = \sum_{x=\#_\Sigma+2}^{L} \frac{\eta_x}{N} f\left(\frac{\mathtt{d}L}{N}\eta_x\right) + \frac{k_\infty}{N} f\left(\frac{\mathtt{d}L}{N}k_\infty\right) + \sum_{i=1}^{m} \frac{k_i}{N}\#_i\, f\left(\frac{\mathtt{d}L}{N}k_i\right)\,,$$

which converges to

$$\lim_{N/L \to \rho} \hat{\mu}^{(\eta)}(f) = \alpha_0 f(0) + \alpha_\infty f(\infty) + \sum_{i=1}^{m} \alpha_i f(p_i) = \nu(f)\,.$$

Now, for every $\mu \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$ there exists a sequence $(\nu_n)_{n\in\mathbf{N}}$ of measures of

the form (5.84) such that $\nu_n \xrightarrow{d} \mu$. Moreover, each ν_n can be approximated by a sequence $(\eta_n^{(L,N)})_{L,N}$ following the above approach. Hence, we can construct a sequence of configurations with the desired property, using a diagonal argument. □

5.3.2 Duality and the hydrodynamic limit

The evolution of a single observable $H(\hat{\mu}_t) := \hat{\mu}_t(h)$, $h \in \mathcal{D}(\hat{A})$, can be expressed by Dynkin's formula as

$$\mathrm{d}\hat{\mu}_t(h) = \hat{\mu}_t(\hat{A}h)\ \mathrm{d}t + \mathrm{d}\hat{M}_t^{(h)}\,, \tag{5.85}$$

with the quadratic variation of the martingale $\hat{M}_t^{(h)}$ given in terms of the carré du champ operator:

$$\left[\hat{M}^{(h)}\right]_t = \int_0^t \left(\hat{\mathcal{L}}H^2(\hat{\mu}_s) - 2H\hat{\mathcal{L}}H(\hat{\mu}_s)\right)\,\mathrm{d}s\,.$$

Note that $\left[\hat{M}^{(h)}\right]$ vanishes because $\hat{\mathcal{L}}H^2 = 2H\hat{\mathcal{L}}H$, cf. (5.15). Hence, $\hat{M}^{(h)} = 0$ and $(\hat{\mu}_t)_{t\geqslant 0}$ solves the ODE

$$\mathrm{d}\hat{\mu}_t(h) = \hat{\mu}_t(\hat{A}h)\ \mathrm{d}t\,. \tag{5.86}$$

Therefore, the absence of interaction in $\hat{\mathcal{L}}$ leads to a deterministic evolution of $(\hat{\mu}_t(h))_{t\geq 0}$, which can be described by a single particle, evolving according to $\hat{A}$. Unlike in the case of $\mathrm{d}L \to \theta < \infty$, we can fully characterise the semigroup of $\hat{\mathcal{L}}$ by considering only the evolution with respect to the single particle generator $\hat{A}$.

Proposition 5.3.5. *Let $g \in C(\overline{\mathbf{R}}_+^n)$ and define $G(\mu) := \mu^{\otimes n}(g)$, then for any $\hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$*

$$G\left(\hat{\mu}_t\right) = \mathbf{E}_{\hat{\mu}_0^{\otimes n}}\left[g(\hat{Z}_t)\right]\,, \quad \forall t \geq 0\,, \tag{5.87}$$

where $(\hat{\mu}_t)_{t\geqslant 0}$ evolves with respect to $\hat{\mathcal{L}}$ and $(\hat{Z}_t)_{t\geq 0}$ is the process consisting of n independent copies generated by the single-particle generator $\hat{A}$ (5.12). Notably, for $n = 1$ we have $\hat{\mu}_t = Law(\hat{Z}_t)$, whenever the initial conditions agree in the sense that $\hat{Z}_0 \sim \hat{\mu}_0$.

Proof. We follow the same steps as in [EK93, Section 6], which we reproduce

here for completeness. First, we define the operator

$$\hat{C}^{(n)} g(z_1, \dots, z_n) := \sum_{k=1}^{n} \hat{A} g(z_1, \dots, z_{k-1}, \cdot, z_{k+1}, \dots, z_n)(z_k)\,, \tag{5.88}$$

acting on functions $g \in C^2(\overline{\mathbf{R}}_+^n)$, which satisfies

$$\hat{\mathcal{L}} H(\mu) = \mu^{\otimes n}\big(\hat{C}^{(n)} g_H\big)\,,$$

for any $H \in \mathcal{D}(\hat{\mathcal{L}})$ of the form $H(\mu) = \mu(h_1) \dots \mu(h_n)$, with $h_i \in \mathcal{D}(\hat{A})$, and $g_H := h_1 \otimes \cdots \otimes h_n \in C^2(\overline{\mathbf{R}}_+^n)$. This implies

$$(\lambda - \hat{\mathcal{L}}) H(\mu) = \mu^{\otimes n}\big((\lambda - \hat{C}^{(n)}) g_H\big)\,, \qquad \forall \lambda > 0\,,\ \mu \in \mathcal{M}_1(\overline{\mathbf{R}}_+)\,. \tag{5.89}$$

Thus, for any $g \in C(\overline{\mathbf{R}}_+^n)$ and $\lambda > 0$, we have

$$\begin{aligned} \mu^{\otimes n}((\lambda - \hat{C}^{(n)})^{-1} g) &= (\lambda - \hat{\mathcal{L}})^{-1}\big(\bullet^{\otimes n}\,((\lambda - \hat{C}^{(n)})(\lambda - \hat{C}^{(n)})^{-1} g)\big)(\mu) \\ &= (\lambda - \hat{\mathcal{L}})^{-1} G(\mu)\,, \end{aligned}$$

where $G(\mu) := \mu^{\otimes}(g)$. Now, we use the fact that

$$\mathbb{E}_{\hat{\mu}_0}[G(\hat{\mu}_t)] = \lim_{m\to\infty} \Big(1 - \frac{t}{m}\hat{\mathcal{L}}\Big)^{-m} G(\hat{\mu}_0) = \lim_{m\to\infty} \hat{\mu}_0^{\otimes n}\Big(\Big(1 - \frac{t}{m}\hat{C}^{(n)}\Big)^{-m} g\Big)\,.$$

Because the expression inside the expectation on the right-hand side converges uniformly in $C(\overline{\mathbf{R}}_+^n)$, we can take the limit inside the expectation. Overall, this yields

$$\mathbb{E}_{\hat{\mu}_0}[G(\hat{\mu}_t)] = \hat{\mu}_0^{\otimes n}\Big(\lim_{m\to\infty} \Big(1 - \frac{t}{m}\hat{C}^{(n)}\Big)^{-m} g\Big) = \mathbf{E}_{\hat{\mu}_0^{\otimes n}}\left[g(\hat{Z}_t)\right]\,, \tag{5.90}$$

where $\hat{Z}$ is the process generated by $\hat{C}^{(n)}$.

Now, let us consider the case $n = 1$, for which (5.90) reads

$$\hat{\mu}_t(h) = \mathbb{E}_{\hat{\mu}_0}\left[\hat{\mu}_t(h)\right] = \mathbf{E}_{\hat{\mu}_0}\left[h(\hat{Z}_t)\right]\,, \quad \forall t \geq 0\,,\ h \in C(\overline{\mathbf{R}}_+)\,, \tag{5.91}$$

where we additionally used the fact that $\hat{\mu}_t(h)$ is deterministic, cf. (5.86). In particular, (5.91) implies that $\hat{\mu}_t = Law(\hat{Z}_t)$, and the measure-valued evolution $(\hat{\mu}_t)_{t\geqslant 0}$ is indeed deterministic. Thus, the expected value on the left-hand side of (5.90) has no affect and can be dropped. □

The duality result in Proposition 5.3.5, and equivalently Theorem 5.1.2,

can be interpreted in the sense of a hydrodynamic limit.

Proposition 5.3.6 (Hydrodynamic limit)**.** *Consider the process $(\hat{\mu}_t)_{t\geq 0}$ generated by $\hat{\mathcal{L}}$ with initial data $\hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbf{R}}_+)$. Then for every $t > 0$, $\hat{\mu}_t$ has a Lebesgue-density $f(t,\cdot)$ on $\mathbf{R}_+$. The evolution of the density $(f(t,\cdot))_{t>0}$ solves*

$$\begin{cases} \partial_t f(t,z) = z\,\partial_z^2 f(t,z) + z\,\partial_z f(t,z) \\ \lim_{z\to 0} f(t,z) = 1 \end{cases} \tag{5.92}$$

with $\lim_{t\to 0}\int_0^\infty h(z)f(t,z)\,\mathrm{d}z = \hat{\mu}_0(h)$ for every $h \in C_c(\mathbf{R}_+)$.

Remark 5.3.7. *The diffusion part of the generator $\hat{A}$, given by*

$$\hat{A}_D h(z) := zh''(z) + (2-z)h'(z)\,. \tag{5.93}$$

$\hat{A}_D$ generates the so called Cox-Ingersoll-Ross model, see e.g. [Mar11], which is a well studied diffusion process in mathematical finance and population genetics.

Proof. First consider the case of an initial condition that has no atom at infinity, i.e. $\|f_0\|_{L^1(\mathbf{R}_+)} = 1$, and let $z_0 \in [0,\infty)$. The Cox-Ingersoll-Ross model generated by $\hat{A}_D$ is known to have a density $g(t,\cdot|z_0)$, for any positive time and initial data [Mar11, Display below (3.2)]. In fact, it is explicitly given and for $z_0 = 0$ it evaluates to

$$g(t,z|0) = \frac{z}{(2\ell_t)^2} e^{-z(2\ell_t)^{-1}} \quad \text{with} \quad \ell_t := \tfrac{1}{2}(1-e^{-t})\,. \tag{5.94}$$

Furthermore, for any $t > 0$ we have $g_t(\cdot|z_0)\big|_{(0,\infty)} \in C^\infty((0,\infty))$ [Mar11, Proposition 3.2]. The resetting mechanism is given by a Poisson jump process, thus, [APZ13, Theorem 1] guarantees that also the process $(\hat{Z}_t)_{t\geq 0}$, generated by $\hat{A}$, has a density which is given by

$$f(t,z|z_0) = e^{-t}g(t,z|z_0) + \int_0^t e^{-s}g(s,z|0)\,\mathrm{d}s\,. \tag{5.95}$$

We note that $f(t,\cdot|z_0)$ inherits the regularity properties of $g(t,\cdot|z_0)$ on $(0,\infty)$. This follows from the change of variable $r = z/(2\ell_s)$ with $\frac{dr}{ds} = -\frac{z}{(2\ell_s^2)}e^{-s}$ which yields for every $z > 0$

$$f(t,z|z_0) = e^{-t}g(t,z|z_0) + \int_{\frac{z}{2\ell_t}}^\infty e^{-r}\,\mathrm{d}r = e^{-t}g(t,z|z_0) + e^{-z(2\ell_t)^{-1}}\,. \tag{5.96}$$

Hence, $f(t,\cdot|z_0) \in C^\infty((0,\infty))$.

It is only left to verify that $(f(t,\cdot))_{t\geq 0}$ indeed solves the given PDE. Using integration by parts, we see that for any $h \in C_c^2(\mathbf{R}_+)$, we have

$$\mu_t(\hat{A}h) = \int_0^\infty f(t,z)\,\hat{A}h(z)\,\mathrm{d}z = \int_0^\infty \hat{A}^* f(t,z)\,h(z)\,\mathrm{d}z\,, \tag{5.97}$$

with the adjoint action defined as

$$\hat{A}^* f(z) := zf''(z) + zf'(z) + \delta_0(z)\big(1 - f(z)\big)\,. \tag{5.98}$$

Hence, the density $f = \big(f(t,\cdot)\big|_{(0,\infty)}\big)_{t\geq 0}$ in (5.96) solves the PDE

$$\partial_t f = \hat{A}^* f\,, \quad f(0,\cdot) = \delta_{z_0}\,. \tag{5.99}$$

From (5.96), it is easy to see that $\lim_{z\to 0} f(t,z) = 1$ for any $t > 0$, since $g(t,0|z_0) = 0$ for all $z_0 \geq 0$. Thus, the boundary term in (5.98) vanishes and we are left with the PDE in the statement.

Lastly, consider $f_0 = \delta_\infty$. The only way $(\hat{Z}_t)_{t\geq 0}$ can escape infinity is via the resetting mechanism. Therefore, (5.96) still applies with $g(t,z|z_0)$ replaced by δ_∞, since $\mathbb{P}(\hat{Z}_t = \infty)$ is equal to the probability that the process has not jumped yet. One can check that also $f(\cdot|\infty)$ solves the given PDE on $(0,\infty)$ with the correct boundary condition. The result for arbitrary initial conditions now follows by integrating over the densities with respect to μ_0 and Leibniz rule. □

Given the explicit form of the density (5.96), we can read of the evolution of the mass process $(\mu_t(\mathbf{R}_+))_{t\geq 0}$.

Corollary 5.3.8. *For all $t \geqslant 0$, we have*

$$\hat{\mu}_t(\mathbf{R}_+) = 1 - (1 - \hat{\mu}_0(\mathbf{R}_+))e^{-t}\,. \tag{5.100}$$

Proof. We integrate the PDE from Proposition 5.3.6 in space, which yields

$$\partial_t \|f(t,\cdot)\|_{L^1(\mathbf{R}_+)} = \int_0^\infty \big(z\,\partial_z^2 f(t,z) + z\,\partial_z f(t,z)\big)\,\mathrm{d}z\,. \tag{5.101}$$

Using integration by parts, the right-hand side simplifies to

$$\mathrm{d}\alpha_t = (1 - \alpha_t)\,\mathrm{d}t\,, \quad \alpha_0 = \mu_0(\mathbf{R}_+)\,. \tag{5.102}$$

Its solution is given by (5.100). □

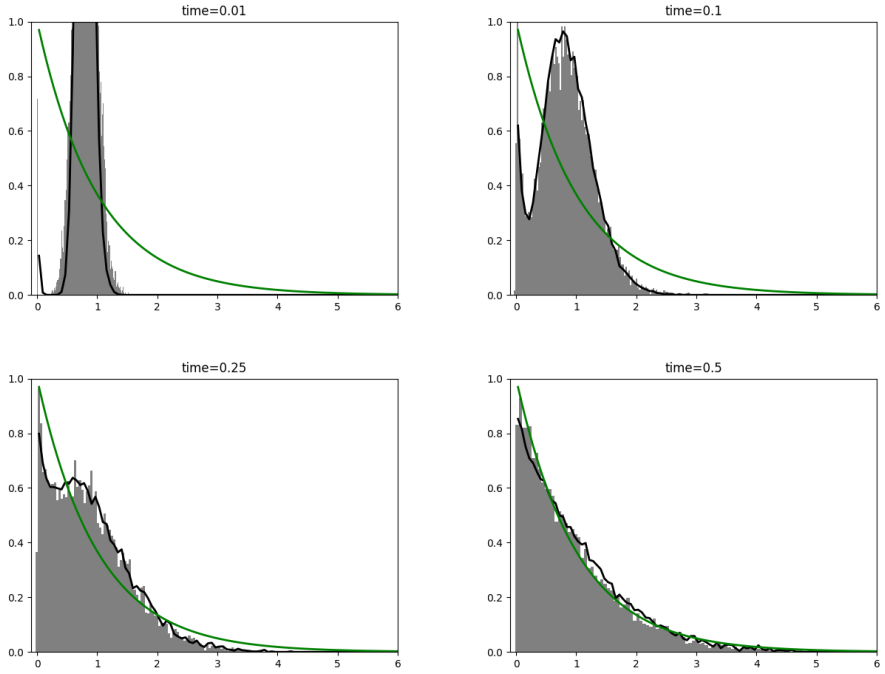


Figure 5.2: Simulations for both the inclusion process ($N = L = 1024$, $\mathrm{d} = L^{-1/2} = \frac{1}{32}$) and the jump diffusion generated by $\hat{A}$ agree, in accordance with Proposition 5.3.5. The black graph shows the density profile of the embedded inclusion process (5.11) (1000 samples), whereas the grey histogram represents the density of the jump diffusion (10000 samples). Both profiles converge rapidly to the unit exponential density (green line), cf. Lemma 5.3.9. We considered an initial condition $\hat{\mu}_0 = \delta_z$, $z \simeq \frac{25}{32}$.

Let us summarise the invariance and exponential ergodicity property of $(\hat{\mu}_t)_{t\geq 0}$.

Lemma 5.3.9. *The process $(\hat{Z}_t)_{t\geq 0}$ satisfies the following properties:*

(i) The exponential distribution $\mathrm{Exp}(1)$ *is the unique invariant probability measure.*

(ii) We have

$$\left\| Law(\hat{Z}_t) - \mathrm{Exp}(1) \right\|_{TV} \leq e^{-t}, \quad \forall t \geq 0\,. \tag{5.103}$$

Proof. The fact that the exponential distribution is invariant can be explicitly proven using integration by parts, but also follows directly from [APZ13, Corollary 1]. The exponential ergodicity is a consequence of [APZ13, Theo-

rem 2]. The same result yields that the process is Harris recurrent, thus, the Exponential distribution is the unique invariant distribution. □

It is an immediate consequence, that the point mass on the Exponential distribution is an absorbing point for the measure valued process.

Corollary 5.3.10. *The atom $\hat{\mathbf{P}} = \delta_{\mathrm{Exp}(1)} \in \mathcal{M}_1(\mathcal{M}_1(\mathbf{R}_+))$ is reversible with respect to the dynamics induced by $\hat{\mathcal{L}}$.*

Proof. Let $H \in \mathcal{D}(\hat{\mathcal{L}})$ be of the form $H(\mu) = \mu(h_1) \cdots \mu(h_n)$. For simplicity we write $\nu := \mathrm{Exp}(1) \in \mathcal{M}_1(\mathbf{R}_+)$. Then

$$\hat{\mathcal{L}}H(\nu) = \sum_{j=1}^{n} \nu(\hat{A}h_i) \prod_{l \neq j} \nu(h_l) = 0\,, \tag{5.104}$$

as ν is invariant with respect to $\hat{A}$, cf. Lemma 5.3.9. Thus, $\hat{\mathbf{P}}(F\hat{\mathcal{L}}H) = 0 = \hat{\mathbf{P}}(H\hat{\mathcal{L}}F)$. □

5.3.3 A natural extension of the PD-diffusion

In this section we prove Theorem 5.1.5, which states that $\mathcal{L}_\theta$ (which is equivalent to the Poisson-Dirichlet diffusion $\mathcal{G}_\theta$) has a limit as $\theta \to \infty$ under appropriate rescaling.

Proof of Theorem 5.1.5. Once more, the statement is a conclusion of the Trotter-Kurtz approximation theorem. We state the essential steps for completeness. Let $H \in \mathcal{D}(\hat{\mathcal{L}})$ be of the form $\mu \mapsto \mu(h_1) \cdots \mu(h_n)$, $h_k \in \mathcal{D}(\hat{A})$, and define $H_\theta \in \mathcal{D}(\mathcal{L}_\theta)$ by

$$E \ni \mu \mapsto \mu(h_1(\theta\,\cdot)) \cdots \mu(h_n(\theta\,\cdot))\,, \tag{5.105}$$

with $h_k(\theta\,\cdot)$'s elements of $C^3([0,1])$. We have

$$\hat{A}h_k(\theta z) - \frac{1}{\theta} A_\theta h_k(\theta\,\cdot)(z) = z\big(\theta z\, h_k''(\theta z) + 2\, h_k'(\theta z)\big)\,. \tag{5.106}$$

If h_k is a constant function, the right-hand side vanishes. On the other hand, if $h_k \in C_c^3(\mathbf{R}_+)$, we can write

$$\begin{aligned}\sup_{\mu \in E} \big|\mu\big((\hat{A}h_k)(\theta\cdot)\big) - \tfrac{1}{\theta}\mu\big(A_\theta h_k(\theta\,\cdot)\big)\big| \leq & C_{h_k} \sup_{\mu \in E} \mu\big(Z\, \mathbb{1}_{\theta Z \in \mathrm{supp}(h_k)}\big) \\ = & C_{h_k} \sup_{z \in [0,1]} \big\{z\, \mathbb{1}_{\theta z \in \mathrm{supp}(h_k)}\big\}\,,\end{aligned}$$

where $Z \sim \mu$ and C_{h_k} is a finite constant, depending on h_k. The right-hand side vanishes as $\theta \to \infty$, due to h_k having compact support. Along the same lines, we can show that the interaction term of $\frac{1}{\theta}\mathcal{L}_\theta$ disappears. Overall we conclude

$$\lim_{\theta\to\infty} \sup_{\mu\in E} \left| \tfrac{1}{\theta}\mathcal{L}_\theta H_\theta(\mu) - (\hat{\mathcal{L}}H)(S_\theta\mu) \right| = 0\,. \tag{5.107}$$

Again, the convergence of generators suffices to conclude weak convergence on the process level. □

5.4 Discussion and outlook

We conclude this chapter with a discussion of boundary cases in the setting considered and outline future directions, as well as work in progress.

Throughout the chapter, we assumed that $\rho = \lim_{N,L\to\infty} N/L \in (0,\infty)$. However, the derived scaling limits do not depend on the actual value of ρ, as we study the distribution of mass after renormalising by N. As long as $N, L \to \infty$, our results extend to the boundary cases $\rho \in \{0,\infty\}$, up to a regime around $\rho = 0$ in the case $\mathtt{d}L \to \infty$. In this regime we see an interesting transition of the clustering behaviour, see Lemma 5.4.1 below.

First consider $\theta < \infty$ and $\mathtt{d} \to 0$, for which both cases $\rho \in \{0,\infty\}$ are covered by our proof. For $\rho = 0$, i.e. $N \ll L$, this is clear intuitively, as an increasing number of empty sites does not affect the dynamics since the total diffusivity per particle is $\mathtt{d}L \to \theta$. On the other hand, if $\rho = \infty$ it may seem surprising that the number of sites L does not play a role (as long they are divergent). Here, the core lies in Lemma 5.2.7, which states that for any thermodynamic limit, we can approximate configurations in $\overline{\nabla}$ (equivalently measures in E) by a sequence of particle configurations. Indeed, having a closer look at the proof of Lemma 5.2.7, we see it is only necessary that a macroscopic excess mass of order $O(N)$ can be distributed uniformly over sites while not being visible under the macroscopic rescaling $\frac{1}{N}$. This is the case, since we can always allocate $O(\frac{N}{L})$-many particles on sites (which might itself diverge), however, under macroscopic rescaling we have $\frac{1}{N}\frac{N}{L} \to 0$.

Our approach can further be adapted, with minimal changes in Lemma 5.2.5, to cover the situation of fixed L and $\mathtt{d} = \mathtt{d}(N) \to 0$, as $N \to \infty$. This has been considered in [BDG17, KS21], to study the metastable dynamics of a single condensate on a large time scale. In our time scale, the system

is described by a Wright-Fisher diffusion (5.69) with a single cluster site as absorbing state, describing convergence to a typical stationary configuration.

Now assume $\theta = \infty$. In the case $\rho \in (0, \infty)$, the results in the present chapter, and also [JCG19], yields that a size-biased chosen chunk (at equilibrium) is approximately exponentially distributed with mean $\sim \frac{N}{\mathtt{d}L}$. In fact, looking at the proof of Theorem 5.1.2, the result remains true as long as $\mathtt{d}L/N \to 0$. This trivially holds when $\rho = \infty$, in which case cluster sizes live on the scale of order

$$\frac{N}{\mathtt{d}L} \gg \frac{1}{\mathtt{d}}. \tag{5.108}$$

On the other hand, if $N/L \to 0$ we have no control over $\mathtt{d}N$ (in contrast to $N/L \to \infty$, which implies $\mathtt{d}N \to \infty$ since $\mathtt{d}L \to \infty$). Therefore, for $\rho = 0$ the convergence in Theorem 5.1.2 remains true, only if $\mathtt{d}L/N \to 0$, i.e. $\mathtt{d} \ll N/L$. Assume on the other hand that $\frac{N}{\mathtt{d}L} \to \gamma \in (0, \infty]$, then we don't expect any clustering of particles on diverging scales. In fact, we see a finer structure emerging in the limit on scales of order one.

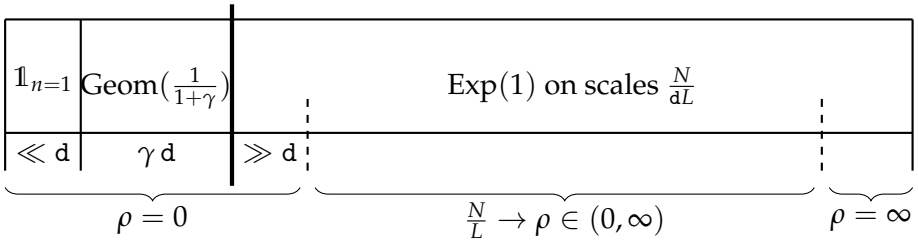


Figure 5.3: Graphical summary of the clustering of particles at equilibrium for the inclusion process when $\mathtt{d}L \to \infty$. The distributions displayed describe the first size-biased marginal $\tilde{\eta}_1$ on the appropriate scale. Note particularly the transition from diverging scales to scales of order 1, when moving from the regime $\rho \gg \mathtt{d}$ into $\rho \sim \gamma\mathtt{d}$.

Lemma 5.4.1. *Assume that* $\frac{N}{L} \to 0$, *as* $N, L \to \infty$, $\mathtt{d} \to 0$ *and* $\mathtt{d}L \to \infty$ *such that* $\frac{N}{\mathtt{d}L} \to \gamma \in [0, \infty)$. *Then*

$$\lim_{N/L \to 0} \pi_{L,N}(\tilde{\eta}_1 \in \cdot) = \mathrm{Geom}\left(\frac{1}{1+\gamma}\right). \tag{5.109}$$

Here $\pi_{L,N}$ *denotes the unique invariant distribution with respect to* $\mathfrak{L}_{L,N}$.

Hence, for $\rho = 0$ and $\mathtt{d}L \to \infty$, there is a critical scaling $N \sim \mathtt{d}L$, below which the equilibrium measure does not exhibit clustering of particles

on diverging scales. Note that the lemma is independent of the underlying graph structure and holds more generally for irreducible and spatially homogeneous dynamics, as in [JCG19].

Before proving the above lemma, we require some notation and representations, which can be found in more detail in [JCG19]. Recall that $\pi_{L,N}$ denotes the unique invariant distribution with respect to $\mathfrak{L}_{L,N}$ supported on $\Omega_{L,N}$, which is given explicitly by

$$\pi_{L,N}(\,\mathrm{d}\eta) = \frac{1}{Z_{L,N}} \prod_{x=1}^{L} w_L(\eta_x)\;\mathrm{d}\eta\,, \tag{5.110}$$

where $\mathrm{d}\eta$ denotes the counting measure, $Z_{L,N}$ the appropriate partition function and w_L describing weights of the form

$$w_L(n) = \frac{\Gamma(n+\mathtt{d})}{n!\Gamma(\mathtt{d})}\,, \tag{5.111}$$

arising from the choice of transition rates in $\mathfrak{L}_{L,N}$. The partition function can be explicitly written in terms of

$$Z_{L,N} = \frac{\Gamma(N+\mathtt{d}L)}{N!\Gamma(\mathtt{d}L)}\,. \tag{5.112}$$

Proof of Lemma 5.4.1. Recall the definition of size-biased sampling from the beginning of the chapter (above Figure 5.1), then

$$\pi_{L,N}(\tilde{\eta}_1 = n) = \sum_{x=1}^{L} \frac{n}{N}\pi_{L,N}(\eta_x = n) = \frac{L}{N} n\, w_L(n) \frac{Z_{L-1,N-n}}{Z_{L,N}}\,, \tag{5.113}$$

which is equal to zero if $n = 0$. Thus, in the following let $n > 0$. We replace the terms in the previous display with the corresponding expressions in (5.111) and (5.112), which yields

$$\begin{aligned}\pi_{L,N}(\tilde{\eta}_1 = n) &= \frac{L}{N}\frac{\Gamma(n+\mathtt{d})}{(n-1)!\Gamma(\mathtt{d})}\frac{\Gamma(N-n+\mathtt{d}(L-1))}{\Gamma(N+\mathtt{d}L)}\frac{N!}{(N-n)!}\frac{\Gamma(\mathtt{d}L)}{\Gamma(\mathtt{d}(L-1))} \\ &\sim \frac{\mathtt{d}L}{N}\frac{\Gamma(n+\mathtt{d})}{\mathtt{d}\,(n-1)!\Gamma(\mathtt{d})}\frac{\Gamma(N-n+\mathtt{d}L)}{\Gamma(N+\mathtt{d}L)}N^n\,.\end{aligned} \tag{5.114}$$

We analyse the remaining terms individually. First,

$$\frac{\Gamma(n+\mathtt{d})}{\mathtt{d}\,(n-1)!\Gamma(\mathtt{d})} = \frac{1}{\mathtt{d}\,(n-1)!}\prod_{k=0}^{n-1}(k+\mathtt{d}) = \prod_{k=1}^{n-1}\frac{k+\mathtt{d}}{k} \to 1\,,$$

and second,

$$\frac{\Gamma(N-n+\mathrm{d}L)}{\Gamma(N+\mathrm{d}L)}N^n = N^n \prod_{k=0}^{n-1} \frac{\Gamma(N-(k+1)+\mathrm{d}L)}{\Gamma(N-k+\mathrm{d}L)} = \prod_{k=1}^{n} \frac{1}{1-\frac{k}{N}+\frac{\mathrm{d}L}{N}}\,.$$

Overall, we conclude that

$$\pi_{L,N}(\tilde{\eta}_1 = n) \sim \frac{\mathrm{d}L}{N}\left(\frac{\frac{N}{\mathrm{d}L}}{1+\frac{N}{\mathrm{d}L}}\right)^n \to \frac{\gamma^{n-1}}{(1+\gamma)^n}\,, \tag{5.115}$$

since $\frac{N}{\mathrm{d}L} \to \gamma$. This finishes the proof. □

A natural next step, in view of the hydrodynamic limit Proposition 5.3.6, when $\mathrm{d}L \to \infty$, is to study fluctuations around the equilibrium. A second moment calculation with respect to the stationary measure yields

$$\pi_{L,N}\left(\left(\hat{\mu}^{(\eta)}(h) - \mathrm{Exp}(1)(h)\right)^2\right) \sim \frac{1}{\mathrm{d}L}\pi_{L,N}\left(\tilde{\eta}_1 \tfrac{\mathrm{d}L}{N} h(\tilde{\eta}_1 \tfrac{\mathrm{d}L}{N})^2\right) \to 0\,, \tag{5.116}$$

for any $h \in C_b(\mathbf{R}_+)$. Hence, in order to see a non-trivial limit, we should investigate fluctuations of order $\sqrt{\mathrm{d}L}$ by studying the limiting behaviour of

$$\sqrt{\mathrm{d}L}\left((\hat{\mu}_{\#}\eta^{(L,N)}(\tfrac{t}{\mathrm{d}L}))(h) - \mathrm{Exp}(1)(h)\right)\,, \quad t \geq 0\,, \tag{5.117}$$

with time slowed down as indicated in Theorem 5.1.2. Due to decoupling of the size-biased marginals, the fluctuations are expected to be Gaussian.

The approach in this chapter (based on the work [CGG23]) should be robust towards perturbation of transition rates, as we do not require the explicit form of the partition function. Throughout the chapter, we have focused on the one parameter family of Poisson-Dirichlet diffusions. There exists a two-parameter extension of the process, which was introduced in [Pet09]. This process has gained a lot of attention over the past years [RW09, FSWX11, Eth14, CBE+17], just to name a few. It would be interesting to investigate the size-biased approach in this setting. The two parameter process has only been studied when fixing finitely many locations/sites and observing the evolution of mass on them, see for example [FPRW23] and also the discussion in Section 5.2.3.

Furthermore, it would be interesting to investigate diffusion limits of the generalised version of the inclusion process with non-trivial bulk, studied in [CGG22]. Numerical simulations and heuristic arguments suggest that the

macroscopic phase evolves under the dynamics described in Theorem 5.1.1. At the same time, one can observe a transfer of mass between the bulk and the condensate, whose evolution is described by a system of ODEs, similar to Corollary 5.3.8.

Bibliography

[ACR21] M. Ayala, G. Carinci, and F. Redig. Condensation of SIP Particles and Sticky Brownian Motion. *Journal of Statistical Physics*, 183(3), 2021.

[AKQ14a] T. Alberts, K. Khanin, and J. Quastel. The continuum directed random polymer. *Journal of Statistical Physics*, 154(1):305–326, 2014.

[AKQ14b] T. Alberts, K. Khanin, and J. Quastel. The intermediate disorder regime for directed polymers in dimension 1+1. *The Annals of Probability*, 42(3):1212–1256, 2014.

[APZ13] K. Avrachenkov, A. Piunovskiy, and Y. Zhang. Markov processes with restart. *Journal of Applied Probability*, 50(4):960 – 968, 2013.

[AZ96] S. Albeverio and X. Y. Zhou. A martingale approach to directed polymers in a random environment. *Journal of Theoretical Probability*, 9(1):171–189, 1996.

[BC95] L. Bertini and N. Cancrini. The stochastic heat equation: Feynman-Kac formula and intermittence. *Journal of Statistical Physics*, 78(5-6):1377–1401, 1995.

[BC98] L. Bertini and N. Cancrini. The two-dimensional stochastic heat equation: renormalizing a multiplicative noise. *Journal of Physics A: Mathematical and General*, 31(2):615–622, 1998.

[BCCH20] Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer. Renormalising SPDEs in regularity structures. *Journal of the European Mathematical Society*, 23(3):869–947, 2020.

[BCJL94] L. Bertini, N. Cancrini, and G. Jona-Lasinio. The stochastic Burgers equation. *Communications in Mathematical Physics*, 165(2):211 – 232, 1994.

[BD00] E. Brunet and B. Derrida. Probability distribution of the free energy of a directed polymer in a random medium. *Physical Review E*, 61(6):6789–6801, 2000.

[BDG17] A. Bianchi, S. Dommers, and C. Giardinà. Metastability in the reversible inclusion process. *Electronic Journal of Probability*, 22, 2017.

[BG97] L. Bertini and G. Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Communications in Mathematical Physics*, 183(3):571–607, 1997.

[BHM+22] A. Bovier, F. den Hollander, S. Marello, E. Pulvirenti, and M. Slowik. Metastability of Glauber dynamics with inhomogeneous coupling disorder. *arXiv preprint arXiv:2209.09827*, 2022.

[BHZ19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Invent. Math.*, 215(3):1039–1156, 2019.

[Bil99] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., 1999.

[BK91] L. Bronsard and R. V. Kohn. Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics. *Journal of Differential Equations*, 90(2):211–237, 1991.

[BK10] D. Blount and M. A. Kouritzin. On convergence determining and separating classes of functions. *Stochastic Processes and their Applications*, 120(10):1898–1907, 2010.

[Bol89] E. Bolthausen. A note on the diffusion of directed polymers in a random environment. *Communications in Mathematical Physics*, 123(4):529–534, 1989.

[Bro04] C. Brouder. Trees, renormalization and differential equations. *BIT Numerical Mathematics*, 44(3):425–438, 2004.

[BU11] V. Betz and D. Ueltschi. Spatial Random Permutations and Poisson-Dirichlet Law of Cycle Lengths. *Electronic Journal of Probability*, 16:1173 – 1192, 2011.

[But63] J. C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *Journal of the Australian Mathematical Society*, 3(2):185–201, 1963.

[But72] J. C. Butcher. An algebraic theory of integration methods. *Mathematics of Computation*, 26(117):79–106, 1972.

[But08] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, Ltd, 2008.

[But18] J. C. Butcher. Trees and b-series. *Numerical Algorithms*, 81(4):1311–1325, 2018.

[Cay57] A. Cayley. XXVIII. on the theory of the analytical forms called trees. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 13(85):172–176, 1857.

[CBE+17] C. Costantini, P. De Blasi, S. N. Ethier, M. Ruggiero, and D. Spanò. Wright–Fisher construction of the two-parameter Poisson–Dirichlet diffusion. *The Annals of Applied Probability*, 27(3), 2017.

[CCG14] J. Cao, P. Chleboun, and S. Grosskinsky. Dynamics of Condensation in the Totally Asymmetric Inclusion Process. *Journal of Statistical Physics*, 155(3):523–543, 2014.

[CCG15] Y.-X. Chau, C. Connaughton, and S. Grosskinsky. Explosive condensation in symmetric mass transport models. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(11):P11031, 2015.

[CD20] S. Chatterjee and A. Dunlap. Constructing a solution of the $(2+1)$-dimensional KPZ equation. *The Annals of Probability*, 48(2), 2020.

[CDR10] P. Calabrese, P. Le Doussal, and A. Rosso. Free-energy distribution of the directed polymer at high temperature. *EPL (Europhysics Letters)*, 90(2):20002, 2010.

[CES21] G. Cannizzaro, D. Erhard, and P. Schönbauer. 2D anisotropic KPZ at stationarity: Scaling, tightness and nontriviality. *The Annals of Probability*, 49(1), 2021.

[CET21] G. Cannizzaro, D. Erhard, and F. Toninelli. Weak coupling limit of the Anisotropic KPZ equation. *arXiv preprint arXiv:2108.09046*, 2021.

[CG14] P. Chleboun and S. Grosskinsky. Condensation in stochastic particle systems with stationary product measures. *Journal of Statistical Physics*, 154:432–465, 2014.

[CGG22] P. Chleboun, S. Gabriel, and S. Grosskinsky. Poisson-Dirichlet asymptotics in condensing particle systems. *Electronic Journal of Probability*, 27:1 – 35, 2022.

[CGG23] P. Chleboun, S. Gabriel, and S. Grosskinsky. Size-biased diffusion limits and the inclusion process. *arXiv preprint arXiv:2304.09722*, 2023.

[CGR21] G. Carinci, C. Giardinà, and F. Redig. Consistent particle systems and duality. *Electronic Journal of Probability*, 26:1–31, 2021.

[CGT23] G. Cannizzaro, M. Gubinelli, and F. Toninelli. Gaussian fluctuations for the stochastic Burgers equation in dimension $d \geq 2$. *arXiv preprint arXiv:2304.05730*, 2023.

[CH16] A. Chandra and M. Hairer. An analytic BPHZ theorem for regularity structures. *arXiv preprint arXiv:1612.08138*, 2016.

[CNN22] C. Cosco, S. Nakajima, and M. Nakashima. Law of large numbers and fluctuations in the sub-critical and l^2 regions for she and kpz equation in dimension $d \geq 3$. *Stochastic Processes and their Applications*, 151:127–173, 2022.

[Com17] F. Comets. *Directed Polymers in Random Environments*. Springer International Publishing, 2017.

[COW22] I. Chevyrev, T. Oh, and Y. Wang. Norm inflation for the cubic nonlinear heat equation above the scaling critical regularity. *arXiv preprint arXiv:2205.14488*, 2022.

[CS19] I. Corwin and H. Shen. Some recent progress in singular stochastic partial differential equations. *Bulletin of the American Mathematical Society*, 57(3):409–454, 2019.

[CSY04] 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*. Mathematical Society of Japan, 2004.

[CSZ14] F. Caravenna, R. Sun, and N. Zygouras. The continuum disordered pinning model. *Probability Theory and Related Fields*, 164(1-2):17–59, 2014.

[CSZ17a] F. Caravenna, R. Sun, and N. Zygouras. Polynomial chaos and scaling limits of disordered systems. *Journal of the European Mathematical Society*, 19(1):1–65, 2017.

[CSZ17b] F. Caravenna, R. Sun, and N. Zygouras. Universality in marginally relevant disordered systems. *The Annals of Applied Probability*, 27(5):3050–3112, 2017.

[CSZ20] F. Caravenna, R. Sun, and N. Zygouras. The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Probab.*, 48(3):1086–1127, 2020.

[CSZ23] F. Caravenna, R. Sun, and N. Zygouras. The critical 2d stochastic heat flow. *Inventiones mathematicae*, pages 1–136, 2023.

[CT85] C. Cocozza-Thivent. Processus des misanthropes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 70(4):509–523, 1985.

[CT16] B. Cloez and M.-N. Thai. Fleming-Viot processes: two explicit examples. *Latin American Journal of Probability and Mathematical Statistics*, 13(1):337, 2016.

[CW17] A. Chandra and H. Weber. Stochastic PDEs, regularity structures, and interacting particle systems. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 26, pages 847–909, 2017.

[CY06] F. Comets and N. Yoshida. Directed polymers in random environment are diffusive at weak disorder. *The Annals of Probability*, 34(5):1746–1770, 2006.

[CZ21] F. Caravenna and L. Zambotti. Hairer's reconstruction theorem without regularity structures. *EMS Surveys in Mathematical Sciences*, 7(2):207–251, 2021.

[CZ23] C. Cosco and O. Zeitouni. Moments of partition functions of 2d gaussian polymers in the weak disorder regime-i. *Communications in Mathematical Physics*, 2023.

[DG22] A. Dunlap and Y. Gu. A forward-backward SDE from the 2d nonlinear stochastic heat equation. *The Annals of Probability*, 50(3), 2022.

[DH82] D. A. Dawson and K. J. Hochberg. Wandering Random Measures in the Fleming-Viot Model. *The Annals of Probability*, 10(3), 1982.

[DK99] P. Donnelly and T. G. Kurtz. Particle Representations for Measure-Valued Population Models. *The Annals of Probability*, 27(1), 1999.

[Dur08] R. Durrett. *Probability Models for DNA Sequence Evolution*. Springer New York, 2008.

[EG87] S. N. Ethier and R. C. Griffiths. The Infinitely-Many-Sites Model as a Measure-Valued Diffusion. *The Annals of Probability*, 15(2), 1987.

[EH05] M. R. Evans and T. Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *Journal of Physics A: Mathematical and General*, 38(19):R195, 2005.

[EK81] S. N. Ethier and T. G. Kurtz. The infinitely-many-neutral-alleles diffusion model. *Advances in Applied Probability*, 13(3):429–452, 1981.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov Processes*. John Wiley & Sons, Inc., 1986.

[EK87] S. N. Ethier and T. G. Kurtz. The Infinitely-Many-Alleles Model with Selection as a Measure-Valued Diffusion. In *Lecture Notes in Biomathematics*, pages 72–86. Springer Berlin Heidelberg, 1987.

[EK93] S. N. Ethier and T. G. Kurtz. Fleming–Viot Processes in Population Genetics. *SIAM Journal on Control and Optimization*, 31(2):345–386, 1993.

[ESS92] L. C. Evans, H. M. Soner, and P. E. Souganidis. Phase transitions and generalized motion by mean curvature. *Communications on Pure and Applied Mathematics*, 45(9):1097–1123, 1992.

[Eth14] S. Ethier. A property of Petrov's diffusion. *Electronic Communications in Probability*, 19, 2014.

[Fen10] S. Feng. *The Poisson-Dirichlet Distribution and Related Topics*. Springer Berlin Heidelberg, 2010.

[Fen12] Z. S. Feng. Diffusivity of rescaled random polymer in random environment in dimensions 1 and 2. *arXiv preprint arXiv:1201.6215*, 2012.

[FH20] P. K. Friz and M. Hairer. *A Course on Rough Paths*. Springer International Publishing, 2020.

[FPRW23] N. Forman, S. Pal, D. Rizzolo, and M. Winkel. Ranked masses in two-parameter fleming–viot diffusions. *Transactions of the American Mathematical Society*, 376(02):1089–1111, 2023.

[Fra78] L. E. Fraenkel. Formulae for high derivatives of composite functions. *Mathematical Proceedings of the Cambridge Philosophical Society*, 83(2):159–165, 1978.

[FSWX11] S. Feng, W. Sun, F.-Y. Wang, and F. Xu. Functional inequalities for the two-parameter extension of the infinitely-many-neutral-alleles diffusion. *Journal of Functional Analysis*, 260(2):399–413, 2011.

[Gab23] S. Gabriel. Central limit theorems for the (2+1)-dimensional directed polymer in the weak disorder limit. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 59(4):1984 – 2016, 2023.

[GIP15] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[GJ14] P. Gonçalves and M. Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Archive for Rational Mechanics and Analysis*, 212:597–644, 2014.

[GKR07] C. Giardinà, J. Kurchan, and F. Redig. Duality and exact correlations for a model of heat conduction. *Journal of mathematical physics*, 48(3):033301, 2007.

[GL12] C. Godrèche and J.-M. Luck. Condensation in the inhomogeneous zero-range process: an interplay between interaction and diffusion disorder. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(12):P12013, 2012.

[GRV11] S. Grosskinsky, F. Redig, and K. Vafayi. Condensation in the Inclusion Process and Related Models. *Journal of Statistical Physics*, 142(5):952–974, 2011.

[GRV13] S. Grosskinsky, F. Redig, and K. Vafayi. Dynamics of condensation in the symmetric inclusion process. *Electronic Journal of Probability*, 18:1–23, 2013.

[GRZ23] S. Gabriel, T. Rosati, and N. Zygouras. The Allen–Cahn equation with weakly critical random initial datum. *arXiv preprint arXiv:2311.04628*, 2023.

[Gu19] Y. Gu. Gaussian fluctuations from the 2d KPZ equation. *Stochastics and Partial Differential Equations: Analysis and Computations*, 8(1):150–185, 2019.

[Gub10] M. Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4):693–721, 2010.

[Gub18] M. Gubinelli. A panorama of singular SPDEs. In *Proceedings of the International Congress of Mathematicians (ICM 2018) (In 4 Volumes) Proceedings of the International Congress of Mathematicians 2018*, pages 2311–2338. World Scientific, 2018.

[GUW11] C. Goldschmidt, D. Ueltschi, and P. Windridge. Quantum Heisenberg models and their probabilistic representations. *Contemporary Mathematics*, pages 177–224, 2011.

[GV64] I. M. Gelfand and N. Ya. Vilenkin. *Generalized functions: Applications of harmonic analysis*, volume 4. Academic press, 1964.

[Hai13] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178(2):559–664, 2013.

[Hai14] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[Har74] A. B. Harris. Effect of random defects on the critical behaviour of ising models. *Journal of Physics C: Solid State Physics*, 7(9):1671–1692, 1974.

[HH85] D. A. Huse and C. L. Henley. Pinning and roughening of domain walls in Ising systems due to random impurities. *Physical Review Letters*, 54(25):2708–2711, 1985.

[HLR23] M. Hairer, K. Lê, and T. Rosati. The Allen–Cahn equation with generic initial datum. *Probability Theory and Related Fields*, pages 1–42, 2023.

[HNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. 1, Nonstiff problems*. Springer-Vlg, 1993.

[HW74] E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing*, 13(1):1–15, 1974.

[IS88] J. Z. Imbrie and T. Spencer. Diffusion of directed polymers in a random environment. *Journal of Statistical Physics*, 52(3-4):609–626, 1988.

[IT20] D. Ioffe and B. Tóth. Split-and-Merge in Stationary Random Stirring on Lattice Torus. *Journal of Statistical Physics*, 180(1-6):630–653, 2020.

[Jan97] S. Janson. *Gaussian hilbert spaces*. Number 129. Cambridge university press, 1997.

[JCG19] W. Jatuviriyapornchai, P. Chleboun, and S. Grosskinsky. Structure of the Condensed Phase in the Inclusion Process. *Journal of Statistical Physics*, 178(3):682–710, 2019.

[Jun21] S. Junk. New characterization of the weak disorder phase of directed polymers in bounded random environments. *Communications in Mathematical Physics*, 389(2):1087–1097, 2021.

[Jun23] S. Junk. Stability of weak disorder phase for directed polymer with applications to limit theorems. *Latin American Journal of Probability and Mathematical Statistics*, 20(1):861, 2023.

[Kal02] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[Kif97] Y. Kifer. The Burgers equation with a random force and a general model for directed polymers in random environments. *Probability Theory and Related Fields*, 108(1):29–65, 1997.

[Kin75] J. F. C. Kingman. Random Discrete Distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(1):1–15, 1975.

[KL99] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*. Springer Berlin Heidelberg, 1999.

[KS91] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 1991.

[KS21] S. Kim and I. Seo. Condensation and metastable behavior of non-reversible inclusion processes. *Communications in Mathematical Physics*, 382:1343–1401, 2021.

[Kup16] A. Kupiainen. Renormalization group and stochastic PDEs. In *Annales Henri Poincaré*, volume 17, pages 497–535. Springer, 2016.

[Kur69] T. G. Kurtz. Extensions of trotter's operator semigroup approximation theorems. *Journal of Functional Analysis*, 3(3):354–375, June 1969.

[Lax02] P. D. Lax. *Functional Analysis*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2002.

[Led01] M. Ledoux. *The Concentration of Measure Phenomenon*. Mathematical surveys and monographs. American Mathematical Society, 2001.

[LL10] G. F. Lawler and V. Limic. *Random Walk: A Modern Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[Lyo98] T. J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

[LZ22] D. Lygkonis and N. Zygouras. Edwards–Wilkinson fluctuations for the directed polymer in the full L^2-regime for dimensions d$\geq$3. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 58(1), 2022.

[LZ23] D. Lygkonis and N. Zygouras. Moments of the 2d directed polymer in the subcritical regime and a generalisation of the erdös-taylor theorem. *Communications in Mathematical Physics*, 401(3):2483–2520, 2023.

[Mar11] S. De Marco. Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions. *The Annals of Applied Probability*, 21(4):1282 – 1321, 2011.

[Mer57] R. H. Merson. An operational method for the study of integration processes. In *Proceedings of a conference on Data processing and automatic computing machines*, volume 1, pages 110–125, Salisbury, South Australia, 1957. Weapons Research Establishment.

[MH99] J. Marsden and M. J. Hoffman. *Basic Complex Analysis*. W.H. Freeman, New York, NY, 3 edition, 1999.

[Mor58] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[MP01] P. Mörters and Y. Peres. *Brownian Motion*. Cambridge University Press, 2001.

[MT93] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993.

[NN23] S. Nakajima and M. Nakashima. Fluctuations of two-dimensional stochastic heat equation and KPZ equation in subcritical regime for general initial conditions. *Electronic Journal of Probability*, 28(none), 2023.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.

[Pet09] L. A. Petrov. Two-parameter family of infinite-dimensional diffusions on the Kingman simplex. *Functional Analysis and Its Applications*, 43(4):279–296, 2009.

[Rud86] W. Rudin. *Real and Complex Analysis*. McGraw-Hill series in higher mathematics. McGraw-Hill Professional, New York, NY, 3 edition, 1986.

[RW09] M. Ruggiero and S. Walker. Countable representation for infinite dimensional diffusions derived from the two-parameter Poisson-Dirichlet process. *Electronic Communications in Probability*, 14, 2009.

[Sch91] B. Schmuland. A result on the infinitely many neutral alleles diffusion model. *Journal of Applied Probability*, 28(2):253–267, 1991.

[Sch05] O. Schramm. Compositions of random transpositions. *Israel Journal of Mathematics*, 147(1):221–243, 2005.

[Shi81] T. Shiga. Diffusion processes in population genetics. *Kyoto Journal of Mathematics*, 21(1), 1981.

[Sin95] Y. Sinai. A remark concerning random walks with random potentials. *Fundamenta Mathematicae*, 147(2):173–180, 1995.

[Spi70] F. Spitzer. Interaction of Markov processes. *Advances in Mathematics*, 5(2):246–290, 1970.

[Str03] R. S. Strichartz. *A Guide to Distribution Theory and Fourier Transforms*. WORLD SCIENTIFIC, 2003.

[SZ96] R. Song and X. Y. Zhou. A remark on diffusion of directed polymers in random environments. *Journal of Statistical Physics*, 85(1-2):277–289, 1996.

[Tao22] R. Tao. Gaussian fluctuations of a nonlinear stochastic heat equation in dimension two. *Stochastics and Partial Differential Equations: Analysis and Computations*, 2022.

[Tro58] H. F. Trotter. Approximation of semi-groups of operators. *Pacific Journal of Mathematics*, 8(4):887 – 919, 1958.

[Uel17] D. Ueltschi. Universal behaviour of 3d loop soup models. *Proceedings of 6th Warsaw School of Statistical Physics*, pages 65–100, 2017.

[Var06] V. Vargas. A local limit theorem for directed polymers in random media: the continuous and the discrete case. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 42(5):521–534, 2006.

[WE12] B. Waclaw and M. R. Evans. Explosive Condensation in a Mass Transport Model. *Physical Review Letters*, 108(7), 2012.

[Wil51] E. Wild. On Boltzmann's equation in the kinetic theory of gases. *Mathematical Proceedings of the Cambridge Philosophical Society*, 47(3):602–609, 1951.

[You36] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica*, 67(0):251–282, 1936.

[Zam21] L. Zambotti. A brief and personal history of stochastic partial differential equations. *Discrete and Continuous Dynamical Systems - A*, 41(1):471–487, 2021.

[ZZMWD04] M. P. W. Zerner, O. Zeitouni, E. Mayer-Wolf, and P. Diaconis. The Poisson-Dirichlet law is the unique invariant distribution for uniform split-merge transformations. *The Annals of Probability*, 32(1B):915–938, 2004.