CORE
Metadata, citation and similar papers at core.ac.uk
Provided by INRIA a CCSD electronic archive server

Coalescent tree based functional representations for some Feynman-Kac particle models

Pierre del Moral, Frédéric Patras, Sylvain Rubenthaler

To cite this version:

Pierre del Moral, Frédéric Patras, Sylvain Rubenthaler. Coalescent tree based functional representations for some Feynman-Kac particle models. [Research Report] PI 1813, 2006, pp.45. inria-00090246

HAL Id: inria-00090246
https://hal.inria.fr/inria-00090246

Submitted on 29 Aug 2006

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

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

I R I S A
INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

PUBLICATION
INTERNE
N° 1813

COALESCENT TREE BASED FUNCTIONAL REPRESENTATIONS FOR SOME FEYNMAN-KAC PARTICLE MODELS

PIERRE DEL MORAL , FRÉDÉRIC PATRAS, SYLVAIN RUBENTHALER

ISSN 1166-8687

IRISA
CAMPUS UNIVERSITAIRE DE BEAULIEU - 35042 RENNES CEDEX - FRANCE

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES
Campus de Beaulieu – 35042 Rennes Cedex – France
Tél. : (33) 02 99 84 71 00 – Fax : (33) 02 99 84 71 71
http://www.irisa.fr

Coalescent tree based functional representations for some Feynman-Kac particle models

Pierre Del Moral[*] [**], Frédéric Patras[*], Sylvain Rubenthaler[*]

Systèmes numériques
Projet ASPI

Publication interne n ° 1813 — Août 2006 — 45 pages

Abstract: We design a theoretic tree-based functional representation of a class of Feynman-Kac particle distributions, including an extension of the Wick product formula to interacting particle systems. These weak expansions rely on an original combinatorial, and permutation group analysis of a special class of forests. They provide refined non asymptotic propagation of chaos type properties, as well as sharp $\mathbb{L}_p$-mean error bounds, and laws of large numbers for U-statistics. Applications to particle interpretations of the top eigenvalues, and the ground states of Schrödinger semigroups are also discussed.

Key-words: Feynman-Kac semigroups, interacting particle systems, trees and forests, automorphism groups, combinatorial enumeration.

(Résumé : tsvp)

The second author was supported by the ANR grant AHBE 05-42234

Mathematics Subject of Classification : Primary: 47D08, 60C05, 60K35, 65C35; Secondary: 31B10, 60J80, 65C05, 92D25.

[*] Université de Nice–Sophia Antipolis, Laboratoire Jean–Alexandre Dieudonné (UMR 6621), Parc Valrose, 06108 NICE, France. e–mail : {delmoral,patras,rubentha}@math.unice.fr
[**] IRISA / INRIA, Campus de Beaulieu, 35042 RENNES, France.

Centre National de la Recherche Scientifique (UMR 6074) Université de Rennes 1 – Insa de Rennes
Institut National de Recherche en Informatique et en Automatique – unité de recherche de Rennes

Représentations fonctionnelles en terme d'arbre coalescent pour des modèles particulaires de Feynman-Kac

Résumé : Nous proposons des représentations fonctionnelles en terme d'arbre coalescent, pour une classe de distributions particulaires de type Feynman-Kac, ce qui inclut en particulier une extension de la formule de Wick aux systèmes de particules en interaction. Les développements polynômiaux en $1/N$ obtenus, où N désigne le nombre de particules, reposent sur une analyse combinatoire originale et basée sur les groupes de permutations, d'une classe particulière de forêts. Ces développements permettent d'obtenir une version fine de propriétés non asymptotiques de type propagation du chaos, ainsi que des bornes d'erreur précises dans $\mathbb{L}_p$, et des lois des grands nombres pour les U-statistiques. Nous considérons aussi des applications à l'interprétation probabiliste de la valeur propre principale et des états propres du semi groupe de Schrödinger.

Mots clés : semigroupes de Feynman-Kac, systèmes de particules en interaction, arbres et forêts, groupes d'automorphisme, dénombrement combinatoire.

Contents

1 Introduction and formulation of the main results **3**
1.1 Introduction . . . 3
1.2 Feynman-Kac semigroups, and some application areas . . . 4
1.3 Mean field interacting particle models . . . 7
1.4 Particle block distributions . . . 8
1.5 Forest based expansion formulas . . . 9

2 Particle measures expansions on forests **12**
2.1 A preliminary stochastic tensor product formula . . . 12
2.2 Laurent expansions . . . 14

3 Combinatorial methods for counting forests and jungles **16**
3.1 Some terminology on trees and forests . . . 16
3.2 On the algebraic structure of trees and forests . . . 18
3.3 Automorphism groups on jungles . . . 20
3.4 Orbit sets of planar forests . . . 20
3.5 An inductive method for counting jungles . . . 22
3.6 Wreath product representation . . . 24
3.7 Hilbert series method for forests enumeration . . . 26
3.8 Some forests expansions . . . 30
3.9 A Wick product formula on forests . . . 30

4 Extension to path-space models **33**
4.1 Colored trees, forests, and jungles . . . 33
4.2 Automorphism groups on colored jungles . . . 35
4.3 Feynman-Kac semigroups . . . 36
4.4 Unnormalized particle measures . . . 37
4.5 Propagations of chaos type expansions . . . 40

1 Introduction and formulation of the main results

1.1 Introduction

The field of Feynman-Kac semigroups, and their interacting particle models is one of the most active contact points between probability, theoretical chemistry, quantum physics, and engineering sciences, including rare event analysis, and advanced signal processing. Whereas it is clearly out of the scope of the present article to review these applications, some of them are briefly pointed out in the further development of this introduction. For a rather thorough discussion, the interested reader is recommended to consult the pair of books [4, 5], and the references therein.

The common aspect of all these applications is that they represent path integration of free Markov particle evolutions, weighted by some potential functions. The models we consider belong to the general class of branching and interacting particle systems. Particles explore the state space as a free Markov evolution; during their exploration particles with low potential are killed, while the ones with high potential value duplicate. From a more probabilistic point of view, these evolutionary genetic type particle models can be interpreted as stochastic linearization, or as sophisticated interacting acceptation-rejection type simulation techniques.

During the last two decades, the asymptotic analysis of these models has been developed in various directions, including propagation of chaos analysis, $\mathbb{L}_p$-mean error estimates, central limit type theorems, and large deviation principles. The purpose of the present work is to develop algebraic tree-based functional representations of particle block distributions, stripped of all analytical superstructure, and probabilistic irrelevancies. These weak expansions rely on an original combinatorial, and permutation group analysis on a special class of classical and colored forests that parametrize naturally the trajectories of interacting particle systems. In order to modelize and compute the corresponding path integrals, we develop an original differential and combinatorial calculus, and take advantage of the natural permutation group action on particle configurations to handle the combinatorial complexity of the problem. Wreath products of permutation groups appear in the process, and new closed algebraic formulae for labelled forests enumeration are obtained. We also discuss Hilbert series

techniques for counting the forests relevant to our analysis, that is with prescribed number of vertices at each level, or with given coalescence degrees.

Let us mention that forests, colored graphs and their combinatorics have appeared recently in various fields such as in theoretical physics and gaussian matrix integral models [6], renormalization theory in high-energy physics or Runge-Kutta methods, two fields where the structure and complexity of perturbative expansions has required the development of new tools [1, 3]. To the best of our knowledge, their introduction in the analysis of Feynman-Kac and interacting particle models is new. The Laurent type integral representations presented in this article do not only reveal the combinatorial essence of interacting particle distributions. They also provide refined non asymptotic propagation of chaos type properties, as well as explicit formulae for $\mathbb{L}_p$-mean error bounds, and laws of large numbers for U-statistics.

The article is divided into four main parts, devoted respectively to the precise description of Feynman-Kac particle models and tree-based integral expansions, to the proof of these formulae, to the combinatorial and group-theoretic analysis, and to the extension to path space models with applications to propagation of chaos type properties .

In section 1.2, we describe a general class of Feynman-Kac semigroups associated with some free Markov evolution, and with a collection of potential functions. We motivate these abstract mathematical models with a brief discussion on a particle absorption model arising in the spectral analysis of Schrödinger operators. The mean field particle models are presented in section 1.3. Section 1.4 is concerned with the description of the corresponding particle block distributions. The main results of the article are housed in section 1.5.

Section 2 is essentially concerned with the proof of the tree-based Laurent expansions presented in section 1.5.

Section 3 is devoted to the combinatorial analysis of the class of forests involved in the tree-based Laurent expansions. The analysis includes some exotic planar forests, the branchs of which are allowed to cross, referred to in the following as jungles. The first subsection, section 3.1, starts with a short review on the vocabulary of trees and forests. In section 3.2, and section 3.3, we provide an algebraic representation of forests and jungles in terms of sequences on maps. Sections 3.4, 3.5 and 3.6 introduce the algebraic tools relevant to the combinatorial analysis. Permutation group actions on jungles are introduced. From the Feynman-Kac modelling point of view, they correspond to the action of the permutation group acting naturally on particle systems. This action is studied in detail. For example, a wreath product representation of the stabilizers is obtained. Closed inductive formulas for interacting particle models path enumeration are derived in section 3.5. Finally, in section 3.7, we design an Hilbert series method for counting forests and the other famillies of relevant combinatorial and geometrical objects. As an application of these results on trees and forests, some explicit calculations of the first order terms of the expansions obtained in section 2 are given in section 3.8, and a generalized Wick formula for forests is obtained in section 3.9.

Section 4 is essentially concerned with the extension of the previous analysis to Feynman-Kac particle models on path spaces. The corresponding tree-based expansions involve a new class of forests, with two distinguished colored vertices. The combinatorial analysis of these objects is slightly more involved, but essentially follows the same line of development as before. This study is summarized in section 4.1, and section 4.2. The next two sections, section 4.3, and section 4.4, are devoted respectively to unnormalized Feynman-Kac semigroups on path spaces, and to the corresponding tree-based Laurent type expansions. In the final section, section 4.5, we use the differential forest expansion machinery developed previously in this article to derive precise propagations of chaos type properties of interacting particle models, including strong expansions of the particle block distributions with respect to Zolotarev type seminorms, sharp $\mathbb{L}_p$-mean error estimates, and a law of large numbers for U-statistics associated to interacting particle models.

1.2 Feynman-Kac semigroups, and some application areas

To avoid any state space restrictions, and any unnecessary topological assumptions, this article is concerned with abstract mathematical models in general measurable state spaces. The main advantage of this general set-up comes from the fact that it applies without further work to the analysis of path space Feynman-Kac models, and their genealogical tree based interpretations.

Before we continue with the body of the introduction, we already present a series of definitions, the benefits of which are reaped in the subsequent sections. We let $(E_n, \mathcal{E}_n)_{n\geq 0}$ be a collection of measurable state spaces. We denote respectively by $\mathcal{M}(E_n)$, $\mathcal{P}(E_n)$, and $\mathcal{B}_b(E_n)$, the set of all finite signed measures on $(E_n, \mathcal{E}_n)$, the convex subset of all probability measures, and the Banach space of all bounded and measurable functions f on E_n, equipped with the uniform norm $\|f\| = \sup_{x_n\in E_n} |f(x_n)|$. We let $\mu_n(f_n) = \int_{E_n} \mu_n(dx_n)\ f_n(x_n)$, be the Lebesgue integral of a function $f_n \in \mathcal{B}_b(E_n)$, with respect to a measure $\mu_n \in \mathcal{M}(E_n)$. We equip $\mathcal{M}(E_n)$ with the total variation norm $\|\mu_n\|_{\mathrm{tv}} = \sup_{f\in\mathcal{B}_b(E_n):\|f\|\leq 1} |\mu_n(f)|$. We recall that a bounded integral operator Q_n

from E_{n-1} into E_n, such that, for any $f_n \in \mathcal{B}_b(E_n)$, the functions

$$Q_n(f_n) \; : \; x_{n-1} \in E_{n-1} \mapsto Q_n(f_n)(x_{n-1}) = \int_{E_{n-1}} Q_n(x_{n-1}, dx_n) \; f_n(x_n) \; \in \mathbb{R}$$

are $\mathcal{E}_{n-1}$-measurable, and bounded, generates a dual operator $\mu_{n-1} \mapsto \mu_{n-1}Q_n$ from $\mathcal{M}(E_{n-1})$ into $\mathcal{M}(E_n)$, and defined by $(\mu_{n-1}Q_n)(f_n) := \mu_{n-1}(Q_n(f_n))$.

For a bounded integral operator Q_1 from E_0 into E_1, and an operator Q_2 from E_1 into E_2, we denote by Q_1Q_2 the composition operator from E_0 into E_2, defined for any $f_2 \in \mathcal{B}_b(E_2)$ by $(Q_1Q_2)(f_2) := Q_1(Q_2(f_2))$. The tensor power $Q_n^{\otimes q}$ represents the bounded integral operator on E_n^q, defined for any $F \in \mathcal{B}_b(E_n^q)$ by

$$Q_n^{\otimes q}(F)(x_n^1, \ldots, x_n^q) = \int_{E_n^q} \left[Q_n(x_n^1, dy_n^1) \; \ldots \; Q_n(x_n^q, dy_n^q)\right] \; F(y_n^1, \ldots, y_n^q)$$

We consider a distribution η_0 on E_0, a collection of Markov transitions $M_n(x_{n-1}, dx_n)$ from E_{n-1} into E_n, and a collection of $\mathcal{E}_n$-measurable, and bounded potential functions G_n on the state spaces E_n. To simplify the presentation, and avoid unnecessary technical discussion, we shall suppose that the potential functions are chosen such that

$$0 < \inf_{x_n \in E_n} G_n(x_n) \leq \sup_{x_n \in E_n} G_n(x_n) < \infty \tag{1.1}$$

We associate to these objects the Feynman-Kac measures defined for any function $f_n \in \mathcal{B}_b(E_n)$ by the following formulae

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E}[f_n(X_n) \prod_{0 \leq k < n} G_k(X_k)] \tag{1.2}$$

In (1.2), $(X_n)_{n \geq 0}$ represents a Markov chain, taking values in the state spaces $(E_n)_{n \geq 0}$, with initial distribution η_0 on E_0, and elementary transitions M_n from E_{n-1} into E_n. The choice of non homogeneous state spaces E_n is not innocent. In several application areas the underlying Markov model is a path-space Markov chain

$$X_n = (X'_0, \ldots, X'_n) \in E_n = (E'_0 \times \ldots \times E'_n) \tag{1.3}$$

The elementary prime variables X'_n represent an elementary Markov chain taking values in some measurable spaces $(E'_n, \mathcal{E}'_n)$. For instance, in macro-molecular analysis, the elementary variables X'_n represents the monomers in a directed chain X_n, with polymerization degree n. In this context, the potential functions G_n reflect the intermolecular attraction or repulsion forces between the monomers, and the corresponding Feynman-Kac model represents the distribution of the polymer configurations in a solvent. In signal processing, and more particularly in filtering problems, the chain X_n represent the paths of a given signal, and the potential functions G_n reflect the likelihood of their terminal values X'_n, with respect to the sequence of observations delivered by the sensors. In this context, the corresponding Feynman-Kac model represents the conditional distribution of the path of the signal, given the sequence of observations. To motivate this article, let us also mention an important, and more classical physical interpretation of these Feynman-Kac flows. In theoretical particle physics, these models represents the distribution of a particle evolving in an absorbing medium, with obstacles related to potential functions G_n, taking values in $[0, 1]$. In this context, the particle X_n^c evolves according to two separate mechanisms. Firstly, it moves from a site $x_{n-1} \in E_{n-1}$, to another $x_n \in E_n$ according to elementary transitions $M_n(x_{n-1}, dx_n)$. Then, it is absorbed with a probability $1 - G_n(x_n)$, and placed in an auxiliary cemetery state $X_n^c = c$; otherwise it remains in the same site. If we let T be the random absorption time, then it is not difficult to check that

$$\gamma_n(f_n) = \mathbb{E}[f_n(X_n^c) \; 1_{T \geq n}] \quad \text{and} \quad \eta_n(f_n) = \mathbb{E}[f_n(X_n^c) \mid T \geq n] \tag{1.4}$$

In time homogeneous settings ($G_n = G$ and $M_n = M$), we have

$$\gamma_n(1) = \mathbb{P}\left(T \geq n\right) \sim e^{-\lambda \; n}$$

The positive constant λ is a measure of the strength and trapping effects of the obstacle. By a lemma of Varadhan's, λ coincide with the logarithmic Lyapunov exponent λ_0 of the transition operator $Q(x, dy) = G(x) \; M(x, dy)$. Whenever it exists, the corresponding eigenfunction h of Q represents the ground state of the operator Q.

These Feynman-Kac models can also be regarded as a natural time discretization of continuous time particle absorption models associated with some Schrödinger operator. To be more precise, let us suppose that

$M(x,dy) = M^{\Delta}(x,dy)$ represent the transitions of a continuous time Markov process with infinitesimal generator L, during a short period of time Δ; in the sense that

$$L(f)(x) = \lim_{\Delta \to 0} \frac{1}{\Delta} \left[M^{\Delta}(f)(x) - f(x) \right]$$

for sufficiently regular functions f on E. Let us also suppose that the potential functions have the form $G(x) = G^{\Delta}(x) = e^{-V(x)\Delta}$, for some non negative energy function V. In this case, it is readily checked that the infinitesimal generator of the particle absorption transitions

$$Q^{\Delta}(x,dy) = e^{-V(x)\Delta} \; M^{\Delta}(x,dy)$$

is given by the Schrodinger operator

$$L(f)(x) - V(x)f(x) = \lim_{\Delta \to 0} \frac{1}{\Delta} \left[Q^{\Delta}(f)(x) - f(x) \right]$$

Under reasonably weak assumptions, the pair (λ, h) coincide respectively with the top eigenvalue, and the ground state of the Schrodinger operator $L^V = L - V$. The estimation of these spectral quantities has recently received a lot of attention in modern numerical physics, and quantum chemistry.

A natural, and common key idea in most of the Monte-Carlo approximation models is to simulate a population of walkers or particles mimicking the evolution of a particle in an absorbing medium. These models can be interpreted as a natural mean field particle approximation of the distribution flow η_n. We end this section with a precise description of these nonlinear models. By the Markov property and the multiplicative structure of (1.2), it is easily checked that the flow $(\eta_n)_{n \geq 0}$ satisfies the following equation

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) \tag{1.5}$$

The transformations $\Phi_{n+1} : \mathcal{P}(E_n) \to \mathcal{P}(E_{n+1})$ are defined for any pair $(\eta_n, f_{n+1}) \in (\mathcal{P}(E_n) \times \mathcal{B}_b(E_{n+1}))$ as follows

$$\Phi_{n+1}(\eta_n)(f_{n+1}) = \frac{\eta_n(Q_{n+1}(f_{n+1}))}{\eta_n(Q_{n+1}(1))} \quad \text{with} \quad Q_{n+1}(x_n, dx_{n+1}) = G_n(x_n) \times M_{n+1}(x_n, dx_{n+1})$$

Notice that the evolution semigroup $(Q_{p,n})_{0 \leq p \leq n}$ of the unnormalized distribution flow is given by

$$\forall 0 \leq p < n \qquad \gamma_n = \gamma_p Q_{p,n} \quad \text{with} \quad Q_{p,n} = Q_{p,n-1} Q_n \tag{1.6}$$

and the unnormalized Feynman-Kac measures γ_n satify $\gamma_n(f_n) = \gamma_{n-1}(Q_n(f_n))$. Equivalently, they can be expressed in terms of the flow $(\eta_p)_{0 \leq p \leq n}$ with the formulae

$$\gamma_n(f_n) = \eta_n(f_n) \; \times \; \gamma_n(1) \quad \text{with} \quad \gamma_n(1) = \prod_{0 \leq p < n} \eta_p(G_p) \tag{1.7}$$

for any $f_n \in \mathcal{B}_b(E_n)$. Returning to the physical interpretation, this product formula shows that the Lyapunov exponent λ is related to the asymptotic mean logarithm value of the potential $G_n = G$ with respect to the measures η_n

$$\frac{1}{n} \sum_{0 \leq p \leq n} \log \eta_p(G) \simeq -\lambda \tag{1.8}$$

Much more is true. In the case where the homogeneous transitions $M_n = M$ are reversible with respect to some measure μ, the ground state represents the limiting fixed point measure of the distribution flow η_n; in the sense that for any bounded measurable function f, we have that

$$\frac{1}{n} \sum_{0 \leq p \leq n} \eta_p(Gf)/\eta_p(G) \simeq \mu(hf)/\mu(h) \tag{1.9}$$

1.3 Mean field interacting particle models

A natural mean field particle model associated with the nonlinear Feyman-Kac flow (1.5) is the E_n^N-valued Markov chain $\xi_n^{(N)} = (\xi_n^{i,N})_{1\leq i\leq N}$ with elementary transitions defined for any $F_n^N \in \mathcal{B}_b(E_n^N)$ by

$$\mathbb{E}\left(F_n^N(\xi_n^{(N)}) \mid \xi_{n-1}^{(N)}\right) = \Phi_n\left(m\left(\xi_{n-1}^{(N)}\right)\right)^{\otimes N}(F_n^N) \quad \text{with} \quad m\left(\xi_{n-1}^{(N)}\right) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi_{n-1}^{i,N}} \tag{1.10}$$

In other terms, given the configuration $\xi_{n-1}^{(N)}$ at rank $(n-1)$, the particle system $\xi_n^{(N)}$ at rank n, consists of N independent and identically distributed random variables with common distribution $\Phi_n\left(m\left(\xi_{n-1}^{(N)}\right)\right)$. The initial configuration $\xi_0^{(N)}$ consists of N independent and identically distributed random variables with distribution η_0. Although the dependency of the model on N is strong, due to the mean-field nature of the model, we abbreviate $(\xi_n^{i,N})_{1\leq i\leq N}$ to $(\xi_n^i)_{1\leq i\leq N}$ when N is fixed, excepted when we want to emphasize explicitely the dependency on N of the model, e.g. as in the end of the present section.

Notice that

$$\Phi_n\left(m\left(\xi_{n-1}^{(N)}\right)\right)(dx_n) = \frac{1}{\sum_{j=1}^{N} G_{n-1}(\xi_{n-1}^j)}\left(\sum_{i=1}^{N} G_{n-1}(\xi_{n-1}^i)\; M_n(\xi_{n-1}^i, dx_n)\right)$$

so that the particle model evolves as a genetic type model with proportional selections, and mutation transitions dictated by the pair of potential-transition (G_{n-1}, M_n). In numerical physics, the selection transition is often called reconfiguration of the population of walkers. In the context of path space Markov chain model (1.3), it is important to notice that the particle interpretation model is again a path particle model. During the selection stage, we select a path particle according to its fitness, and the mutation stage consists in extending the selected path with an elementary transition dictated by the Markov transitions of the elementary prime chain. In this context, the particle model can be interpreted as a genealogical particle evolution model.

The approximation measures (η_n^N, γ_n^N) associated with the pair of Feynman-Kac measures (η_n, γ_n) are defined by the empirical occupation measures

$$\eta_n^N = m\left(\xi_n^{(N)}\right)$$

and the unnormalized particle distributions γ_n^N are defined for any $f_n \in \mathcal{B}_b(E_n)$ by

$$\gamma_n^N(f_n) = \eta_n^N(f_n) \;\times\; \gamma_n^N(1) \quad \text{with} \quad \gamma_n^N(1) = \prod_{0\leq p<n} \eta_p^N(G_p) \tag{1.11}$$

For a rather complete asymptotic analysis of these measures we again refer the reader to [4], and references therein. In particular, for any $f_n \in \mathcal{B}_b(E_n)$, we have the following almost convergence results

$$\lim_{N\to\infty} \eta_n^N(f_n) = \eta_n(f_n) \quad \text{and} \quad \lim_{N\to\infty} \gamma_n^N(f_n) = \gamma_n(f_n)$$

Therefore, the particle interpretations of the pair objects (λ, h) introduced in (1.8) and (1.9), are defined by simply with replacing η_n by the particle occupation measure η_n^N. It is also well known that γ_n^N is an unbias approximation measure of γ_n, in the sense that

$$\mathbb{E}(\gamma_n^N(f_n)) = \gamma_n(f_n)$$

However, it turns out that for any $n \geq 1$

$$\mathbb{E}(\eta_n^N(f_n)) = \mathbb{E}(f_n(\xi_n^{1,N})) \neq \eta_n(f_n)$$

This means that the mean field particle interpretation model is not an exact sampling algorithm of the distributions η_n. In practice, it is clearly important to analyze these quantities, and more generally the distribution of particle blocks of any size $q \leq N$

$$\mathbb{P}_{n,q}^N = \text{Law}\left(\xi_n^{1,N}, \ldots, \xi_n^{q,N}\right) \in \mathcal{P}(E_n^q)$$

1.4 Particle block distributions

For any pair of integers $(l,m) \in (\mathbb{N}^\star)^2$, we set $[l] = \{1,\ldots,l\}$, and $[l]^{[m]}$ the set of mappings a from $[m]$ into $[l]$. By $|a|$, we denote the cardinality of the set $a([m])$, and for any $1 \leq p \leq l$ we set

$$[l]_p^{[m]} = \{a \in [l]^{[m]} \ : \ |a| = p\}$$

The number $m-|a|$ will be refered to as the coalescence number of a. To simplify the presentation, when $l > m$, we also denote by $\langle m,l\rangle \left(= [l]_m^{[m]}\right)$ the set of all

$$(l)_m := \frac{l!}{(l-m)!}$$

one to one mappings from $[m]$ into $[l]$, and by $\mathcal{G}_l = \langle l,l\rangle$ the symmetric group of all permutations of $[l]$.

From the pure mathematical point of view, the measures $\mathbb{P}_{n,q}^N$ are better understood when they are connected with the q-tensor product occupation measures (resp. the restricted q-tensor product occupation measures) on E_n^q defined by

$$(\eta_n^N)^{\otimes q} = \frac{1}{N^q} \sum_{a\in[N]^{[q]}} \delta_{(\xi_n^{a(1)},\ldots,\xi_n^{a(q)})} \quad \text{and} \quad (\eta_n^N)^{\odot q} = \frac{1}{(N)_q} \sum_{a\in\langle q,N\rangle} \delta_{(\xi_n^{a(1)},\ldots,\xi_n^{a(q)})}$$

Notice that these measures are, by construction, symmetry-invariant. That is, for any $F \in \mathcal{B}_b(E_n^q)$

$$\left(\eta_n^N\right)^{\otimes q}(F) = \left(\eta_n^N\right)^{\otimes q}((F)_{\mathrm{sym}}) \quad \text{and} \quad \left(\eta_n^N\right)^{\odot q}(F) = \left(\eta_n^N\right)^{\odot q}((F)_{\mathrm{sym}})$$

with the symmetrization operator

$$\forall q \geq 1 \qquad F \in \mathcal{B}_b(E_n^q) \mapsto (F)_{\mathrm{sym}} := \frac{1}{p!} \sum_{\sigma\in\mathcal{G}_q} D_\sigma F$$

In the above displayed definition, the operator D_b stands for the Markov transition from E_n^q into itself, associated with a mapping $b \in [q]^{[q]}$, and defined by

$$D_b(F)(x_n^1,\ldots,x_n^q) = F(x_n^{b(1)},\ldots,x_n^{b(q)})$$

for any $F \in \mathcal{B}_b(E_n^q)$, and $(x^1,\ldots,x^q) \in E_n^q$. In particular, we may assume, without restriction, in our forthcoming computations on q-tensor product occupation measures, that F is a symmetric function. That is, for all $s \in \mathcal{G}_q$, $D_sF = F$. We write from now on, $\mathcal{B}_b^{sym}(E_n^q)$ for the set of all symmetric functions in $\mathcal{B}_b(E_n^q)$.

Notice also the following symmetry property, essential in view of all the forthcoming computations. Since, conditional to $\xi_{n-1}^{(N)}$, the ξ_n^i are i.i.d., for any $F \in \mathcal{B}_b(E_n^q)$ and any $a,b \in \langle q,N\rangle^2$, we have

$$\mathbb{E}\left[F(\xi_n^{a(1)},\ldots,\xi_n^{a(q)})\right] = \mathbb{E}\left[F(\xi_n^{b(1)},\ldots,\xi_n^{b(q)})\right].$$

For instance, we have that

$$\mathbb{E}\left(\left(\eta_n^N\right)^{\odot q}(F)\right) = \mathbb{E}\left(F\left(\xi_n^1,\ldots,\xi_n^q\right)\right) = \mathbb{P}_{n,q}^N(F) \tag{1.12}$$

As in the case of elementary particle blocks of size one, the precise analysis of these q-tensor product measures is intimately related to the unnormalized pair particle measures $((\gamma_n^N)^{\odot q}, (\gamma_n^N)^{\otimes q})$ defined for any $F \in \mathcal{B}_b(E_n^q)$ by

$$(\gamma_n^N)^{\odot q}(F) = (\eta_n^N)^{\odot q}(F) \times (\gamma_n^N(1))^q \quad \text{and} \quad (\gamma_n^N)^{\otimes q}(F) = (\eta_n^N)^{\otimes q}(F) \times (\gamma_n^N(1))^q$$

We consider now the nonnegative measure $\mathbb{Q}_{n,q}^N$, indexed by the particle block sizes, on the product state spaces E_n^q, and defined for any $F \in \mathcal{B}_b(E_n^q)$ by

$$\mathbb{Q}_{n,q}^N(F) := \mathbb{E}((\gamma_n^N)^{\otimes q}(F))$$

In the further development of this article, we shall prove the following formula (2.4)

$$\mathbb{E}((\gamma_n^N)^{\odot q}(F)) = \mathbb{E}((\gamma_{n-1}^N)^{\odot q} Q_n^{\otimes q}(F)) = \mathbb{Q}_{n-1,q}^N(Q_n^{\otimes q}(F))$$

This simple observation indicates that the distribution flow $(\mathbb{Q}_{n,q}^N)_{n\geq 0}$ has a linear evolution semigroup, for any particle block size $q \leq N$. Much more is true. As another consequence of the results presented in [4], the sequence of distributions $\mathbb{Q}_{n,q}^N$ converges, as N tends to infinity, to the distribution of q non absorbed, and independent particles $(X_n^{c,i})_{1\leq i\leq q}$ evolving in the original absorbing medium. That is, we have that

$$\lim_{N\to\infty} \mathbb{Q}_{n,q}^N(F) \quad = \quad \gamma_n^{\otimes q}(F) = \mathbb{E}\left(F(X_n^{c,1},\ldots,X_n^{c,q})\ 1_{T^1\geq n}\ \ldots\ 1_{T^q\geq n}\right)$$

for any $F \in \mathcal{B}_b(E_n^q)$, and where T^i stands for the random absorption time sequence of the chain $(X_k^{c,i})_{k\geq 0}$, with $1 \leq i \leq q$.

This article is mainly concerned with explicit functional expansions of the deterministic measures $\mathbb{P}_{n,q}^N \in \mathcal{P}(E_n^q)$, and $\mathbb{Q}_{n,q}^N \in \mathcal{M}(E_n^q)$, with respect to the precision parameter N. These (Laurent type) expansions reflect the complete interaction structure of the particle model. Roughly speaking, the k-th order terms of these integral representations represent the $1/N^k$-contributions of mean field particle scenarios with an interaction degree k (see Thm. 1.2).

To describe more precisely these functional representations, it is convenient to introduce the abstract definition of the derivative of a sequence of finite signed measures.

Definition 1.1. *We let $(\Theta^N)_{N\geq 1} \in \mathcal{M}(E)^{\mathbb{N}}$, be a uniformly bounded sequence of signed measures on a measurable space $(E,\mathcal{E})$, in the sense that $\sup_{N\geq 1} \|\Theta^N\|_{\mathrm{tv}} < \infty$. We suppose that Θ^N strongly converges to some measure $\Theta \in \mathcal{M}(E)$, as $N \uparrow \infty$, in the sense that*

$$\forall F \in \mathcal{B}_b(E) \qquad \lim_{N\uparrow\infty} \Theta^N(F) = \Theta(F)$$

The discrete derivative of the sequence $(\Theta^N)_{N\geq 1}$ is the sequence of measures $(\partial\Theta^N)_{N\geq 1}$ defined by

$$\partial\Theta^N := N\ \left[\Theta^N - \Theta\right]$$

We say that Θ^N is differentiable, if $\partial\Theta^N$ is uniformly bounded, and if it strongly converges to some measure $\partial\Theta \in \mathcal{M}(E)$, as $N \uparrow \infty$. The discrete derivative of a discrete derivative $\partial\Theta^N$ (of a differentiable sequence) is called the second discrete derivative and it is denoted by

$$\partial^2\Theta^N = N\left[\partial\Theta^N - \partial\Theta\right]$$

The discrete derivative $\partial\Theta^N$ of a differentiable sequence can itself be differentiable. In this situation, the derivative of the discrete derivative is called the second derivative and it is denoted by $\partial^2\Theta = \partial(\partial\Theta)$, and so on.

At this point, it is convenient to make some comments. Firstly, the strong form of the weak convergence of measures considered above corresponds to the s-topology on the set of measures, defined as the coarsest topology for which all functionals $\mu \in \mathcal{M}(E) \mapsto \mu(\varphi) \in \mathbb{R}$, $\varphi \in \mathcal{B}_b(E)$, are continuous. For a more detailed discussion on the s-topology in the context of probability measures, we refer the reader to the article [9]. A sequence of real numbers can be interpreted as a sequence of measures with a constant value, so that the above definition applies to any sequence of real numbers. We also notice that the Euler's derivation operators ∂^k satisfy the usual linearity properties. Finally, we notice that a sequence Θ^N that is differentiable up to order $(k+1)$, has the following integral representation

$$\Theta^N = \sum_{0\leq l\leq k} \frac{1}{N^l}\ \partial^l\Theta + \frac{1}{N^{k+1}}\ \partial^{k+1}\Theta^N$$

with $\sup_{N\geq 1} \|\partial^{k+1}\Theta^N\|_{\mathrm{tv}} < \infty$, and the convention $\partial^0\Theta = \Theta$, for $l = 0$.

1.5 Forest based expansion formulas

For transparency, we will often use terms associated with botanical, and genealogical trees. The forthcoming measures expansions will be therefore given in terms of signed measures indexed by collections of trees, colored forests and other analogous combinatorial and geometrical objects. More generally, most computations on particle measures will be rephrased in the course of the article in terms of enumerative problems on tree-like objects. In order to avoid lenghtening the present introduction, their definition and study are postponed to the next sections, and we present here our main results in a classical set-theoretic langage.

Let $\mathcal{A}_{n,q} =_{\text{def.}} ([q]^{[q]})^{n+1}$ be the set of $(n+1)$-sequences of mappings $\mathbf{a} = (a_p)_{0\leq p\leq n}$ from $[q]$ into itself.

$$[q] \xleftarrow{a_0} [q] \xleftarrow{a_1} \cdots \longleftarrow [q] \xleftarrow{a_{n-1}} [q] \xleftarrow{a_n} [q]$$

Notice that we write in bold the symbols for sequences (of maps, integers...) such as $\mathbf{a}$.

We let $\Delta_{n,q}$ be the nonnegative measure valued functional on $\mathcal{A}_{n,q}$ defined by

$$\Delta_{n,q} \ : \ \mathbf{a} \in \mathcal{A}_{n,q} \mapsto \Delta^{\mathbf{a}}_{n,q} = \left(\eta_0^{\otimes q} D_{a_0} Q_1^{\otimes q} D_{a_1} \ldots Q_n^{\otimes q} D_{a_n}\right) \in \mathcal{M}(E_n^q) \tag{1.13}$$

The Markov operator D_b can be seen as a coalescent, or a selection type transition. In this interpretation, the population $(x_n^{b(1)}, \ldots, x_n^{b(q)})$ results from a selection of the individuals with labels in the set $b([q])$. In addition, arguing as in (1.4), the non negative integral tensor product operators $Q_n^{\otimes q}$ can be seen as the overlapping of an absorption transition from $(E_{n-1} \cup \{c\})^q$ into itself, and an exploration transition from $(E_{n-1} \cup \{c\})^q$ into $(E_n \cup \{c\})^q$. We then have that

$$\Delta^{\mathbf{a}}_{n,q}(F) = \mathbb{E}\left(F(X_n^{\mathbf{a},c,1}, \ldots, X_n^{\mathbf{a},c,q})\ 1_{T^{\mathbf{a},1}\geq n}\ \ldots\ 1_{T^{\mathbf{a},q}\geq n}\right) \tag{1.14}$$

In the above display, $(X_n^{\mathbf{a},c,1}, \ldots, X_n^{\mathbf{a},c,q})$ represents the absorbed Markov chain on $(E_n \cup \{c\})^q$ with transitions $Q_n^{\otimes q} D_{a_n}$, and initial distribution $\eta_0^{\otimes q} D_{a_0}$.

Notice that the measures $\Delta^{\mathbf{a}}_{n,q}$ inherit a remarquable invariance property from their set-theoretic definition. Namely, let us introduce the natural left action of the group $\mathcal{G}_q^{n+2}$ on $\mathcal{A}_{n,q}$ defined for all $\mathbf{a} \in \mathcal{A}_{n,q}$ and all $\mathbf{s} = (s_0, ..., s_{n+1}) \in \mathcal{G}_q^{n+2}$ by

$$\mathbf{s}(\mathbf{a}) := (s_0 a_0 s_1^{-1}, s_1 a_1 s_2^{-1}, ..., s_n a_n s_{n+1}^{-1})$$

Then, for any $F \in \mathcal{B}_b^{sym}(E_n^q)$ we have

$$\Delta^{\mathbf{b}}_{n,q}(F) = \Delta^{\mathbf{s}(\mathbf{b})}_{n,q}(F) \tag{1.15}$$

The identity wouldn't hold if F wasn't a symmetric function. However, as already mentioned in Sect. 1.4, this is not a serious restriction as far as the determination of the q-tensor occupation measures is concerned. The action of $\mathcal{G}_q^{n+2}$ on $\mathcal{A}_{n,q}$ induces a partition of $\mathcal{A}_{n,q}$ into orbits or, equivalently, a partition into equivalence classes, where $\mathbf{a}$ and $\mathbf{b}$ are equivalent, $\mathbf{a} \approx \mathbf{b}$, if and only if there exists $\mathbf{s} \in \mathcal{G}_q^{n+2}$ such that $\mathbf{a} = \mathbf{s}(\mathbf{b})$. We write $\mathcal{F}_{n,q}$ for the set of orbits or equivalence classes. As we shall see later, $\mathcal{F}_{n,q}$ is nothing but (up to a canonical isomorphism) a remarquable subset of the classical set of forests. If $\mathbf{a} \in \mathcal{A}_{n,q}$, the corresponding equivalence class is written $\overline{\mathbf{a}}$. We also write $Stab(\mathbf{a})$ for the stabilizer of $\mathbf{a}$ in $\mathcal{G}_q^{n+2}$. According to the class formula, the number of elements in the equivalence class $\overline{\mathbf{a}}$, written $\#(\mathbf{a})$ or $\#(\overline{\mathbf{a}})$ is given by:

$$\#(\overline{\mathbf{a}}) = \frac{(q!)^{n+2}}{|Stab(\mathbf{a})|},$$

where we write $|Stab(\mathbf{a})|$ for the cardinal of $Stab(\mathbf{a})$ and, more generally, $|S|$ for the cardinal of any set S.

Explicit formulas for $\#(\overline{\mathbf{a}})$ and for the various quantities associated to the action of $\mathcal{G}_q^{n+2}$ on $\mathcal{A}_{n,q}$ such as the number of orbits (and much more) will be given later, and form, from the combinatorial point of view, one of the cores of the article. As it will appear, due to its general form, our analysis paves the way for a systematical combinatorial treatment of all mean field path dependent approximations -e.g. without restrictions on the number of particles at each step of the approximation process.

We are now nearly in a position to give our main results on Feynman-Kac particle models, but firstly, to simplify the presentation, need to introduce a pair of definitions. We start with some more or less traditional, and simplifying multi index notation. For any sequences of integers $\mathbf{p} = (p_k)_{0\leq k\leq n}$, and $\mathbf{l} = (l_k)_{0\leq k\leq n}$, we write $\mathbf{p} \leq \mathbf{l}$ if and only if $p_k \leq l_k$ for all $0 \leq k \leq n$. We write $|\mathbf{p}|$ for $(p_0 + ... + p_n)$. Assuming now that $\mathbf{p} \leq \mathbf{l}$, we use the multi-index notation

$$(\mathbf{l})_{\mathbf{p}} = \prod_{0\leq k\leq n} (l_k)_{p_k}\ , \quad \mathbf{p}! = \prod_{0\leq k\leq n} p_k!\ , \quad \text{and} \quad s(\mathbf{l}, \mathbf{p}) = \prod_{k=0}^{n} s(l_k, p_k)$$

where the $s(l_k, p_k)$ are Stirling numbers of the first kind. The difference $(\mathbf{p} - \mathbf{l})$, and respectively the addition $(\mathbf{p}+\mathbf{l})$ of two sequences is the sequence $(p_k - l_k)_{0\leq k\leq n}$, and respectively $(p_k + l_k)_{0\leq k\leq n}$. When no confusions can arise, we write $\mathbf{N}$, and $\mathbf{q}$, for the constant sequences $(N)_{0\leq i\leq n}$, and $(q)_{0\leq i\leq n}$. We also write $\mathbf{1}$, and respectively $\mathbf{0}$, for the sequence of unit integers, and respectively null integers. The above definitions are extended to infinite sequence of integers $\mathbf{p} = (p_k)_{k\geq 0} \in \mathbb{N}^{\mathbb{N}}$, with a finite numbers of strictly positive terms. Any function $\alpha : \mathbb{N} \mapsto \mathbb{N}$ on the set of integers into itself, is extended to integer sequences $\mathbf{p} \in \mathbb{N}^{\mathbb{N}}$, by setting $\alpha(\mathbf{p}) = (\alpha(p_k))_{k\geq 0}$.

For $\mathbf{a} \in \mathcal{A}_{n,q}$, we write $|\mathbf{a}|$ the sequence $(|a_i|)_{0\leq i\leq n}$. The coalescence sequence of $\mathbf{a}$ is the sequence $\mathbf{q}-|\mathbf{a}|$ of the coalescence numbers of the a_is. The coalescence degree of $\mathbf{a}$ is

$$coal(\mathbf{a}) := (|\mathbf{q}| - |\mathbf{a}|) = \sum_{i=0}^{n}(q - |a_i|)$$

The subset of $\mathcal{A}_{n,q}$ of sequences $\mathbf{a}$ such that $|\mathbf{a}| \geq \mathbf{q} - \mathbf{p}$ is written $\mathcal{A}_{n,q}(\mathbf{p})$. Notice that, if $\mathbf{a} \approx \mathbf{b}$, $|\mathbf{a}| = |\mathbf{b}|$, so that the notions of coalescence degree and coalescence sequence go over to the set $\mathcal{F}_{n,q}$ of forests. In particular, notation such as $|\overline{\mathbf{a}}|$ or $coal(\overline{\mathbf{a}})$ is well-defined. In view of fla 1.15, for any choice $(a_0, ..., a_n)$ of a representative of a forest $\mathbf{f}$ in $\mathcal{F}_{n,q}$, we set

$$\Delta^{\mathbf{f}}_{n,q} = \Delta^{\mathbf{a}}_{n,q} \quad \text{and} \quad |\mathbf{f}| = |\mathbf{a}|$$

The subset of $\mathcal{F}_{n,q}$ associated to $\mathcal{A}_{n,q}(\mathbf{p})$ is written $\mathcal{F}_{n,q}(\mathbf{p})$.

Theorem 1.2. *The non negative measure $\mathbb{Q}^N_{n,q}$ represents the distribution of the coalescent, and non absorbed Markov chain introduced in (1.14), associated with a sequence $\mathbf{A} = (A_k)_{0\leq k\leq n}$ of independent random mappings from $[q]$ into itself, with common distribution $\frac{1}{N^q}\sum_{a\in[q]^{[q]}}\frac{(N)_{|a|}}{(q)_{|a|}}\delta_a$. That is, for all $F \in \mathcal{B}^{sym}_b(E^q_n)$, we have that*

$$\mathbb{Q}^N_{n,q}(F) = \Delta^{\mathbf{A}}_{n,q}(F) = \mathbb{E}\left(F(X^{\mathbf{A},c,1}_n, \ldots, X^{\mathbf{A},c,q}_n)\ 1_{T^{\mathbf{A},1}\geq n}\ \ldots\ 1_{T^{\mathbf{A},q}\geq n}\right)$$

In addition, for any $1 \leq q \leq N$, we have the polynomial expansion

$$\mathbb{Q}^N_{n,q} = \gamma^{\otimes q}_n + \sum_{1\leq k\leq (q-1)(n+1)} \frac{1}{N^k}\ \partial^k\mathbb{Q}_{n,q}$$

with the collection of signed, and weak derivative measures $\partial^k\mathbb{Q}_{n,q}$ given by the formula

$$\partial^k\mathbb{Q}_{n,q} = \sum_{\mathbf{r}<\mathbf{q}:|\mathbf{r}|=k}\ \sum_{\mathbf{f}\in\mathcal{F}_{n,q}(\mathbf{r})} \frac{s(|\mathbf{f}|, \mathbf{q}-\mathbf{r})\ \#(\mathbf{f})}{(\mathbf{q})_{|\mathbf{f}|}}\ \Delta^{\mathbf{f}}_{n,q}$$

We shall present several consequences of this theorem, including explicit descriptions of the first two order terms in the polynomial expansion, and a new extension of the Wick product formula to forests.

The functional expansions of the distributions $\mathbb{P}^N_{n,q}$ rely on the analysis of tensor product particle measures on path spaces. These models can be studied following the same line of arguments as the ones we used for the distributions $\mathbb{Q}^N_{n,q}$. As a parenthesis, we observe that these two measures coincide for constant potential functions

$$(\forall n \geq 0 \quad \forall x_n \in E_n \quad G_n(x_n) = 1) \Longrightarrow \left(\forall n \geq 0 \quad \forall q \leq N \quad \mathbb{P}^N_{n,q} = \mathbb{Q}^N_{n,q}\right)$$

To deal with general potential functions, we need to re-normalize the particle distributions γ^N_n, and to extend the previous functional analysis to the path-space distribution of the particle measures defined for any integer sequence $\mathbf{l} = (l_k)_{0\leq k\leq n}$, by the following formula

$$\gamma^{\mathbf{l},N}_n = (\gamma^N_0)^{\otimes l_0} \otimes (\gamma^N_1)^{\otimes l_1} \otimes \cdots \otimes (\gamma^N_n)^{\otimes l_n}$$

The analytic functional structure of these path space models relies on natural combinatorial techniques on colored rooted trees, and forests. Our result basically read as follows.

Theorem 1.3. *For any q, the sequence of probability measures $(\mathbb{P}^N_{n+1,q})_{N\geq q}$ is differentiable up to any order. Besides, at any order $N \geq q$ we have the closed formula*

$$\mathbb{P}^N_{n+1,q} = \eta^{\otimes q}_{n+1} + \sum_{1\leq k<\lfloor (N-q)/2\rfloor} \frac{1}{N^k}\ \partial^k\mathbb{P}_{n+1,q} + \frac{1}{N^{\lfloor (N-q)/2\rfloor}}\ \partial^{\lfloor (N-q)/2\rfloor}\mathbb{P}^N_{n+1,q}$$

for some collection of signed derivative measures $\partial^k\mathbb{P}_{n,q}$, whose values can be explicitly described in terms of a class of colored forests, with maximal coalescent degree k.

As before, we shall derive several consequences of these expansions, including an explicit description of the first order term, and a Wick product formula on colored forests. The latter provides sharp estimates for $\mathbb{L}_q$-mean errors between the particle occupation measures η^N_n and the limiting Feynman-Kac measures η_n. Incidentally,

combining this Wick product formula with the Borel-Cantelli lemma we obtain the almost sure convergence result

$$\lim_{N\to\infty} (\eta_n^N)^{\odot q}(F_{n,q}) = 0$$

for any bounded symmetric function F on E_n^q, with $\int_{E_n} F(x_n^1,\ldots,x_n^q)\eta_n(dx_n^q) = 0$. This result is an extension of the law of large numbers for U-statistics obtained by W. Hoeffding [7] for independent and identically distributed random variables to interacting particle models.

2 Particle measures expansions on forests

2.1 A preliminary stochastic tensor product formula

The link between the two, usual and restricted, tensor products measures, $(\eta_n^N)^{\otimes q}$ and $(\eta_n^N)^{\odot q}$ relies in the end on a simple observation, that will appear to be fundamental for all our forthcoming computations. Throughout this section, integers $N \geq q \geq 1$ and a mesurable state space E are fixed once for all.

Consider first the surjection

$$\pi : [q]^{[q]} \times \langle q, N\rangle \longrightarrow [N]^{[q]}$$
$$(s,a) \longmapsto as := a \circ s$$

Lemma 2.1. *Let $b \in [N]^{[q]}$, then the cardinal $|\pi^{-1}(\{b\})|$ of $\pi^{-1}(\{b\})$ only depends on the cardinal $|b|$ of the image of b. It is given by*

$$|\pi^{-1}(\{b\})| = (N-|b|)_{q-|b|}(q)_{|b|}$$

Proof. Observe indeed that, with the same notations as above, for b and a fixed, the equation $b = as$ can have a solution only if the image of b is contained in the image of a. Let us consider such a a. Since a is injective, it has a left inverse a' for the composition of maps such that moreover $a'a$ acts as the identity on $[q]$. The equation has therefore a unique solution $s = a'b$. Since the number n_p of injections from $[q]$ to $[N]$ the image of which contains a fixed subset of $[N]$ of cardinal $p < q$ is

$$n_p = (q)_p(N-p)_{q-p},$$

the lemma follows. □

The lemma translates almost immediately into a functional relation between the two tensor product measures. Notice first the identity

$$D_a D_b = D_{ab}$$

that holds for any pair of mappings $(a,b) \in ([q]^{[q]})^2$. Let u, v be any linear combinations $\sum_{i\in I} \alpha_i a_i$, (resp. $\sum_{j\in J} \beta_i b_i$) of elements of $[q]^{[q]}$, or, in other terms, elements of the monoid algebra of $[q]^{[q]}$. We extend linearly the definition of D, so that:

$$D_u := \sum_{i\in I} \alpha_i D_{a_i}$$

and

$$D_u D_v = D_{uv} = \sum_{i\in I, j\in J} \alpha_i\beta_j D_{a_i b_j}$$

We denote now by $m(\mathbf{x})$ the empirical measure associated with an N-uple $\mathbf{x} = (x^i)_{1\leq i\leq N} \in E^N$

$$m(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^N \delta_{x^i}$$

For any integer $q \leq N$, we also consider the empirical measures on E^q defined by

$$m(\mathbf{x})^{\otimes q} = \frac{1}{N^q}\sum_{a\in[N]^{[q]}} \delta_{(x^{a(1)},\ldots,x^{a(q)})}$$
$$m(\mathbf{x})^{\odot q} = \frac{1}{(N)_q}\sum_{a\in\langle q,N\rangle} \delta_{(x^{a(1)},\ldots,x^{a(q)})}$$

Although the notation is self-explanatory, notice that, in the sequel, we will write simply $\mathbf{x}^a$ for $(x^{a(1)},\ldots,x^{a(q)})$ and $\delta_{\mathbf{x}^a}$ for $\delta_{(x^{a(1)},\ldots,x^{a(q)})}$.

Corollary 2.2. *We have*

$$m(\mathbf{x})^{\otimes q} = m(\mathbf{x})^{\odot q} D_{L_q^N}$$

where

$$L_q^N = \frac{1}{N^q} \sum_{a \in [q]^{[q]}} \frac{(N)_{|a|}}{(q)_{|a|}} \; a$$

Proof. In view of the following identity, that holds for any function $F \in \mathcal{B}_b^{sym}(E^q)$, with $q \leq N$,

$$\delta_{\mathbf{x}^a} D_b(F) = D_b(F)(x^{a(1)}, ..., x^{a(n)}) = F(x^{ab(1)}, ..., x^{ab(n)}) = \delta_{\mathbf{x}^{ab}}(F)$$

the corollary follows from lemma 2.1 and the identity

$$\frac{1}{(N)_q}(N)_p(N-p)_{q-p} = 1 \quad \text{where} \quad N \geq q \geq p$$

□

Corollary 2.3. *The tensor product measure $m(\mathbf{x})^{\otimes q}$ has a Laurent expansion:*

$$m(\mathbf{x})^{\otimes q} = m(\mathbf{x})^{\odot q} \cdot \sum_{0 \leq k < q} \frac{1}{N^k} D_{\partial^k L_q}$$

where

$$L_q^N = \sum_{0 \leq k < q} \frac{1}{N^k} \; \partial^k L_q$$

and

$$\partial^k L_q = \sum_{q-k \leq p \leq q} s(p, q-k) \; \frac{1}{(q)_p} \sum_{a \in [q]_p^{[q]}} \; a$$

Notice that, in the expansion of $\partial^k L_q$, the double sum involves only elements $a \in [q]^{[q]}$ the coalescence number of which is less or equal to k.

The corollary follows from corollary 2.2 and from the Stirling formula

$$(N)_p = \sum_{1 \leq k \leq p} s(p,k) N^k$$

For $q > 1$, and recalling that $s(q, q-1) = -\begin{pmatrix} q \\ 2 \end{pmatrix}$, we get

$$\begin{aligned}
\partial^0 L_q &= \frac{1}{q!} \sum_{a \in \langle q,q \rangle} \; a \\
\partial^1 L_q &= -\frac{q(q-1)}{2} \; \frac{1}{q!} \sum_{a \in \langle q,q \rangle} \; a + \frac{1}{q!} \sum_{a \in [q]_{q-1}^{[q]}} \; a \\
\partial^2 L_q &= s(q,q-2) \; \frac{1}{q!} \sum_{a \in \langle q,q \rangle} \; a - \frac{(q-1)(q-2)}{2} \; \frac{1}{q!} \sum_{a \in [q]_{q-1}^{[q]}} \; a + \frac{1}{(q)_{q-2}} \sum_{a \in [q]_{q-2}^{[q]}} \; a
\end{aligned}$$

Notice in particular that, since the composition map

$$\langle q, q \rangle \times \langle q, N \rangle \longrightarrow \langle q, N \rangle$$

is a surjection with all fibers of cardinal $q!$, we have

$$m(\mathbf{x})^{\odot q} D_{\partial^0 L_q} = m(\mathbf{x})^{\odot q}$$

so that, asymptotically in N

$$m(\mathbf{x})^{\otimes q} \sim m(\mathbf{x})^{\odot q}$$

More precisely, we have, assuming from now on that $\mathbf{x}$ is generic (that is, $x_i \neq x_j$ if $i \neq j$):

Corollary 2.4. *The following formulas hold for the expansion of* $m(\mathbf{x})^{\otimes q}$

$$N\ ||m(\mathbf{x})^{\otimes q} - m(\mathbf{x})^{\odot q}||_{tv} = 2\ \frac{N^q - (N)_q}{N^{q-1}} \underset{N\to+\infty}{\longrightarrow} q(q-1)$$

$$||m(\mathbf{x})^{\odot q} D_{\partial^k L_q}||_{tv} = \sum_{p=q-k}^{q} |s(p,q-k)|\ S(q,p)$$

Here, $||\ \ ||_{tv}$ stands fort the total variation, that is, for any linear operator Q on $\mathcal{B}_b(E^q)$,

$$||Q||_{tv} := \sup_{||f||_\infty=1} |Q(f)|$$

Recall that the Stirling number of the second kind $S(q,p)$ is the number of partitions of $[q]$ into p non-empty subsets, so that

$$|[q]_p^{[q]}| = S(q,p)(q)_p$$

$$N^q = \sum_{1\le k\le q} S(q,k)(N)_k$$

The corollary follows then from our previous computations and from the following lemma.

Lemma 2.5. *Let* $Q = \sum_{0\le p\le q} u_p \sum_{a\in[N]_p^{[q]}} \delta_{\mathbf{x}^a}$*, where the* u_p *are arbitrary real coefficients, then we have*

$$||Q||_{tv} = \sum_{0\le p\le q} |u_p|\ (N)_p\ S(q,p)$$

Proof. Indeed, we have $||Q||_{tv} \le \sum_{0\le p\le q} |u_p|\ (N)_p\ S(q,p)$ by direct inspection, since $||\delta_{\mathbf{x}^a}||_{tv} = 1$ for all $a \in [q]^{[q]}$. Let us write $sgn(k)$ for the sign of u_k, and let us introduce the function $\phi \in \mathcal{B}_b(E^q)$ defined by:

$$\phi(y_1,...,y_q) = \sum_{0\le k\le q} sgn(k)\delta^k_{\lambda(y_1,...,y_q)}$$

where $\lambda(y_1,...,y_q) := |\{y_1,\dots,y_q\}|$ and $\delta^k_{\lambda(y_1,...,y_q)}$ stands for the Dirac function. Then, $Q(\phi) = \sum_{0\le p\le q} |u_p|\ (N)_p\ S(q,p)$, and the lemma follows. □

2.2 Laurent expansions

The goal of this section is to derive a Laurent expansion of the measures $\mathbb{Q}^N_{n,q} \in \mathcal{M}(E_n^q)$, $\mathbb{Q}^N_{n,q}(F) := \mathbb{E}((\gamma_n^N)^{\otimes q}(F))$, $F \in \mathcal{B}_b(E_n^q)$, with respect to the population size parameter N. The following proposition is fundamental. Recall from Sect. 1.5 that we write $\mathcal{F}_{n,q}$ for the set of equivalence classes in $\mathcal{A}_{n,q}$ under the action of the permutation group $\mathcal{G}_q^{n+2}$.

Proposition 2.6. *For any integers* $q \le N$*, any time parameter* $n \in \mathbb{N}$ *and any* $F \in \mathcal{B}_b^{sym}(E_n^q)$ *we have the Laurent expansion*

$$\mathbb{Q}^N_{n,q}(F) = \frac{1}{N^{q(n+1)}} \sum_{\mathbf{f}\in\mathcal{F}_{n,q}} \frac{(\mathbf{N})_{|\mathbf{f}|}}{(\mathbf{q})_{|\mathbf{f}|}}\ \#(\mathbf{f})\ \Delta^{\mathbf{f}}_{n,q}(F)$$

with the mappings $\Delta_{n,q}$ *introduced in (1.13).*

Proof. Indeed, combining the definition of the particle model with corollary 2.2, we first find that

$$\begin{aligned}
\mathbb{E}\left(\left(\gamma_n^N\right)^{\otimes q}(F)\ \middle|\ \xi_{n-1}^{(N)}\right) &= \left(\gamma_n^N(1)\right)^q \times \mathbb{E}\left(\left(\eta_n^N\right)^{\otimes q}(F)\ \middle|\ \xi_{n-1}^{(N)}\right) \\
&= \left(\gamma_n^N(1)\right)^q \times \mathbb{E}\left(\left(\eta_n^N\right)^{\odot q} D_{L_q^N}(F)\ \middle|\ \xi_{n-1}^{(N)}\right) \\
&= \left(\gamma_n^N(1)\right)^q \times \mathbb{E}\left(\Phi_n(\eta_{n-1}^N)^{\otimes q} D_{L_q^N}(F)\ \middle|\ \xi_{n-1}^{(N)}\right)
\end{aligned}$$

Using the fact that, conditionnaly to $\xi_{n-1}^{(N)}$ and for $G \in \mathcal{B}_b^{sym}(E_n^q)$

$$\Phi_n(\eta_{n-1}^N)^{\otimes q}(G) = \frac{\left(\eta_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}G)}{\left(\eta_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}(1))} = \frac{\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}G)}{\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}(1))}$$

and

$$\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}(1)) = \left(\gamma_{n-1}^N Q_n(1)\right)^q = \left(\gamma_{n-1}^N(G_{n-1})\right)^q = \left(\gamma_n^N(1)\right)^q$$

we arrive at

$$\begin{aligned}\mathbb{E}\left(\left(\gamma_n^N\right)^{\otimes q}(F) \;\middle|\; \xi_{n-1}^{(N)}\right) &= \left(\gamma_n^N(1)\right)^q \times \mathbb{E}\left(\frac{\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}D_{L_q^N}F)}{\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}(1))} \;\middle|\; \xi_{n-1}^{(N)}\right) \\ &= \mathbb{E}\left(\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}D_{L_q^N}F) \;\middle|\; \xi_{n-1}^{(N)}\right) \qquad (2.1)\end{aligned}$$

Integrating over the past, this yields that

$$\mathbb{E}\left(\left(\gamma_n^N\right)^{\otimes q}(F)\right) = \mathbb{E}\left(\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q}D_{L_q^N}F)\right)$$

Using a simple induction, we readily obtain the formulae

$$\begin{aligned}\mathbb{E}\left(\left(\gamma_n^N\right)^{\otimes q}(F)\right) &= \mathbb{E}\left(\left((\gamma_0^N)^{\otimes q}Q_1^{\otimes q}D_{L_q^N}\ldots Q_n^{\otimes q}D_{L_q^N}\right)(F)\right) \\ &= \mathbb{E}\left(\left(\eta_0^{\otimes q}D_{L_q^N}Q_1^{\otimes q}D_{L_q^N}\ldots Q_n^{\otimes q}D_{L_q^N}\right)(F)\right) \\ &= \frac{1}{N^{q(n+1)}}\sum_{\mathbf{a}\in\mathcal{A}_{n,q}}\frac{(\mathbf{N})_{|\mathbf{a}|}}{(\mathbf{q})_{|\mathbf{a}|}}\,\Delta_{n,q}^{\mathbf{a}}(F)\end{aligned}$$

where $\mathcal{A}_{n,q}$ is the set of sequences of maps introduced in section 1.5 and where we use the multi-index notation introduced in the same section. In view of equation (1.15), we know that

$$\Delta_{n,q}^{\mathbf{b}}(F) = \Delta_{n,q}^{\mathbf{s(b)}}(F)$$

whenever $\mathbf{a}$ and $\mathbf{b}$ belong to the same equivalence class in $\mathcal{A}_{n,q}$, and the Proposition follows. □

We are now in position to derive Theorem 1.2

Theorem 2.7. *For any time horizon $n \in \mathbb{N}$, any population size $N \geq 1$, and any integer $1 \leq q \leq N$, we have the Laurent expansion*

$$\mathbb{Q}_{n,q}^N(F) = \gamma_n^{\otimes q} + \sum_{1\leq k\leq (q-1)(n+1)}\frac{1}{N^k}\,\partial^k\mathbb{Q}_{n,q}(F) \qquad (2.2)$$

with the signed measures $\partial^k\mathbb{Q}_{n,q}$ such that

$$\partial^k\mathbb{Q}_{n,q}(F) = \sum_{\mathbf{r}<\mathbf{q},\;|\mathbf{r}|=k}\;\sum_{\mathbf{f}\in\mathcal{F}_{n,q}(\mathbf{r})}\frac{s(|\mathbf{f}|,\mathbf{q}-\mathbf{r})\;\#(\mathbf{f})}{(\mathbf{q})_{|\mathbf{f}|}}\;\Delta_{n,q}^{\mathbf{f}}(F)$$

for all $F \in \mathcal{B}_b^{sym}(E_n^q)$.

Proof. Indeed, according to Proposition 2.6

$$\mathbb{Q}_{n,q}^N(F) = \frac{1}{N^{q(n+1)}}\sum_{\mathbf{p}\in[1,q]^{n+1}}\frac{(\mathbf{N})_{\mathbf{p}}}{(\mathbf{q})_{\mathbf{p}}}\sum_{\mathbf{f}\in\mathcal{F}_{n,q},|\mathbf{f}|=\mathbf{p}}\#(\mathbf{f})\;\Delta_{n,q}^{\mathbf{f}}(F)$$

Using the Stirling formula

$$(\mathbf{N})_{\mathbf{p}} = \sum_{\mathbf{l}\leq\mathbf{p}} s(\mathbf{p},\mathbf{l})N^{|\mathbf{l}|} \qquad (2.3)$$

we also have that

$$\mathbb{Q}^N_{n,p}(F) = \sum_{\mathbf{l}\leq\mathbf{q}} \sum_{\mathbf{l}\leq\mathbf{p}\leq\mathbf{q}} s(\mathbf{p},\mathbf{l})\ \frac{1}{N^{|\mathbf{q}-\mathbf{l}|}} \frac{1}{(\mathbf{q})_{\mathbf{p}}} \sum_{\mathbf{f}\in\mathcal{F}_{n,q},|\mathbf{f}|=\mathbf{p}} \#(\mathbf{f})\ \Delta^{\mathbf{f}}_{n,q}(F)$$

from which we conclude that

$$\mathbb{Q}^N_{n,q}(F) = \sum_{\mathbf{r}<\mathbf{q}} \sum_{\mathbf{q}-\mathbf{r}\leq\mathbf{p}\leq\mathbf{q}} s(\mathbf{p},\mathbf{q}-\mathbf{r})\ \frac{1}{N^{|\mathbf{r}|}} \frac{1}{(\mathbf{q})_{\mathbf{p}}} \sum_{\mathbf{f}\in\mathcal{F}_{n,q}|\mathbf{f}|=\mathbf{p}} \#(\mathbf{f})\ \Delta^{\mathbf{f}}_{n,q}(F)$$

Observe then that $\#(\mathbf{f}) = (q!)^{n+1}$ if $|\mathbf{f}| = \mathbf{q}$ since, in that case, $\mathbf{f}$ is nothing but the equivalence class of all sequences of bijections in $\mathcal{A}_{n,q}$. This ends the proof of the theorem. □

Corollary 2.8. *Assuming that $||\Delta^{\mathbf{f}}_{n,q}||_{tv} \leq 1$ for any $\mathbf{f} \in \mathcal{F}_{n,q}$, the following formula holds for the expansion of $\mathbb{Q}^N_{n,q}$*

$$\lim_{N\longrightarrow\infty} N\ ||\mathbb{Q}^N_{n,q} - \gamma_n^{\otimes q}||_{tv} \leq (n+1)q(q-1)$$

Proof. We have (see proof of theorem 2.7) that $\#(\mathbf{f}) = (q!)^{n+1}$ if $|\mathbf{f}| = \mathbf{q}$ and besides, $|\{\mathbf{r},\ |\mathbf{r}| = 1\}| = n+1$, and for any such $\mathbf{r}$, $s(\mathbf{q},\mathbf{q}-\mathbf{r}) = -\binom{q}{2}$. In conclusion, the case $|\mathbf{f}| = \mathbf{q}$ contributes to $(n+1)q(q-1)/2$ to the asymptotic evaluation of $N\ ||\mathbb{Q}^N_{n,q} - \gamma_n^{\otimes q}||_{tv}$.

Now, let us consider a sequence $\mathbf{r}$ with $|\mathbf{r}| = 1$, e.g. the one, written $\mathbf{r}_i$ with $r_i = 1$, for a given $0 \leq i \leq n$. In that case, there is a unique $\mathbf{f}$ with $|\mathbf{f}| = \mathbf{q}-\mathbf{r}$, which is the equivalence class of all sequences of maps $\mathbf{a}$ in $\mathcal{A}_{n,q}$ with $a_j \in \mathcal{G}_q,\ j \neq i$ and $|a_i| = q-1$. In particular $\#(\mathbf{f}) = (q!)^{n+1}\binom{q}{2}$, so that, on the whole, the case $|\mathbf{f}| = \mathbf{r}_i$ contributes to $q(q-1)/2$ to the evaluation.

The corollary follows then from the Laurent expansion of Theorem 2.7. □

We conclude this section with noting that the above functional expansions also apply to the dot-tensor product measures

$$\left(\gamma_n^N\right)^{\odot q}(F) = \left(\gamma_n^N(1)\right)^q \times \left(\eta_n^N\right)^{\odot q}(F)$$

More precisely, by definition of the particle model we have that

$$\mathbb{E}\left(\left(\gamma_n^N\right)^{\odot q}(F)\right) = \mathbb{E}\left(\left(\gamma_{n-1}^N\right)^{\otimes q} Q_n^{\otimes q}(F)\right) = \mathbb{Q}^N_{n-1,q}(Q_n^{\otimes q}F) \tag{2.4}$$

3 Combinatorial methods for counting forests and jungles

In the present section, we face the problem of computing the cardinals $\#(\mathbf{f})$ involved in our Laurent expansions. We also derive various identities relevant for the fine asymptotical analysis of mean field particle interacting models, such as the number of classes of elements in $\mathcal{F}_{n,q}$ with a given coalescence degree.

As already pointed out in the Introduction, we have choosen the langage of botanical and genealogical trees, both for technical reasons (since forests appear to be the most natural parametrization of elements in $\mathcal{F}_{n,q}$), and also for the clarity of the exposition.

3.1 Some terminology on trees and forests

In this section, we detail the vocabulary of trees, that will be of constant use later in the article.

Definition 3.1. *A tree (respectively a planar tree) is a (isomorphism class of) finite non-empty oriented connected (and respectively planar) graph T without loops such that any vertex of T has at most one outgoing edge. Paths are oriented from the vertices to the root. There is no restriction on the number of incoming edges. The empty graph $\emptyset$ is viewed as a tree and is called the empty tree. The set of all trees (resp. planar trees) is denoted by $\mathcal{T}$ (resp. $\mathcal{PT}$). The vertices of a tree without incoming edges are called the leaves; the vertices with both incoming edges and an outgoing edge are called the internal vertices or the nodes; the (necessarily unique) vertex without outgoing edge is called the root.*

Definition 3.2. *A forest* $\mathbf{f}$ *is a multiset of trees, that is a set of trees, with repetitions of the same tree allowed or, equivalently, an element of the commutative monoid* $\langle \mathcal{T} \rangle$ *on* $\mathcal{T}$*, with the empty graph* $T_0 = \emptyset$ *as a unit. Since the algebraic notation is the most convenient, we write*

$$\mathbf{f} = T_1^{m_1} ... T_k^{m_k}$$

for the forest with the tree T_i *appearing with multiplicity* m_i*,* $i \leq k$*. When* $T_i \neq T_j$ *for* $i \neq j$*, we say that* $\mathbf{f}$ *is written in normal form. A planar forest* $\mathbf{f}'$ *is an ordered sequence of planar trees. Planar forests can be represented by noncommutative monomials (or words) on the set of planar trees. The sets of forests and planar forests are written* $\mathcal{F}$ *and* $\mathcal{PF}$*.*

For example, if T_1' and T_2' are two planar trees, $\mathbf{f}' = T_1' T_2' T_1'$ is the planar forest obtained by left-to-right concatenation of T_1', T_2' and another copy of T_1'.

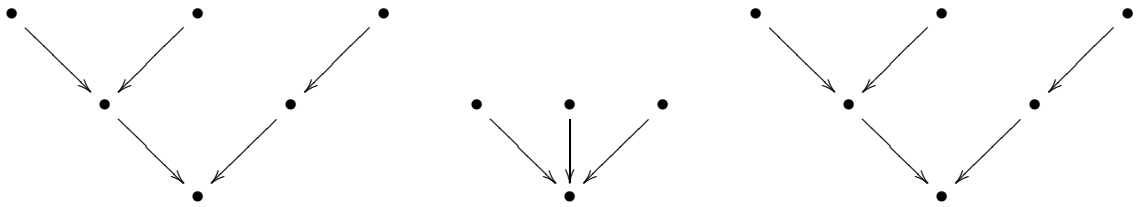

Figure 1: a graphical representation of a planar forest $\mathbf{f}' = T_1' T_2' T_1'$.

We say that a tree T is a subtree of T', and we write $T \subset T'$, if the graph of T is a subgraph of the graph of T', and if the root of T is also the root of T'. In a more pedantic way, a subtree of a tree is a connected subgraph of T' containing the root. A subforest $\mathbf{f} \subset \mathbf{f}'$ of a forest $\mathbf{f}'$ is defined accordingly, as a collection of pairwise disjoint subtrees of the trees in $\mathbf{f}'$.

The distance between two vertices in a tree is the minimal number of edges of a path joigning them. The height of a vertex is its distance to the root. We also say that a vertex with height k is a vertex at level k in the tree. The height $ht(T)$ of a tree T is the maximal distance between a leaf and the root. The height $ht(\mathbf{f})$ of a forest $\mathbf{f}$ is the maximal height $ht(T)$ of the trees $T \subset \mathbf{f}$ in the forest

$$ht(\mathbf{f}) = \sup_{T \subset \mathbf{f}} ht(T)$$

For the unit empty tree, we also take the convention $ht(\emptyset) = -1$.

We write $v_k(\mathbf{f})$ for the number of vertices in a forest $\mathbf{f}$ at level $k \geq 0$. The number of vertices at each level in a forest is encoded in the mapping :

$$v \ : \ \mathbf{f} \in \mathcal{F} \mapsto v(\mathbf{f}) = (v_k(\mathbf{f}))_{k \geq 0}$$

Notice that for any pair of forests $(\mathbf{f}, \mathbf{g}) \in \mathcal{F}$, we have that

$$v(\mathbf{fg}) = v(\mathbf{f}) + v(\mathbf{g})$$

We write $\mathbf{V}$ for $v(\mathcal{F})$, which coincides with the set of connected initial integer sequences, that is integer sequences $\mathbf{p} = (p_k)_{k \geq 0} \in \mathbb{N}^{\mathbb{N}}$ satisfying the following property

$$\exists ht(\mathbf{p}) \in \mathbb{N} \quad \text{s.t.} \quad \inf_{k \leq ht(\mathbf{p})} p_k > 0 \quad \text{and} \quad \sup_{k > ht(\mathbf{p})} p_k = 0$$

For the unit empty tree $T_0 = \emptyset$, we use the conventions

$$v(\emptyset) = \mathbf{0} = (0, 0, \ldots) \quad \text{and} \quad ht(\mathbf{0}) = -1$$

For any $\mathbf{p} \in \mathbf{V}$, we write $\mathcal{F}_{\mathbf{p}}$ for the set of forests $\mathbf{f}$ such that $v(\mathbf{f}) = \mathbf{p}$. Our notation is consistent, since the height $ht(\mathbf{f})$ of a forest $\mathbf{f} \in \mathcal{F}_{\mathbf{p}}$ clearly coincides with the height $ht(\mathbf{p})$ of the integer sequence $\mathbf{p}$. For the null sequence $\mathbf{0}$, we have that $\mathcal{F}_{\mathbf{0}} = \{\emptyset\}$.

When $\mathbf{p} \in \mathbf{V}$ is chosen so that $p_0 = 1$, the set $\mathcal{F}_{\mathbf{p}}$ reduces to the set $\mathcal{T}_{\mathbf{p}}$ of all trees T such that $v(T) = \mathbf{p}$. The notation on trees, and forests goes over to planar trees, and planar forests in a self-explanatory way. For example, for any sequence $\mathbf{p} \in \mathbf{V}$, $\mathcal{PF}_{\mathbf{p}}$ is the set of planar forests $\mathbf{f}$, with $v_k(\mathbf{f}) = p_k$ vertices at level $k \geq 0$.

For any $n \in (\mathbb{N} \cup \{-1\})$, we denote by $\mathbf{V}_n \subset \mathbf{V}$ the subset of sequences $\mathbf{p}$ such that $ht(\mathbf{p}) = n$. Notice that $\mathbf{V}_{-1} = \{\mathbf{0}\}$. In this notation, the sets of all forests, and trees with height n are given by the sets

$$\mathcal{F}(n) := \coprod_{\mathbf{p} \in V_n} \mathcal{F}_p, \quad \mathcal{T}(n) := \coprod_{\mathbf{p} \in V_n, p_0 = 1} \mathcal{T}_p$$

The shift operator

$$B \ : \ \mathbf{p} = (p_k)_{k \geq 0} \in \mathbf{V} \mapsto B(\mathbf{p}) = (q_l)_{l \geq 0} \in \mathbf{V}, \ q_l := p_{l+1}$$

induces a canonical bijection, still denoted by B, between the set of trees $\mathcal{T}_{\mathbf{p}}$ (resp. planar trees $\mathcal{PT}_{\mathbf{p}}$), with $\mathbf{p} \in \mathbf{V}_{n+1}$ and the set of forests $\mathcal{F}_{B(\mathbf{p})}$ (resp. planar forests $\mathcal{PF}_{B(\mathbf{p})}$), obtained by removing the root of the tree

$$B \ : \ T \in \mathcal{T}_{\mathbf{p}} \ \text{ (resp. } \mathcal{PT}_{\mathbf{p}}) \mapsto B(T) \in \mathcal{F}_{B(\mathbf{p})} \ \left(\text{resp. } \mathcal{PF}_{B(\mathbf{p})}\right)$$

From the graph-theoretic point of view, the number of coalescences $c_k(\mathbf{f})$ at each level $k \geq 0$ in a forest $\mathbf{f} \in \mathcal{F}$ is defined by the mapping

$$c \ : \ \mathbf{f} \in \mathcal{F} \mapsto c(\mathbf{f}) = B(v(\mathbf{f})) - |\mathbf{f}|$$

The sequence $c(\mathbf{f})$ is called the coalescence sequence of $\mathbf{f}$. In the above displayed formula, the sequence $|\mathbf{f}| = (|\mathbf{f}|_k)_{k \geq 0}$ represents the number of vertices at level $k \geq 0$ minus the number of leaves at the same level, that is the number of vertices with an ingoind edge, whereas $B(v(\mathbf{f}))_k$ stands, by definition, for the number of edges between the levels $k+1$ and k. Finally observe that for any pair of forests $(\mathbf{f}, \mathbf{g}) \in \mathcal{F}$, we have

$$(\ |\mathbf{fg}| = |\mathbf{f}| + |\mathbf{g}| \quad \text{and} \quad B(v(\mathbf{f}) + v(\mathbf{g})) = B(v(\mathbf{f})) + B(v(\mathbf{g}))\ ) \Rightarrow c(\mathbf{fg}) = c(\mathbf{f}) + c(\mathbf{g})$$

We also say that the coalescence order of a vertex in a tree or a forest is its number of incoming edges minus 1. The coalescence degree of a tree or a forest $\mathbf{f}$ is the sum of the $c_k(\mathbf{f})$ or, equivalently, the sum of the coalescence orders of its vertices. We say that a tree is *trivial* if its coalescence degree is 0.

3.2 On the algebraic structure of trees and forests

In a planar forest $\mathbf{f}'$, vertices at the same level $k \geq 0$, are naturally ordered from left to right, and therefore in bijection with $[v_k(\mathbf{f}')]$. Planar forests (resp. planar trees) $\mathbf{f}' \in \mathcal{PF}$ of height $ht(\mathbf{f}') = (n+1)$ are therefore canonically in bijection with sequences of maps $\mathbf{a} = (a_0, ..., a_n)$, where a_k is a weakly increasing map from $[p_{k+1}]$ to $[p_k]$, with $k = 0, ..., n$, and $\mathbf{p} \in \mathbf{V}_{n+1}$ (resp. and $p_0 = 1$).

To clarify the presentation, we write $a \uparrow$ to indicate that a given mapping a between sets of integers is weakly increasing. In this notation, for any $\mathbf{p} \in \mathbf{V}_{n+1}$ we have that

$$\mathcal{PF}_{\mathbf{p}} \simeq \mathcal{IA}_{\mathbf{p}} := \left\{(a_0, ..., a_n) \in \left([p_0]^{[p_1]} \times \ldots \times [p_n]^{[p_{n+1}]}\right) \ : \ \forall 0 \leq k \leq n \ \ a_k \uparrow\right\}$$

We write $PF(\mathbf{a})$ (resp. $PT(\mathbf{a})$) for the planar forest associated to a sequence in $\mathcal{IA}_{\mathbf{p}}$ (resp. with $p_0 = 1$). In this interpretation, when $\mathbf{f}' \in \mathcal{PF}_{\mathbf{p}}$ is represented by a sequence of mappings $\mathbf{a} = (a_0, ..., a_n) \in \mathcal{IA}_{\mathbf{p}}$, the number of coalescent vertices at level k is given by

$$c_k(\mathbf{f}') = v_{k+1}(\mathbf{f}') - |a_k|$$

Notice that the definition is coherent with the definition of the coalescence sequence associated to $\mathbf{a}$. In addition, when the integer sequence $\mathbf{p} \in \mathbf{V}_{n+1}$ is chosen such that $p_0 = 1$, the mapping B introduced in the previous Section can be rewritten as follows:

$$B \ : \ PT(\mathbf{a}) \in \mathcal{PT}_{\mathbf{p}} \mapsto B(PT(\mathbf{a})) := PF(B(\mathbf{a})) \in \mathcal{PF}_{B(\mathbf{p})}$$

with the shift operator $B(a_0, \ldots, a_n) := (a_1, \ldots, a_n)$. These definitions and notations can be extended to arbitrary sequences of maps and arbitrary trees and forests as follows.

Definition 3.3. *Let $\mathbf{a} = (a_0, a_1, ..., a_n)$ be any sequence of maps $a_k : [p_{k+1}] \to [p_k]$, with $k = 0, \ldots, n$. To a sequence $\mathbf{a}$ is naturally associated a forest $F(\mathbf{a})$: the one with one vertex for each element of $\coprod_{k=0}^{n+1}[p_k]$, and a edge for each pair $(i, a_k(i)), i \in [p_{k+1}]$. The sequence can also be represented graphically uniquely by a planar graph $J(\mathbf{a})$, where however the edges between vertices at level $k+1$ and k are allowed to cross. We call such a planar graph, where paths between vertices are entangled, a jungle. The set of jungles is written $\mathcal{J}$.*

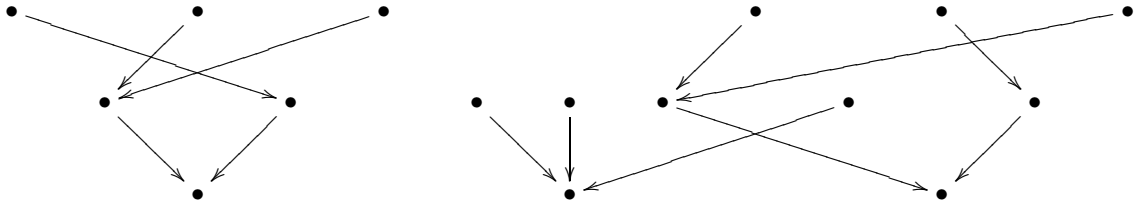

Figure 2: The entangled graph representation of a jungle with the same underlying graph as the planar forest in Fig. 1.

Notice that planar forests are particular examples of jungles, so that, for any sequence of weakly increasing maps $\mathbf{a}$ we have $PF(\mathbf{a}) = J(\mathbf{a})$. Notation such as $\mathcal{F}_{\mathbf{p}}$, with $\mathbf{p} \in \mathbf{V}_{n+1}$, extends from forests to jungles in a self-explanatory way. Jungles $j \in \mathcal{J}_{\mathbf{p}}$ of height $(n+1)$ are therefore canonically in bijection with sequences of maps $(a_0, ..., a_n)$, where a_k is a map from $[p_{k+1}]$ to $[p_k]$, with $k = 0, ..., n$. More formally, we have that

$$\mathcal{J}_{\mathbf{p}} \simeq \mathcal{A}_{\mathbf{p}} := \left([p_0]^{[p_1]} \times \ldots \times [p_n]^{[p_{n+1}]} \right)$$

To make these algebraic representations of planar forests and jungles more precise, it is convenient to introduce another round of notation. For any integers $p, q, r, s \geq 1$, and any pair of maps $(\alpha, \beta) \in \left([q]^{[p]} \times [s]^{[r]} \right)$, we let $(\alpha \vee \beta)$ be the weakly increasing map from $[p+r]$ into $[q+s]$ defined by

$$(\alpha \vee \beta)(i) = \begin{cases} \alpha(i) & \text{if} \quad i \in [p] \\ q + \beta(i) & \text{if} \quad i = p + j \text{ with } j \in [r] \end{cases}$$

If α and β are weakly increasing, so is $\alpha \vee \beta$. For null integers $p = 0$, or $p = 0 = q$ we use the conventions $[q]^{[p]} = \{\emptyset\}$, and the composition rules $(\emptyset \vee \emptyset) = \emptyset$, $(\alpha \vee \emptyset) = \alpha = (\emptyset \vee \alpha)$. Observe that for any $(\sigma, \tau) \in (\mathcal{G}_q \times \mathcal{G}_s)$, any $(\alpha, \beta) \in \left([q]^{[p]} \times [s]^{[r]} \right)$ and any $(\rho, \theta) \in (\mathcal{G}_p \times \mathcal{G}_r)$ we have the composition formula

$$(\sigma \vee \tau) \circ (\alpha \vee \beta) \circ (\rho \vee \theta) \quad = \quad (\sigma \circ \alpha \circ \rho) \vee (\tau \circ \beta \circ \theta) \tag{3.1}$$

For any $\rho \in \mathcal{G}_{p+r}$ with α and β as above, we also quote the following factorization rule

$$(\alpha \vee \beta) \circ \rho \in \left([q]^{[p]} \vee [s]^{[r]} \right) \Longrightarrow \rho \in (\mathcal{G}_p \vee \mathcal{G}_r)$$

To check this final assertion, we suppose there exists some $i \in [p]$ such that $\rho(i) = p + j$, for some $j \in [r]$. In this case, we would find the following contradiction

$$(\alpha \vee \beta)(\rho(i)) = (\alpha \vee \beta)(p + j) = q + \beta(j) \notin [q]$$

By a simple induction argument, for any $\alpha_i \in [q_i]^{[p_i]}, i = 1, ..., m$, we also have that

$$(\alpha_1 \vee \ldots \vee \alpha_m) \circ \rho \in \left([q_1]^{[p_1]} \vee \ldots \vee [q_m]^{[p_m]} \right) \Longrightarrow \rho \in (\mathcal{G}_{p_1} \vee \ldots \vee \mathcal{G}_{p_m}) \tag{3.2}$$

for any permutation $\rho \in \mathcal{G}_{p_1 + \ldots + p_m}$, and $m \geq 1$.

More generally, for any sequences of integers $\mathbf{p} \in \mathbf{V}$, and $\mathbf{r} \in \mathbf{V}$, and any sequences of maps $\alpha = (\alpha_k)_{k \geq 0} \in \mathcal{A}_{\mathbf{p}}$, and $\beta = (\beta_k)_{k \geq 0} \in \mathcal{A}_{\mathbf{r}}$ we denote by $(\alpha \vee \beta)$ the sequence of maps

$$\alpha \vee \beta := (\alpha_k \vee \beta_k)_{k \geq 0} \in \mathcal{A}_{\mathbf{p}+\mathbf{r}}$$

Notice that, if $\alpha \in \mathcal{IA}_{\mathbf{p}}$ and $\beta \in \mathcal{IA}_{\mathbf{q}}$, then $\alpha \vee \beta \in \mathcal{IA}_{\mathbf{p}+\mathbf{q}}$. In the above displayed formula, we have used the conventions $\alpha_k = \emptyset$, and $\beta_l = \emptyset$, for any $k > ht(\mathbf{p})$, and $l > ht(\mathbf{r})$.

Recall that sequences of weakly increasing maps $\alpha \in \mathcal{IA}_{\mathbf{p}}$, and $\beta \in \mathcal{IA}_{\mathbf{r}}$ can be represented graphically by the planar forests $PF(\alpha) \in \mathcal{PF}_{\mathbf{p}}$, and $PF(\beta) \in \mathcal{PF}_{\mathbf{r}}$. In this interpretation, the composition formula on the noncommutative monoid $\mathcal{PF}$ is now given by

$$PF(\alpha)PF(\beta) = PF(\alpha \vee \beta) \in \mathcal{PF}_{\mathbf{p}+\mathbf{r}} \tag{3.3}$$

The same constructions hold for non necessarily increasing maps. In this situation, the sequences of maps $\alpha \in \mathcal{A}_{\mathbf{p}}$, and $\beta \in \mathcal{A}_{\mathbf{r}}$ can be represented graphically by the jungles $J(\alpha) \in \mathcal{J}_{\mathbf{p}}$, and $J(\beta) \in \mathcal{J}_{\mathbf{r}}$, and we have the (noncommutative) product formula on the set $\mathcal{J}$

$$J(\alpha)J(\beta) = J(\alpha \vee \beta) \in \mathcal{J}_{\mathbf{p}+\mathbf{r}} \tag{3.4}$$

3.3 Automorphism groups on jungles

For any given sequence of integers $\mathbf{p} \in \mathbf{V}_{n+1}$, the product permutation group

$$\mathcal{G}_{\mathbf{p}} := \left(\mathcal{G}_{p_0} \times \mathcal{G}_{p_1} \times ... \times \mathcal{G}_{p_{n+1}}\right)$$

acts naturally on sequences of maps $\mathbf{a} = (a_0, a_1, ..., a_n) \in \mathcal{A}_{\mathbf{p}}$, and on jungles $J(\mathbf{a})$ in $\mathcal{J}_{\mathbf{p}}$ by permutation of the vertices at each level. More formally, for any $\mathbf{s} = (s_0, ..., s_{n+1}) \in \mathcal{G}_{\mathbf{p}}$ this pair of actions is given by

$$\mathbf{s}(\mathbf{a}) := (s_0 a_0 s_1^{-1}, ..., s_n a_n s_{n+1}^{-1}) \quad \text{and} \quad \mathbf{s}J(\mathbf{a}) := J(\mathbf{s}(\mathbf{a}))$$

An automorphism $\mathbf{s} \in \mathcal{G}_{\mathbf{p}}$ of a given jungle $J(\mathbf{a}) \in \mathcal{J}_{\mathbf{p}}$ is a sequence of permutations that preserves the jungle, in the sense that $\mathbf{s}J(\mathbf{a}) = J(\mathbf{a})$. The set of automorphisms of a given jungle $J(\mathbf{a}) \in \mathcal{J}_{\mathbf{p}}$ coincides with the stabilizer of $J(\mathbf{a})$ with respect to the group action.

$$Stab(J(\mathbf{a})) = \{\mathbf{s} \in \mathcal{G}_{\mathbf{p}} \ : \ \mathbf{s}J(\mathbf{a}) = J(\mathbf{a})\}$$

By the definition of a jungle, we also have that

$$J(\mathbf{a}) = J(\mathbf{b}) \Longleftrightarrow \mathbf{a} = \mathbf{b}$$

from which we find $Stab(J(\mathbf{a}))$ coincide with the stabilizer of $\mathbf{a}$ with respect to the group action. The stabilizer theorem also provides a one to one correspondence $\varphi_{\mathbf{a}}$ between the orbit set

$$Orb(J(\mathbf{a})) := \{\mathbf{s}J(\mathbf{a}) \ : \ \mathbf{s} \in \mathcal{G}_{\mathbf{p}}\} \subset \mathcal{J}_{\mathbf{p}}$$

and the quotient $\mathcal{G}_{\mathbf{p}}/Stab(J(\mathbf{a}))$

$$\varphi_{\mathbf{a}} \ : \ J(\mathbf{b}) \in Orb(J(\mathbf{a})) \mapsto \mathbf{s}_{(\mathbf{a},\mathbf{b})} Stab(J(\mathbf{a})) \in \mathcal{G}_{\mathbf{p}}/Stab(J(\mathbf{a}))$$

for any choice of $\mathbf{s}_{(\mathbf{a},\mathbf{b})} \in \mathcal{G}_{\mathbf{p}}$ such that $\mathbf{s}_{(\mathbf{a},\mathbf{b})}(\mathbf{a}) = \mathbf{b}$.

Notice that if two sequences $\mathbf{a}$ and $\mathbf{b}$ differ only by the order of the vertices in $J(\mathbf{a})$ and $J(\mathbf{b})$, that is by the action of an element of $\mathcal{G}_{\mathbf{p}}$, then the associated forests are identical: $F(\mathbf{a}) = F(\mathbf{b})$. Moreover, the converse is true: if $F(\mathbf{a}) = F(\mathbf{b})$, then $J(\mathbf{a})$ and $J(\mathbf{b})$ differ only by the ordering of the vertices, since they have the same underlying non planar graph. In this situation, $\mathbf{a}$ and $\mathbf{b}$ belong to the same orbit under the action of $\mathcal{G}_{\mathbf{p}}$. In particular, we get a result already mentioned in the Introduction

Lemma 3.4. *The set of equivalence classes of jungles in $\mathcal{J}_{\mathbf{p}}$ under the action of the permutation groups $\mathcal{G}_{\mathbf{p}}$ is in bijection with the set of forests $\mathcal{F}_{\mathbf{p}}$. In particular, due to the equivalence between the two notions of jungles and sequences of maps, it follows that the set $\mathcal{F}_{n,q}$ of equivalence classes in $\mathcal{A}_{n,q}$ under the natural action of $\mathcal{G}_{\mathbf{q}}$ is canonically in bijection with the set of forests with q vertices at each level $0 \leq k \leq n+1$.*

Here, $\mathbf{q}$ denotes, once again, the q-constant sequence of lenght $n+1$.

Since the root mapping a_0 of a planar tree is constant and equal to 1, we finally observe that for any $\mathbf{a} = (a_0, \ldots, a_n) \in \mathcal{IA}_{\mathbf{p}}$, we have

$$Stab(J(\mathbf{a}) = Stab(PT(\mathbf{a})) = \{1_1\} \times Stab(B(PT(\mathbf{a})))$$

3.4 Orbit sets of planar forests

From the Feynman-Kac mean-field approximation point of view, jungles in $\mathcal{A}_{n,q}$ describe trajectories of families of particles with prescribed coalescence properties. In particular, two trajectories that are equivalent under the action of $\mathcal{G}_{\mathbf{q}}$ have the same statistical properties. In the present section, we face the general problem of computing the cardinals $\#(\mathbf{f})$, defined as the number of jungles in $\mathbf{f} \in \mathcal{F}_{\mathbf{p}}$, for some integer sequence $\mathbf{p} \in \mathbf{V}_{n+1}$, when $\mathbf{f}$ is viewed as an equivalence class according to the previous lemma. Let us write $\mathbf{f}$ in normal form (with the T_is all distincts) as a commutative monomial of trees:

$$\mathbf{f} = T_1^{m_1} ... T_k^{m_k}, \ m_1 + ... + m_k = p_0$$

Let us also choose arbitrary representations $T_i' \in \mathcal{PT}$ of the T_is as planar trees, so that we can view $\mathbf{f}$ as the forest associated to the planar forest

$$\mathbf{f}' = (T_1')^{m_1} ... (T_k')^{m_k}$$

obtained by left-to-right concatenation of m_1 copies of T'_1, m_2 copies of T'_2,..., m_k copies of T'_k. We write $T''_1...T''_{p_0}$ for the expansion of the noncommutative monomial

$$\mathbf{f}' = (T'_1)^{m_1}...(T'_k)^{m_k} = T''_1...T''_{p_0}$$

as a product of planar trees (without exponents).

From previous considerations, we have that

$$\#(\mathbf{f}) = \#Orb(\mathbf{f}') = \#\left\{\mathbf{g} \in \mathcal{J}_{\mathbf{p}} \ : \ \exists \mathbf{s} \in \mathcal{G}_{\mathbf{p}} \quad \text{s.t.} \quad \mathbf{g} = \mathbf{s}\mathbf{f}'\right\}$$

Due to the class formula, we also know that

$$\#(\mathbf{f}) = \#Orb(\mathbf{f}') = \frac{|\mathcal{G}_{\mathbf{p}}|}{Stab(\mathbf{f}')} = \frac{\mathbf{p}!}{Stab(\mathbf{f}')}$$

where we remind that $Stab(\mathbf{f}')$ stands for the stabilizer of the jungle $\mathbf{f}' \in \mathcal{J}_{\mathbf{p}}$.

The computation of $\#(\mathbf{f})$ will be done by induction on $(n+1) := ht(\mathbf{f})$. Let us assume that we know $\#(\mathbf{g})$ for any forest of height less or equal n. Notice that, then, we also know $\#(\mathbf{t})$ for any tree $\mathbf{t}$ of height $(n+1)$, due to the canonical bijection B between trees and forests: $\#(\mathbf{t}) = \#(B\mathbf{t})$. From the previous discussion, the problem amounts to compute the cardinals of the stabilizers $Stab(\mathbf{f}')$ inductively with respect to the height parameter. The following technical lemma is instrumental.

Lemma 3.5. *There is a natural isomorphism*

$$Stab(T''_1...T''_{p_0}) \cap \left(\{1_{p_0}\} \times \mathcal{G}_{(p_1,\ldots,p_{n+1})}\right) \sim Stab(T''_1) \vee \ldots \vee Stab(T''_{p_0})$$

where 1_{p_0} *stands for the identity in* $\mathcal{G}_{p_0}$.

Proof. We let $\mathbf{a}^i = (a^i_0, \ldots, a^i_n)$ be a sequence of weakly increasing maps such that $T''_i = PT(\mathbf{a}^i)$, with $1 \leq i \leq p_0$. In this notation, we have that

$$\mathbf{f}' \quad = \quad T''_1...T''_{p_0} = PT(\mathbf{a}^1)\ldots PT(\mathbf{a}^{p_0}) = PF(\mathbf{a})$$

where $PF(\mathbf{a})$ stands for the planar forest associated with the sequence of weakly increasing maps

$$\mathbf{a} = (\mathbf{a}^1 \vee \ldots \vee \mathbf{a}^{p_0}) = (a_0, \ldots, a_n) \quad \text{with} \quad \forall 0 \leq l \leq n \qquad a_l = a^1_l \vee \ldots \vee a^{p_0}_l$$

Notice that the stabilizer of the planar forest $PF(\mathbf{a})$ is given by

$$Stab(PF(\mathbf{a})) = \{\mathbf{s} \in \mathcal{G}_{\mathbf{p}} \ : \ \mathbf{s}(\mathbf{a}) = \mathbf{a}\}$$

and for any $\mathbf{s} = (s_0, \ldots, s_{n+1}) \in \mathcal{G}_{\mathbf{p}}$ we have that

$$\mathbf{s}(\mathbf{a}) = \mathbf{a} \Leftrightarrow \left(\forall 0 \leq l \leq n \qquad s_l \circ (a^1_l \vee \ldots \vee a^{p_0}_l) \circ s^{-1}_{l+1} = a^1_l \vee \ldots \vee a^{p_0}_l\right)$$

Therefore, using the composition formula (3.1) and the factorization rule following the formula, it is not difficult to check that any sequence

$$\mathbf{s} = (1_{p_0}, s_1, \ldots, s_{n+1}) \in Stab(\mathbf{f}')$$

can be represented in an unique way $\mathbf{s} = \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right)$, in terms of some permutation sequences

$$\mathbf{s}^{(i)} = (1_1, s^{(i)}_1, \ldots, s^{(i)}_{n+1})$$

between the vertices at each level of the planar tree $PT(\mathbf{a}^i)$, with $1 \leq i \leq p_0$. In terms of the level parameter $1 \leq l \leq (n+1)$, we also notice that

$$s_l = s^{(1)}_l \vee \ldots \vee s^{(p_0)}_l$$

To take the final step, we observe that

$$\begin{aligned} s_l a_l s^{-1}_{l+1} &= \left(s^{(1)}_l \vee \ldots \vee s^{(p_0)}_l\right) \circ (a^1_l \vee \ldots \vee a^{p_0}_l) \circ \left((s^{(1)}_{l+1})^{-1} \vee \ldots \vee (s^{(p_0)}_{l+1})^{-1}\right) \\ &= (s^{(1)}_l \circ a^1_l \circ (s^{(1)}_{l+1})^{-1}) \vee \ldots \vee (s^{(p_0)}_l \circ a^{p_0}_l \circ (s^{(p_0)}_{l+1})^{-1}) \end{aligned}$$

This implies that

$$\mathbf{s}(\mathbf{a}) = \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right)(\mathbf{a}^1 \vee \ldots \vee \mathbf{a}^{p_0}) = (\mathbf{s}^{(1)}(\mathbf{a}^1) \vee \ldots \vee \mathbf{s}^{(p_0)}(\mathbf{a}^{p_0}))$$

from which we find that $\mathbf{s}J(\mathbf{a}) = J(\mathbf{a})$ if, and only if, we have

$$J(\mathbf{s}^{(1)}(\mathbf{a}^1)) \ldots J(\mathbf{s}^{(p_0)}(\mathbf{a}^{p_0})) \quad = \quad J(\mathbf{a}^1) \ldots J(\mathbf{a}^{p_0})$$

This is clearly equivalent to the fact that

$$J(\mathbf{s}^{(i)}(\mathbf{a}^i)) = \mathbf{s}^{(i)} J(\mathbf{a}^i) = J(\mathbf{a}^i)$$

for any $1 \leq i \leq p_0$. This yields that for any $1 \leq i \leq p_0$, we have

$$\mathbf{s}^{(i)} = (1_1, s_1^{(i)}, \ldots, s_{n+1}^{(i)}) \in Stab(PT(\mathbf{a}^i))$$

from which we conclude that the mapping

$$\mathbf{s} \in Stab(PT(\mathbf{a}^1) \ldots PT(\mathbf{a}^{p_0})) \cap \left(\{1_{p_0}\} \times \mathcal{G}_{(p_1,\ldots,p_{n+1})}\right)$$

$$\downarrow$$

$$\left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right) \in Stab(T(\mathbf{a}^1)) \vee \ldots \vee Stab(T(\mathbf{a}^{p_0}))$$

is an isomorphism. This ends the proof of the lemma. ∎

The first coordinate mapping

$$\pi \ : \ \mathbf{s} = (s_0, \ldots, s_{n+1}) \in Stab((T_1')^{m_1} \ldots (T_k')^{m_k}) \mapsto \pi(\mathbf{s}) = s_0 \in (\mathcal{G}_{m_1} \vee \ldots \vee \mathcal{G}_{m_k})$$

is a surjective map (it even has a natural section, the construction of which is omitted). We therefore have that

$$Ker(\pi) = \pi^{-1}(1_{p_0}) \sim Stab(T_1'') \vee \ldots \vee Stab(T_{p_0}'') \sim Stab(T_1')^{m_1} \times \ldots \times Stab(T_k')^{m_k}$$

This yields the isomorphism formula

$$Stab((T_1')^{m_1} \ldots (T_k')^{m_k}) / \left(Stab(T_1')^{m_1} \times \ldots \times Stab(T_k')^{m_k}\right) \sim (\mathcal{G}_{m_1} \vee \ldots \vee \mathcal{G}_{m_k})$$

from which we readily deduce the following recursive formula for the computation of $\#(\mathbf{f})$.

Proposition 3.6. *We have*

$$|Stab((T_1')^{m_1} \ldots (T_k')^{m_k})| = \prod_{i=1}^{k} \left(m_i! \ |Stab\ (B(T_i'))|^{m_i}\right)$$

3.5 An inductive method for counting jungles

We conclude this section by giving a closed formula for $\#(\mathbf{f})$. Let us introduce for that purpose some further notations.

Definition 3.7. *Let $\mathbf{f}$ be a forest written in normal form $\mathbf{f} = T_1^{m_1} \ldots T_k^{m_k}$. In this situation, we say that the unordered k-uplet $(m_1, \ldots, m_k)$ is the symmetry multiset of the tree $T = B^{-1}(\mathbf{f})$ and write*

$$\mathbf{s}(T) = \mathbf{s}(B^{-1}(T_1^{m_1} \ldots T_k^{m_k})) = (m_1, \ldots, m_k)$$

The symmetry multiset of the forest $\mathbf{s}(\mathbf{f}) = T_1^{m_1} \ldots T_k^{m_k}$ is the disjoint union of the symmetry multisets of its trees

$$\mathbf{s}(T_1^{m_1} \ldots T_k^{m_k}) = \left(\underbrace{\mathbf{s}(T_1), \ldots, \mathbf{s}(T_1)}_{m_1 \text{ terms}}, \ldots, \underbrace{\mathbf{s}(T_k), \ldots, \mathbf{s}(T_k)}_{m_k \text{ terms}}\right)$$

For example, the symmetry multiset of the first tree of the forest $\mathbf{f}$ displayed in Fig. 1 is $(1,1)$; the symmetry multiset of $B^{-1}(\mathbf{f})$ is $(2,1)$; and the symmetry multiset of $\mathbf{f}$ is $(1,1,3,1,1)$.

We also extend the map B from trees to forests to a map between forests and, if $\mathbf{f} = T_1^{m_1}...T_k^{m_k}$ is a forest, set

$$B(\mathbf{f}) = B(T_1)^{m_1}...B(T_k)^{m_k}$$

From the point of view of graphs, the operations amounts to removing all the roots and all the edges that have the root as terminal vertex from the graph defining $\mathbf{f}$. The definition extends naturally to planar trees and planar forests. In particular, if $\mathbf{f}'$ is a planar forest with $\mathbf{f}$ as its underlying forest, we set $\mathbf{s}(\mathbf{f}') := \mathbf{s}(\mathbf{f})$.

Theorem 3.8. *The number of jungles (or, equivalently, of Feynman-Kac mean-field type trajectories with the same statistics) in $\mathbf{f} \in \mathcal{F}_{\mathbf{p}}$, with $\mathbf{p} \in \mathbf{V}_{n+1}$, is given by*

$$\#(\mathbf{f}) = \frac{\mathbf{p}!}{\prod_{i=-1}^{n} \mathbf{s}(B^i(\mathbf{f}))!}$$

where we use the usual multi-index notation to define $\mathbf{s}(B^i(\mathbf{f}))!$.

Proof. It clearly suffices to prove that

$$|Stab(\mathbf{f}')| = \prod_{i=-1}^{n} \mathbf{s}(B^i(\mathbf{f}'))!$$

for any planar forest $\mathbf{f}'$ of height $(n+1)$. We check this assertion by induction on the height parameter. First, we observe that a planar forest $\mathbf{f}'$ of height 1 can be represented as a noncommutative monomial

$$\mathbf{f}' = (T_1')^{m_1}...(T_k')^{m_k}$$

with different planar trees T_i' of height 1, and some sequence of integers m_i. In that case, the planar forests $B(T_i')$ reduce to elementary planar forests with null height, with $S(T_i')$ elementary planar trees with null height. This yields that

$$ht(T_i) = 1 \Longrightarrow |Stab(B(T_i'))| = \mathbf{s}(T_i')!$$

By proposition 3.6, we conclude that

$$|Stab(\mathbf{f}')| = \prod_{j=1}^{k} \left(m_j!\ (\mathbf{s}(T_j')!)^{m_j}\right) = \mathbf{s}(B^{-1}(\mathbf{f}'))!\ \mathbf{s}(\mathbf{f}')!$$

This ends the proof of the formula at rank 1. Suppose now that the assertion is satisfied for any planar forest with height n. By proposition 3.6, for any planar forest $\mathbf{f}' = (T_1')^{m_1}...(T_k')^{m_k}$, with height $(n+1)$, and written in terms of distinct planar trees T_i, we have

$$|Stab(\mathbf{f}')| = \mathbf{s}(B^{-1}(\mathbf{f}'))! \prod_{i=1}^{k} |Stab(B(T_i'))|^{m_i}$$

Since the planar forests $B(T_i')$ have height n, the induction hypothesis implies that

$$|Stab(B(T_i'))| = \prod_{j=-1}^{n-1} \mathbf{s}(B^j(B(T_i'))) = \prod_{j=0}^{n} \mathbf{s}(B^j(T_i'))!$$

This yields that

$$|Stab(\mathbf{f}')| = \mathbf{s}(B^{-1}(\mathbf{f}'))! \prod_{i=1}^{k} \left(\prod_{j=0}^{n} \mathbf{s}(B^j(T_i'))! \right)^{m_i} = \mathbf{s}(B^{-1}(\mathbf{f}'))! \prod_{j=0}^{n} \left(\prod_{i=1}^{k} \mathbf{s}(B^j(T_i'))! \right)^{m_i}$$

Recalling that $B^j(\mathbf{f}') = B^j(T_1')^{m_1}...B^j(T_k')^{m_k}$, we also find that

$$\mathbf{s}(B^j(\mathbf{f}')) = (\underbrace{\mathbf{s}(B^j(T_1')),\ldots,\mathbf{s}(B^j(T_1'))}_{m_1 \text{ terms}},\ldots,\underbrace{\mathbf{s}(B^j(T_k')),\ldots,\mathbf{s}(B^j(T_k'))}_{m_k \text{ terms}})$$

We conclude that

$$|Stab(\mathbf{f}')| = \mathbf{s}(B^{-1}(\mathbf{f}'))! \prod_{j=0}^{n} \mathbf{s}(B^j(\mathbf{f}'))!$$

This shows that the formula is also true for planar forests with height $(n+1)$. The inductive proof of the theorem is now completed. ∎

3.6 Wreath product representation

This section is concerned with an original wreath product interpretation of the stabilizer of planar forests. To be self-contained, let us recall (see e.g. [8]) that the wreath product, written $\mathcal{G}_k < G >$ of an arbitrary group G with the symmetric group of order k, $\mathcal{G}_k$, is the group that identifies, as a set, with $\mathcal{G}_k \times G^k$, and where the product is defined by:

$$(s, g_1, ..., g_k) \odot (s', g'_1, ..., g'_k) := (ss', g_1 g'_{s^{-1}(1)}, ..., g_k g'_{s^{-1}(k)})$$

In addition, if a group G acts on a set S, and if $\mathcal{G}_k$ acts by permutation on S^k, that is, if for any $s \in \mathcal{G}_k$ and any $(x_1, ..., x_k) \in S^k$ the action of s on $(x_1, ..., x_k)$ is given by

$$s(x_1, ..., x_k) = (x_{s^{-1}(1)}, ..., x_{s^{-1}(k)})$$

we then have, for any $(g_1, ..., g_k) \in G^k$

$$\begin{aligned} s(g_1, ..., g_k)s^{-1}(x_1, ..., x_k) &= s(g_1, ..., g_k)(x_{s(1)}, ..., x_{s(k)}) \\ &= s(g_1 x_{s(1)}, ..., g_k x_{s(k)}) \\ &= (g_{s^{-1}(1)} x_1, ..., g_{s^{-1}(k)} x_k) \end{aligned}$$

We are now in position to state and prove the main result of this section.

Theorem 3.9. *For any planar forest $(T'_1)^{m_1}...(T'_k)^{m_k} \in \mathcal{PF}_{\mathbf{p}}$, with height $(n+1)$ written in terms of distinct planar trees T'_i, for some sequence of integers $\mathbf{p} \in \mathbf{V}_{n+1}$, there is a natural isomorphism*

$$Stab((T'_1)^{m_1}...(T'_k)^{m_k}) \sim (\mathcal{G}_{m_1} < Stab(T'_1) >) \vee \ldots \vee (\mathcal{G}_{m_k} < Stab(T'_k) >)$$

so that, in particular, the stabilizer of any jungle in the equivalence class of $\mathbf{f}$ is conjugated in $\mathcal{G}_{\mathbf{p}}$ to $(\mathcal{G}_{m_1} < Stab(T'_1) >) \vee \ldots \vee (\mathcal{G}_{m_k} < Stab(T'_k) >)$.

Proof. We write $T''_1...T''_{p_0}$ for the expansion of the non commutative monomial $(T'_1)^{m_1}...(T'_k)^{m_k}$ as a product of planar trees (without exponents). Let us also choose an arbitrary

$$\mathbf{s} = (s_0, ..., s_{n+1}) \in Stab(\mathbf{f}')$$

This implies in particular that the root permutation $s_0 \in \mathcal{G}_{p_0}$ is of the form

$$s_0 = (s_0^{(1)} \vee \ldots \vee s_0^{(k)}) \in (\mathcal{G}_{m_1} \vee \ldots \vee \mathcal{G}_{m_k})$$

Also observe that s_0 acts on the planar forest $T''_1...T''_{p_0}$ by permutation of the planar trees:

$$s_0[T''_1...T''_{p_0}] := (T''_{s_0^{-1}(1)}...T''_{s_0^{-1}(p_0)})$$

This action by permutation can be viewed as a permutation (by blocks, according to the permutation of the planar trees) of the vertices of the planar forest, and this action can therefore be rewritten uniquely under this process as the action of an element

$$\mathbf{s}_0 := (s_0, s_{0,1}, \ldots, s_{0,n+1}) \in \mathcal{G}_{\mathbf{p}}$$

on the vertices of the planar graph of $T''_1...T''_{p_0}$

$$s_0[T''_1...T''_{p_0}] = \mathbf{s}_0(T''_1...T''_{p_0})$$

Notice that $\mathbf{s}_0$ depends on the planar forest $T_1''...T_{p_0}''$, since it acts non trivially on the vertices at levels $i \geq 1$. In general, the embedding $s_0 \in \mathcal{G}_{p_0} \hookrightarrow \mathbf{s}_0 \in \mathcal{G}_{\mathbf{p}}$ is different from the canonical embedding $s_0 \in \mathcal{G}_{p_0} \hookrightarrow (s_0, 1_{p_1}, ..., 1_{p_{n+1}}) \in \mathcal{G}_{\mathbf{p}}$.

We consider then the map (of sets)

$$\begin{array}{rcl} \mu \ : \ Stab(\mathbf{f}') & \mapsto & (\mathcal{G}_{m_1} \vee \ldots \vee \mathcal{G}_{m_k}) \times \left(Stab(\mathbf{f}') \cap \left(\{1_{p_0}\} \times \mathcal{G}_{(p_1,\ldots,p_{n+1})}\right)\right) \\ \mathbf{s} = (s_0, ..., s_{n+1}) & \mapsto & \mu(\mathbf{s}) = (s_0, \mathbf{s}\mathbf{s}_0^{-1}) \end{array}$$

In lemma 3.5, we have proved that $\left(Stab(T_1''...T_{p_0}'') \cap \left(\{1_{p_0}\} \times \mathcal{G}_{(p_1,\ldots,p_{n+1})}\right)\right)$ is isomorphic to the permutation group $Stab(T_1'') \vee \ldots \vee Stab(T_{p_0}'')$. Therefore, the permutation sequence

$$\mathbf{s}\mathbf{s}_0^{-1} = (1_{p_0}, s_1 s_{0,1}^{-1}, \ldots, s_{n+1} s_{0,n+1}^{-1})$$

can be alternatively represented as

$$\mathbf{s}\mathbf{s}_0^{-1} = \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right) \in Stab(T_1'') \vee \ldots \vee Stab(T_{p_0}'')$$

Using these representations, we have that

$$\begin{array}{rcl} \mu(\mathbf{s}) & = & \left((s_0^{(1)} \vee \ldots \vee s_0^{(k)}), \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right)\right) \\ & = & \left(s_0^{(1)}, \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(m_1)}\right)\right) \vee \left(s_0^{(2)}, \left(\mathbf{s}^{(m_1+1)} \vee \ldots \vee \mathbf{s}^{(m_1+m_2)}\right)\right) \\ & & \vee \ldots \vee \left(s_0^{(k)}, \left(\mathbf{s}^{(m_1+\ldots+m_{k-1}+1)} \vee \ldots \vee \mathbf{s}^{(m_1+\ldots+m_{k-1}+m_k)}\right)\right) \end{array}$$

On the other hand, we have

$$\begin{array}{rcl} \mu(\mathbf{s}\mathbf{t}) & = & (s_0 t_0, \mathbf{s}\mathbf{t}(\mathbf{s}_0\mathbf{t}_0)^{-1}) \\ & = & (s_0 t_0, \mathbf{s}\mathbf{t}\mathbf{t}_0^{-1}\mathbf{s}_0^{-1}) = (s_0 t_0, (\mathbf{s}\mathbf{s}_0^{-1})\mathbf{s}_0(\mathbf{t}\mathbf{t}_0^{-1})\mathbf{s}_0^{-1}) \end{array}$$

and for any level index $1 \leq l \leq n$ we find that

$$\begin{array}{rcl} s_{0,l}(t_l t_{0,l}^{-1}) s_{0,l}^{-1} & = & s_{0,l}(t_l^{(1)} \vee \ldots \vee t_l^{(p_0)}) s_{0,l}^{-1} \\ & = & (t_l^{((s_0^{(1)})^{-1}(1))} \vee \ldots \vee t_l^{((s_0^{(1)})^{-1}(m_1))}) \\ & & \vee (t_l^{((s_0^{(2)})^{-1}(m_1+1))} \vee \ldots \vee t_l^{((s_0^{(2)})^{-1}(m_1+m_2))}) \\ & & \vee \ldots \vee (t_l^{((s_0^{(k)})^{-1}(m_1+\ldots+m_{k-1}+1))} \vee \ldots \vee t_l^{((s_0^{(k)})^{-1}(m_1+\ldots+m_{k-1}+m_k))}) \end{array}$$

This implies that

$$\mathbf{s}_0(\mathbf{t}\mathbf{t}_0^{-1})\mathbf{s}_0^{-1} = (1_{p_0}, r_1, \ldots, r_{n+1})$$

with a sequence of mappings

$$(1_{p_0}, r_1, \ldots, r_{n+1}) = \left(\mathbf{r}^{(1)} \vee \ldots \vee \mathbf{r}^{(p_0)}\right) \in Stab(T_1'') \vee \ldots \vee Stab(T_{p_0}'')$$

given by

$$\left\{ \begin{array}{rcl} \mathbf{r}^{(1)} \vee \ldots \vee \mathbf{r}^{(m_1)} & = & \mathbf{t}^{((s_0^{(1)})^{-1}(1))} \vee \ldots \vee \mathbf{t}^{((s_0^{(1)})^{-1}(m_1))} \\ \mathbf{r}^{(m_1+1)} \vee \ldots \vee \mathbf{r}^{(m_1+m_2)} & = & \mathbf{t}^{((s_0^{(2)})^{-1}(m_1+1))} \vee \ldots \vee \mathbf{t}^{((s_0^{(2)})^{-1}(m_1+m_2))} \\ & \vdots & \\ \mathbf{r}^{(m_1+\ldots+m_{k-1}+1)} \vee \ldots \vee \mathbf{r}^{(m_1+\ldots+m_k)} & = & \mathbf{t}^{((s_0^{(k)})^{-1}(m_1+\ldots+m_{k-1}+1))} \vee \ldots \vee \mathbf{t}^{((s_0^{(k)})^{-1}(m_1+\ldots+m_k))} \end{array} \right.$$

This yields that

$$\begin{array}{rcl} (\mathbf{s}\mathbf{s}_0^{-1}) \circ \left(\mathbf{s}_0(\mathbf{t}\mathbf{t}_0^{-1})\mathbf{s}_0^{-1}\right) & = & \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(p_0)}\right) \circ \left(\mathbf{r}^{(1)} \vee \ldots \vee \mathbf{r}^{(p_0)}\right) \\ & = & \left((\mathbf{s}^{(1)} \circ \mathbf{r}^{(1)}) \vee \ldots \vee (\mathbf{s}^{(p_0)} \circ \mathbf{r}^{(p_0)})\right) \end{array}$$

We finally obtain the following formula

$$\begin{aligned}\mu(\mathbf{st}) &= \left(s_0^{(1)} \circ t_0^{(1)}, \left((\mathbf{s}^{(1)} \circ \mathbf{t}^{((s_0^{(1)})^{-1}(1))}) \vee \ldots \vee (\mathbf{s}^{(m_1)} \circ \mathbf{t}^{((s_0^{(1)})^{-1}(m_1))})\right)\right) \\ &\quad \vee \left(s_0^{(2)} \circ t_0^{(2)}, \left((\mathbf{s}^{(m_1+1)} \circ \mathbf{t}^{((s_0^{(2)})^{-1}(m_1+1))}) \vee \ldots \vee (\mathbf{s}^{(m_1+m_2)} \circ \mathbf{t}^{((s_0^{(2)})^{-1}(m_1+m_2))})\right)\right) \\ &\quad \vee \ldots \\ &= \left\{\left(s_0^{(1)}, \left(\mathbf{s}^{(1)} \vee \ldots \vee \mathbf{s}^{(m_1)}\right)\right) \odot \left(t_0^{(1)}, \left(\mathbf{t}^{(1)} \vee \ldots \vee \mathbf{t}^{(m_1)}\right)\right)\right\} \\ &\quad \vee \left\{\left(s_0^{(2)}, \left(\mathbf{s}^{(m_1+1)} \vee \ldots \vee \mathbf{s}^{(m_1+m_2)}\right)\right) \odot \left(t_0^{(2)}, \left(\mathbf{t}^{(m_1+1)} \vee \ldots \vee \mathbf{t}^{(m_1+m_2)}\right)\right)\right\} \\ &\quad \vee \ldots\end{aligned}$$

and the Theorem follows. ∎

3.7 Hilbert series method for forests enumeration

In the present subsection, we face the problem of computing other cardinals relevant to the analysis of Feynman-Kac mean field particle models, according to our Theorem 1.2. For example, we want to be able to compute the number of forests in any $\mathcal{F}_{\mathbf{p}}$; the cardinals $|\mathcal{F}_{n,q}(\mathbf{r})|$; or the number of forests in $\mathcal{F}_{n,q}$ with a given coalescence degree. Notice, by the way, that the notions of coalescence sequence and degree, as defined in the Introduction for sequences of maps in $\mathcal{A}_{n,q}$, and forests in $\mathcal{F}_{n,q}$, go over to arbitrary forests and trees.

We let $\langle \mathcal{S} \rangle$ be the commutative monoid generated by a subset of forests $\mathcal{S} \in \mathcal{F}$

$$\langle \mathcal{S} \rangle = \{f_1^{m_1} \ldots f_k^{m_k} \ : \ k \geq 0, \quad \text{and} \quad \forall 1 \leq i \leq k \quad m_i \geq 0 \quad \text{and} \quad f_i \in \mathcal{S}\}$$

with the convention $f_1^{m_1} \ldots f_k^{m_k} = \emptyset$, for $k = 0$. Notice that there is a canonical map from $\langle \mathcal{S} \rangle$ to $\mathcal{F}$, which is an embedding, e.g. if $\mathcal{S}$ is a set of (distinct) trees. We also consider a set of variables $\chi(T)$ indexed by the set of trees $T \in \mathcal{T}$, and we let $M[\mathcal{T}]$ be the set of monomials defined as follows

$$M[\mathcal{T}] = \{\chi(T_1)^{m_1} \ldots \chi(T_k)^{m_k} \ : \ k \geq 0, \quad \text{and} \quad \forall 1 \leq i \leq k \quad m_i \geq 0 \quad \text{and} \quad T_i \in \mathcal{T}\}$$

For any mapping χ on the set of forests

$$\chi \ : \ \mathbf{f} \in \mathcal{F} \mapsto \chi(\mathbf{f}) \in M[\mathcal{T}] \quad \text{such that} \quad \forall (\mathbf{f}, \mathbf{g}) \in \mathcal{F}^2 \quad \chi(\mathbf{fg}) = \chi(\mathbf{f})\ \chi(\mathbf{g})$$

we have the Hilbert series expansions

$$\frac{1}{1 - \chi(\mathbf{f})} = \sum_{\mathbf{g} \in \langle \mathbf{f} \rangle} \chi(\mathbf{g}) \quad \text{and} \quad \prod_{\mathbf{f} \in \mathcal{S}} \frac{1}{1 - \chi(\mathbf{f})} = \sum_{\mathbf{g} \in \langle \mathcal{S} \rangle} \chi(\mathbf{g}) \tag{3.5}$$

We let $\mathbf{V}_n^\star$ be the subset of all multi-indices $\mathbf{p} \in \mathbf{V}$, with height $ht(\mathbf{p}) \leq n$

$$\mathbf{V}_n^\star := \cup_{-1 \leq k \leq n} \mathbf{V}_k$$

Definition 3.10. *The (multidegree) Hilbert series of forests, $\mathbb{H}_{\mathcal{F}}^n(\mathbf{x})$, is the Hilbert series associated to the partition of the set of forests of height less or equal than n into subsets according to the number of vertices at each level, that is,*

$$\mathbb{H}_{\mathcal{F}}^n(\mathbf{x}) := \sum_{\mathbf{p} \in \mathbf{V}_n^\star} |\mathcal{F}_{\mathbf{p}}|\ \mathbf{x}^{\mathbf{p}} = \mathbb{H}_{\mathcal{F}}^{n-1}(\mathbf{x}) + \sum_{\mathbf{p} \in \mathbf{V}_n} |\mathcal{F}_{\mathbf{p}}|\ \mathbf{x}^{\mathbf{p}}$$

where we write $\mathbf{x}^{\mathbf{p}}$ as a shorthand for $x_0^{p_0} x_1^{p_1} \ldots$.

We let $\partial_{\mathbf{p}} = \partial_{x_0}^{p_0} \partial_{x_1}^{p_1} \ldots$, and we consider the mapping B^{-1} from $\mathbf{V}$ into itself defined for any $k \geq 0$ by

$$B^{-1} \ : \ \mathbf{p} \in \mathbf{V}_k \mapsto B^{-1}(\mathbf{p}) = (1, \mathbf{p}) = (1, p_0, p_1, \ldots) \in \mathbf{V}_{k+1}$$

Since forests $\mathbf{f}$ of height 0 are characterized by the number of roots in $\mathbf{f}$, we clearly have the formula

$$(\forall \mathbf{p} \in \mathbf{V}_0 \quad |\mathcal{F}_{\mathbf{p}}| = 1) \Longrightarrow \mathbb{H}_{\mathcal{F}}^0(x_0) = \frac{1}{1 - x_0} = \sum_{p \geq 0} x_0^p$$

Proposition 3.11.

$$\forall n \geq 1 \qquad \mathbb{H}^n_{\mathcal{F}}(\mathbf{x}) = \prod_{\mathbf{p}\in\mathbf{V}^\star_{n-1}} \left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{|\mathcal{F}_{\mathbf{p}}|} = \mathbb{H}^{n-1}_{\mathcal{F}}(\mathbf{x}) \times \prod_{\mathbf{p}\in\mathbf{V}_{n-1}} \left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{\partial_{\mathbf{p}}\mathbb{H}^{n-1}_{\mathcal{F}}(\mathbf{0})}$$

In addition, the generating function associated with the number of forests with multidegree $\mathbf{p} \in \mathbf{V}_n$*, and built with trees with height* n *is given by*

$$\forall n \geq 1 \qquad \sum_{\mathbf{p}\in\mathbf{V}_n} |\langle \mathcal{T}(n)\rangle \cap \mathcal{F}_{\mathbf{p}}|\ \mathbf{x}^{\mathbf{p}} = \prod_{\mathbf{p}\in\mathbf{V}_{n-1}} \left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{\partial_{\mathbf{p}}\mathbb{H}^{n-1}_{\mathcal{F}}(\mathbf{0})}$$

Before getting into the proof of this result, we notice that the formula stated in the above proposition makes the Hilbert series computable at any finite vertex order, and at any height using any formal computation software. The first two orders can be handly computed. For $n = 0$, we have already seen that

$$\mathbb{H}^0_{\mathcal{F}}(\mathbf{x}) = \frac{1}{1-x_0} = \sum_{p\geq 0} x_0^p \Rightarrow \forall \mathbf{p} \in \mathbf{V}_0 \quad \partial_{\mathbf{p}}\mathbb{H}^0_{\mathcal{F}}(\mathbf{0}) = |\mathcal{F}_{\mathbf{p}}| = 1$$

By the recursion formula we find that

$$\begin{aligned}
\mathbb{H}^1_{\mathcal{F}}(\mathbf{x}) &= \frac{1}{1-x_0} \times \prod_{n\geq 1} \frac{1}{1-x_0 x_1^n} \\
&= \prod_{n\geq 0} \left(\sum_{k_n\geq 0} (x_0\ x_1^n)^{k_n}\right) = \sum_{(k_n)_{n\geq 0}\in\mathbb{N}^{\mathbb{N}}} x_0^{\sum_{n\geq 0} k_n}\ x_1^{\sum_{n\geq 0} n k_n} \\
&= \sum_{p_0\geq 0} x_0^{p_0} + \sum_{\mathbf{p}\in\mathbf{V}_1} |\mathcal{F}_{\mathbf{p}}|\ x_0^{p_0} x_1^{p_1}
\end{aligned}$$

This readily yields that for any $\mathbf{p} \in \mathbf{V}_1$, we have

$$|\mathcal{F}_{\mathbf{p}}| = \#\left\{(k_n)_{n\geq 0} \in \mathbb{N}^{\mathbb{N}} \ :\ \sum_{n\geq 0} k_n = p_0 \quad \text{and} \quad \sum_{n\geq 0} n\ k_n = p_1\right\}$$

More generally, the $(n+1)$-th order Hilbert series are given by the formula

$$\mathbb{H}^{n+1}_{\mathcal{F}}(\mathbf{x}) \quad = \quad \prod_{\mathbf{p}\in\mathbf{V}^\star_n} \left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{|\mathcal{F}_{\mathbf{p}}|}$$

Recalling that for any $m \geq 1$, we have the following formal series expansion

$$\left(\frac{1}{1-u}\right)^m = \sum_{k\geq 0} \frac{(m-1+k)!}{(m-1)!\ k!}\ u^k$$

we find that

$$\begin{aligned}
\mathbb{H}^{n+1}_{\mathcal{F}}(\mathbf{x}) &= \prod_{\mathbf{p}\in\mathbf{V}^\star_n} \sum_{k_{\mathbf{p}}\geq 0} \frac{(|\mathcal{F}_{\mathbf{p}}|-1+k_{\mathbf{p}})!}{(|\mathcal{F}_{\mathbf{p}}|-1)!\ k_{\mathbf{p}}!}\ \mathbf{x}^{k_{\mathbf{p}}\ B^{-1}(\mathbf{p})} \\
&= \sum_{k\in\mathbb{N}^{\mathbf{V}^\star_n}} \left(\prod_{\mathbf{p}\in\mathbf{V}^\star_n} \frac{(|\mathcal{F}_{\mathbf{p}}|-1+k(\mathbf{p}))!}{(|\mathcal{F}_{\mathbf{p}}|-1)!\ k(\mathbf{p})!}\right) \mathbf{x}^{\sum_{\mathbf{p}\in\mathbf{V}^\star_n} k(\mathbf{p})\ B^{-1}(\mathbf{p})}
\end{aligned}$$

This implies that that for any $\mathbf{p} \in \mathbf{V}_{n+1}$, we have

$$|\mathcal{F}_{\mathbf{p}}| = \sum_{\substack{k:\mathbf{q}\in\mathbf{V}^\star_n\mapsto k(\mathbf{q})\in\mathbb{N} \\ \sum_{\mathbf{q}\in\mathbf{V}^\star_n} k(\mathbf{q})\ B^{-1}(\mathbf{q})=\mathbf{p}}} \left(\prod_{\mathbf{q}\in\mathbf{V}^\star_n} \frac{(|\mathcal{F}_{\mathbf{q}}|-1+k(\mathbf{q}))!}{(|\mathcal{F}_{\mathbf{q}}|-1)!\ k(\mathbf{q})!}\right) \tag{3.6}$$

Now, we come to the proof of proposition 3.11.
Proof of proposition 3.11: We consider the set $\mathcal{S}(n)$ of all trees with height less or equal to n

$$\begin{aligned}\mathcal{S}(n)=\mathcal{T}^\star(n) &:= \cup_{-1\leq k\leq n}\mathcal{T}(k)=\{\emptyset\}\cup\left(\cup_{\mathbf{p}\in\mathbf{V}_n^\star:p_0=1}\mathcal{T}_{\mathbf{p}}\right)\\ &= \{\emptyset\}\cup\left(\cup_{0\leq k\leq n}\cup_{\mathbf{p}\in\mathbf{V}_k:p_0=1}\mathcal{T}_{\mathbf{p}}\right)\end{aligned}$$

The set of forests generated by $\mathcal{S}(n)$ coincide with the set $\mathcal{F}^\star(n)$ of all the forests with height less or equal to n. That is, we have that

$$\langle\mathcal{S}(n)\rangle=\mathcal{F}^\star(n):=\cup_{-1\leq k\leq n}\mathcal{F}(k)=\cup_{-1\leq k\leq n}\cup_{\mathbf{p}\in\mathbf{V}_k}\mathcal{F}_{\mathbf{p}}$$

To take the final step, we consider the mapping

$$\forall\mathbf{p}\in\mathbf{V}_n\quad\forall\mathbf{f}\in\mathcal{F}_{\mathbf{p}}\qquad\chi(\mathbf{f})=\mathbf{x}^{v(\mathbf{f})}=x_0^{p_0}x_1^{p_1}\ldots x_n^{p_n}$$

From previous decompositions, and using the Hilbert series expansions (3.5), we find that

$$\sum_{\mathbf{f}\in\langle\mathcal{S}(n)\rangle}\mathbf{x}^{v(\mathbf{f})}=\sum_{-1\leq k\leq n}\sum_{\mathbf{p}\in\mathbf{V}_k}|\mathcal{F}_{\mathbf{p}}|\,\mathbf{x}^{\mathbf{p}}=\sum_{\mathbf{p}\in\mathbf{V}_n^\star}|\mathcal{F}_{\mathbf{p}}|\,\mathbf{x}^{\mathbf{p}}$$

Recalling that

$$B\ :\ \mathbf{p}\in\mathbf{V}_k\to B(\mathbf{p})\in\mathbf{V}_{k-1}\quad\text{and}\quad|\mathcal{T}_{\mathbf{p}}|=|B(\mathcal{T}_{\mathbf{p}})|=|\mathcal{F}_{B(\mathbf{p})}|$$

we also have that

$$\begin{aligned}\prod_{\mathbf{f}\in\mathcal{S}(n)}\frac{1}{1-\mathbf{x}^{v(\mathbf{f})}} &= \frac{1}{1-x_0}\prod_{\mathbf{p}\in\mathbf{V}_n^\star:p_0=1}\left(\frac{1}{1-\mathbf{x}^{\mathbf{p}}}\right)^{|\mathcal{T}_{\mathbf{p}}|}\\ &= \frac{1}{1-x_0}\prod_{0\leq k\leq n}\prod_{\mathbf{p}\in\mathbf{V}_k:p_0=1}\left(\frac{1}{1-\mathbf{x}^{\mathbf{p}}}\right)^{|\mathcal{T}_{\mathbf{p}}|}\\ &= \frac{1}{1-x_0}\prod_{0\leq k\leq n}\prod_{\mathbf{p}\in\mathbf{V}_k:p_0=1}\left(\frac{1}{1-x_0\ B(\mathbf{x})^{B(\mathbf{p})}}\right)^{|B(\mathcal{T}_{\mathbf{p}})|}\end{aligned}$$

with the shift operator $B((x_k)_{k\geq0})=(x_{k+1})_{k\geq0}$. This implies that

$$\prod_{\mathbf{f}\in\mathcal{S}(n)}\frac{1}{1-\mathbf{x}^{v(\mathbf{f})}}=\prod_{0\leq k\leq n}\prod_{\mathbf{p}\in\mathbf{V}_{k-1}}\left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{|\mathcal{F}_{\mathbf{p}}|}=\prod_{\mathbf{p}\in\mathbf{V}_{n-1}^\star}\left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}}\right)^{|\mathcal{F}_{\mathbf{p}}|}$$

The end of the proof follows then from the Hilbert series formula (3.5). ∎

Recall now that the signed measures in the Laurent expansion of the distributions $\mathbb{Q}_{n,q}^N$ are indexed by the coalescence degrees of forests. To compute or estimate these Laurent expansions, we are therefore interested in a more precise Hilbert series, namely the one taking into account, besides the multidegrees of forests, their coalescence numbers.

Definition 3.12. *We denote by $\mathcal{F}_{\mathbf{p}}[\mathbf{q}]:=\mathcal{F}_{\mathbf{p}}\cap c^{-1}(\mathbf{q})$ the set of forests in $\mathcal{F}_{\mathbf{p}}$ with a prescribed coalescence sequence $\mathbf{q}$ (we also use the convention $\mathcal{F}_{\mathbf{0}}[\mathbf{0}]=\{\emptyset\}$).*

We let $\mathbb{C}_{\mathcal{F}}^n(\mathbf{x},\mathbf{y})$ be the (multidegree) Hilbert series of forests associated to the partition of the set of forests of height less or equal than n, and prescribed coalescence sequences into subsets parametrized by multidegrees, that is

$$\mathbb{C}_{\mathcal{F}}^n(\mathbf{x},\mathbf{y}):=\sum_{\mathbf{p}\in\mathbf{V}_n^\star}\sum_{\mathbf{q}\in C(\mathbf{p})}|\mathcal{F}_{\mathbf{p}}[\mathbf{q}]|\,\mathbf{x}^{\mathbf{p}}\,\mathbf{y}^{\mathbf{q}}$$

In the above display, $C(\mathbf{p})$ stands for the set of coalescence multidegrees

$$\forall\mathbf{p}\in\mathbf{V}\qquad C(\mathbf{p}):=c(\mathcal{F}_{\mathbf{p}})=\left\{\mathbf{q}\in\mathbb{N}^{\mathbb{N}}\ :\ \text{and}\ \ \mathbf{q}\leq(B(\mathbf{p})-1)_+\right\}$$

with $(B(\mathbf{p})-1)_+=((p_k-1)_+)_{k\geq1}$, for any $\mathbf{p}=(p_k)_{k\geq0}\in\mathbf{V}$.

Proposition 3.13. *The multidegree Hilbert series of coalescent forests $\mathbb{C}^n_{\mathcal{F}}(\mathbf{x},\mathbf{y})$ satisfies the recursive formulae*

$$\mathbb{C}^n_{\mathcal{F}}(\mathbf{x},\mathbf{y}) \;=\; \mathbb{C}^{n-1}_{\mathcal{F}}(\mathbf{x},\mathbf{y}) \prod_{\mathbf{p}\in\mathbf{V}_{n-1}} \prod_{\mathbf{q}\in C(B^{-1}(\mathbf{p}))} \left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}\ \mathbf{y}^{\mathbf{q}}}\right)^{\left|\partial_{\mathbf{p}}\partial_{B(\mathbf{q})}\mathbb{C}^{n-1}_{\mathcal{F}}(\mathbf{0},\mathbf{0})\right|}$$

Using the same lines of arguments as before, this proposition can used to derive a recursive formula for the explicit combinatorial calculation of the number of forests prescribed heights, and coalescence multi-indices. Notice that the first two orders are given by

$$\mathbb{C}^0_{\mathcal{F}}(\mathbf{x},\mathbf{y}) = \frac{1}{1-x_0} \quad \text{and} \quad \mathbb{C}^1_{\mathcal{F}}(\mathbf{x},\mathbf{y}) = \prod_{p_0\geq 0} \frac{1}{(1-x_0x_1^{p_0}\ y_0^{(p_0-1)_+})}$$

Now, we come to the proof of proposition 3.13.

Proof. We use the same notation as the one used in the proof of proposition 3.11. Firstly, we notice that $\mathcal{F}_{\mathbf{p}} = \cup_{\mathbf{q}\in C(\mathbf{p})}\mathcal{F}_{\mathbf{p}}[\mathbf{q}]$, and $ht(\mathbf{q}) \leq (ht(\mathbf{p})-1)$, for any $\mathbf{q}\in C(\mathbf{p})$, and where $ht(\mathbf{q}) := \inf\{n\geq 0 : \sup_{m>n} q_m = 0\}$, for any $\mathbf{q} = (q_m)_{m\geq 0}$. Also observe that

$$\mathcal{F}_{B^{-1}(\mathbf{p})} = \mathcal{T}_{(1,\mathbf{p})} \Longrightarrow C(B^{-1}(\mathbf{p})) = c(\mathcal{T}_{(1,\mathbf{p})}) = \{p_0-1\}\times C(\mathbf{p})$$

We next consider the mapping χ

$$\forall \mathbf{p}\in\mathbf{V}_n \quad \forall \mathbf{f}\in\mathcal{F}_{\mathbf{p}}[\mathbf{q}] \qquad \chi(\mathbf{f}) = \mathbf{x}^{v(\mathbf{f})}\ \mathbf{y}^{c(\mathbf{f})} = (x_0^{p_0}x_1^{p_1}\ldots x_n^{p_n})\ \ (y_0^{q_0}y_1^{q_1}\ldots y_n^{q_{n-1}})$$

Using the Hilbert series expansions (3.5), we readily check that

$$\prod_{\mathbf{f}\in\mathcal{S}(n)} \frac{1}{1-\mathbf{x}^{v(\mathbf{f})}\ \mathbf{y}^{c(\mathbf{f})}} = \prod_{\mathbf{f}\in\mathcal{S}(n)} \left(\sum_{\mathbf{g}\in\langle\mathbf{f}\rangle} \mathbf{x}^{v(\mathbf{g})}\ \mathbf{y}^{c(\mathbf{g})}\right) = \sum_{\mathbf{g}\in\langle\mathcal{S}(n)\rangle} \mathbf{x}^{v(\mathbf{g})}\ \mathbf{y}^{c(\mathbf{g})}$$

Arguing as in the proof of proposition 3.11, we find that

$$\begin{aligned}\sum_{\mathbf{g}\in\langle\mathcal{S}(n)\rangle} \mathbf{x}^{v(\mathbf{g})}\ \mathbf{y}^{c(\mathbf{g})} &= 1+\sum_{\mathbf{p}\in\mathbf{V}_0}|\mathcal{F}_{\mathbf{p}}|\ \mathbf{x}^{\mathbf{p}} + \sum_{k=1}^{n}\sum_{\mathbf{p}\in\mathbf{V}_k}\sum_{\mathbf{q}\in C(\mathbf{p})}|\mathcal{F}_{\mathbf{p}}[\mathbf{q}]|\ \mathbf{x}^{\mathbf{p}}\ \mathbf{y}^{\mathbf{q}}\\ &= \sum_{\mathbf{p}\in\mathbf{V}^{\star}_n}\sum_{\mathbf{q}\in C(\mathbf{p})}|\mathcal{F}_{\mathbf{p}}[\mathbf{q}]|\ \mathbf{x}^{\mathbf{p}}\ \mathbf{y}^{\mathbf{q}}\end{aligned}$$

In much the same way, we have that

$$\prod_{\mathbf{f}\in\mathcal{S}(n)} \frac{1}{1-\mathbf{x}^{v(\mathbf{f})}\ \mathbf{y}^{c(\mathbf{f})}} \;=\; \frac{1}{1-x_0}\prod_{k=1}^{n}\prod_{\mathbf{p}\in\mathbf{V}_k:p_0=1}\prod_{\mathbf{q}\in C(\mathbf{p})}\left(\frac{1}{1-\mathbf{x}^{\mathbf{p}}\ \mathbf{y}^{\mathbf{q}}}\right)^{|\mathcal{F}_{\mathbf{p}}[\mathbf{q}]|}$$

Therefore,

$$\begin{aligned}&\textstyle\prod_{\mathbf{f}\in\mathcal{S}(n)}\frac{1}{1-\mathbf{x}^{v(\mathbf{f})}\ \mathbf{y}^{c(\mathbf{f})}}\\ &= \textstyle\prod_{k=0}^{n}\prod_{\mathbf{p}\in\mathbf{V}_k:p_0=1}\prod_{\mathbf{q}\in C(\mathbf{p})}\left(\frac{1}{1-x_0B(\mathbf{x})^{B(\mathbf{p})}\ y_0^{(p_1-1)}\ B(\mathbf{y})^{B(\mathbf{q})}}\right)^{\left|\mathcal{F}_{(1,B(\mathbf{p}))}[(p_1-1,B(\mathbf{q}))]\right|}\end{aligned}$$

This yields that

$$\begin{aligned}\prod_{\mathbf{f}\in\mathcal{S}(n)} \frac{1}{1-\mathbf{x}^{v(\mathbf{f})}\ \mathbf{y}^{c(\mathbf{f})}} &= \prod_{k=0}^{n}\prod_{\mathbf{p}\in\mathbf{V}_{k-1}}\prod_{\mathbf{q}\in C(B^{-1}(\mathbf{p}))}\left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}\ \mathbf{y}^{\mathbf{q}}}\right)^{|\mathcal{F}_{B^{-1}(\mathbf{p})}[\mathbf{q}]|}\\ &= \prod_{\mathbf{p}\in\mathbf{V}^{\star}_{n-1}}\prod_{\mathbf{q}\in C(B^{-1}(\mathbf{p}))}\left(\frac{1}{1-\mathbf{x}^{B^{-1}(\mathbf{p})}\ \mathbf{y}^{\mathbf{q}}}\right)^{|\mathcal{F}_{\mathbf{p}}[B(\mathbf{q})]|}\end{aligned}$$

The proposition follows. ∎

Notice that the Hilbert series technique can be developed to any order of refinement. For example, it could be used to take into account, besides the number of vertices or of coalescences at each level, the cardinals $\#(\mathbf{f})$, so that the coefficients of the Laurent expansion of the measures $\mathbb{Q}_{n,q}^N$ could be read, in the end, entirely on the corresponding Hilbert series. We leave the task of expressing the recursive formula to the interested reader, and simply point out that the technique allows an easy, systematic, recursive, computation of the coefficients of the expansion of the measures $\mathbb{Q}_{n,q}^N$ at any order, both in n and q. The observation can be usefull, especially in view of the systematic development of numerical schemes and numerical approximations based on Feynman-Kac particle models.

3.8 Some forests expansions

In the present section, and the forthcoming one, we take advantage of the langage of trees and of the results obtained on their statistics to compute the first orders of the Laurent functional representation of $\mathbb{Q}_{n,q}^N$ and, respectively, to derive a natural generalization to Feynman-Kac particle models of the classical Wick product formula.

Let us fix n and $q \geq 4$, so that the notation $\mathbf{q}$ denotes, once again, the constant sequence of lenght n associated to q. As we have already pointed out, there is only one forest in $\mathcal{F}_{\mathbf{q}}$ without coalescence, which is the product of q trivial trees of height n. There is only one forest in $\mathcal{F}_{\mathbf{q}}$ with only one coalescence, at level i, that will be written $\mathbf{f}_{1,i}$. Its only non trivial tree is the tree with one coalescence at level i and two leaves at level n. There are two forests with coalescence degree 2 and the two coalescences at levels i, written $\mathbf{f}_{2,i}^1$ and $\mathbf{f}_{2,i}^2$. The notation $\mathbf{f}_{2,i}^1$ denotes the forest with only one non trivial tree with coalescence degree 2, a vertex with coalescence order 2 at level i and its three leaves at level n. The notation $\mathbf{f}_{2,i}^2$ denotes the forest with two non trivial trees with coalescence degree 1 and the coalescence at level i. There are four forests with coalescence degree 2 and the two coalescences at levels $i < j$, written $\mathbf{f}_{2,i,j}^k$, $k = 1...4$. The forest $\mathbf{f}_{2,i,j}^1$ has one non trivial tree with coalescences at levels i and j and its tree leaves at level n. The forest $\mathbf{f}_{2,i,j}^2$ has one non trivial tree with coalescences at levels i and j and its three leaves at the levels j, n, n. The forest $\mathbf{f}_{2,i,j}^3$ has two non trivial trees with one coalescence at level i, resp. j and their two leaves at level n. The forest $\mathbf{f}_{2,i,j}^4$ has two non trivial trees with one coalescence at level i, resp. j and their two respective leaves at level j, n, resp. n, n.

Expanding the formulae for $\partial^i \mathbb{Q}_{q,n}$ and using the formulae obtained in Thm 3.8 for the cardinals $\#(\mathbf{f})$ and using that $s(q, q-2) = \left(\begin{array}{c} q \\ 3 \end{array} \right) \frac{3q-1}{4}$ (see for instance [2], on page 63), we get the following result.

Corollary 3.14. *The first three order terms in the polynomial functional representation of $\mathbb{Q}_{n,q}^N$ (2.2) are given by the following formulae*

$$
\begin{aligned}
\partial^0 \mathbb{Q}_{n,q} &= \gamma_n^{\otimes q} \\
\partial^1 \mathbb{Q}_{n,q} &= \frac{q(q-1)}{2} \sum_{0 \leq k \leq n} \left(\Delta_{n,q}^{\mathbf{f}_{1,k}} - \gamma_n^{\otimes q} \right) \\
\partial^2 \mathbb{Q}_{n,q} &= \frac{q!}{(q-3)!\ 3!} \sum_{0 \leq k \leq n} \left(\Delta_{n,q}^{\mathbf{f}_{2,k}^1} + \frac{3}{4}\ (q-3)\ \Delta_{n,q}^{\mathbf{f}_{2,k}^2} - \frac{3}{2}\ (q-1)\ \Delta_{n,q}^{\mathbf{f}_{1,k}} + \frac{(3q-1)}{4}\ \gamma_n^{\otimes q} \right) \\
&\quad + \left(\frac{q(q-1)}{2} \right)^2 \sum_{0 \leq k < l \leq n} \left(\gamma_n^{\otimes q} - \left(\Delta_{n,q}^{\mathbf{f}_{1,l}} + \Delta_{n,q}^{\mathbf{f}_{1,k}} \right) \right) \\
&\quad + \frac{q(q-1)}{2} \sum_{0 \leq k < l \leq n} \left(\Delta_{n,q}^{\mathbf{f}_{2,k,l}^2} + (q-2)\ (\Delta_{n,q}^{\mathbf{f}_{2,k,l}^1} + \Delta_{n,q}^{\mathbf{f}_{2,k,l}^4}) + \frac{(q-2)(q-3)}{2} \Delta_{n,q}^{\mathbf{f}_{2,k,l}^3} \right)
\end{aligned}
$$

3.9 A Wick product formula on forests

Let now $\mathcal{B}_0^{sym}(E_n^q) \subset \mathcal{B}_b(E_n^q)$ be the set of symmetric functions F on E_n^q such that

$$
(D_{1_{q-1}} \otimes \gamma_n)(F)(x_1, ..., x_{q-1}) = \int F(x_1, \ldots, x_{q-1}, x_n)\ \gamma_n(dx_n) = 0
$$

Notice that $\mathcal{B}_0^{sym}(E_n^q)$ contains the set of functions $F = (P)_{\rm sym}$, with $P \in {\rm Poly}(E_n^q)$, where ${\rm Poly}(E_n^q) \subset \mathcal{B}_b(E_n^q)$ stands for the subset of polynomial functions of the form

$$P = \sum_{a \in I} c(a)\ f^a \quad \text{with} \quad f^a = (f^{a(1)} \otimes \ldots \otimes f^{a(q)})$$

In the above display, I is a finite subset of $\mathbb{N}^{[q]}$, $c \in \mathbb{R}^I$, and the elementary functions $f^p \in \mathcal{B}_b(E_n)$ are choosen such that $\gamma_n(f^p) = 0$. For instance, we can take

$$f^p = (g^p - \eta_n(g^p)) \quad \text{with} \quad g^p \in \mathcal{B}_b(E_n)$$

Let us recall that, when explicited in the particle models setting, the traditional Wick product formula reads, for any integer q and any $F \in \mathcal{B}_0^{sym}(E_0^q)$

$$\partial^i \mathbb{Q}_{0,q}(F) = 0,\ i < \frac{q}{2}$$

and, if q is even,

$$\partial^{\frac{q}{2}} \mathbb{Q}_{0,q}(F) = \frac{q!}{2^{q/2}\ (q/2)!}\ \Delta_{0,q}^{\mathbf{f}}(F)$$

where $\mathbf{f}$ is the forest in $\mathcal{F}_{0,q}$ containing $\frac{q}{2}$ copies of the tree of unit height, with two vertices at level 1, and $\frac{q}{2}$ copies of the tree with the root as unique vertex. For symmetric tensor product functions $F = (f^1 \otimes \ldots \otimes f^q)_{\rm sym}$, associated with a collection of functions $f^i \in \mathcal{B}_b(E_0)$, such that $\eta_0(f^i) = \gamma_0(f^i) = 0$, for any $1 \leq i \leq q$, we readily find that

$$\Delta_{0,q}^{\mathbf{f}}(F) = \frac{2^{q/2}\ (q/2)!}{q!} \sum_{\mathcal{I}_q} \left(\prod_{\{i,j\} \in \mathcal{I}_q} \eta_0(f^i f^j) \right)$$

as soon as q is even. In the above displayed formula, $\mathcal{I}_q$ ranges over all partitions of $[q]$ into pairs. In a more probabilistic language, the above formula can be interpreted as the q-th order central moment

$$\partial^{\frac{q}{2}} \mathbb{Q}_{0,q}(F) = \mathbb{E}(W_0(f^1) \ldots W_0(f^q)) = \sum_{\mathcal{I}_q} \left(\prod_{\{i,j\} \in \mathcal{I}_q} \mathbb{E}(W_0(f^i) W_0(f^j)) \right)$$

of a Gaussian field W_0 on the Banach space of functions $\mathcal{B}_b(E_0)$, such that for any pair of functions $(\varphi, \psi) \in \mathcal{B}_b(E_0)$, $\mathbb{E}(W_0(\varphi)) = 0$, and $\mathbb{E}(W_0(\varphi) W_0(\psi)) = \eta_0(\varphi\psi)$. In terms of the Laplace moment generating function

$$\mathbf{x} = (x^i)_{1 \leq i \leq q} \mapsto \exp\left(\frac{1}{2} \mathbf{x}' C_0(f) \mathbf{x} \right) = \frac{1}{((2\pi)^q |C_0(f)|)^{1/2}} \int_{\mathbb{R}^q} d\mathbf{y} \exp\left(-\frac{1}{2} \mathbf{y}' C_0(f)^{-1} \mathbf{y} + \mathbf{y}' \mathbf{x} \right)$$

with the gaussian covariance matrix $C_0(f) = (C_0(f^i, f^j))_{1 \leq i,j \leq q}$, we also have that

$$\partial^{\frac{q}{2}} \mathbb{Q}_{0,q}(F) = \frac{\partial^q}{\partial \mathbf{x}^q} \exp\left(\frac{1}{2} \mathbf{x}' C_0(f) \mathbf{x} \right)_{|\mathbf{x}=\mathbf{0}} \quad \text{with} \quad \frac{\partial^q}{\partial \mathbf{x}^q} := \frac{\partial}{\partial x^1} \cdots \frac{\partial}{\partial x^q}$$

In the present section, we will show that the Wick formula generalizes to forests in $\mathcal{F}_{n,q}$ of arbitrary height. Let us start by listing various straightforward properties of trees and forests. A tree T with coalescence degree d has exactly $(d+1)$ leaves. A forest with coalescence degree d has at most d non trivial trees, and the equality holds if and only if all its non trivial trees have coalescence degree 1. In particular, if a forest in $\mathcal{F}_{n,q}$ has coalescence degree d, it has at most $(2d)$ leaves belonging to non trivial trees so that, if $d < \frac{q}{2}$, there is at least one vertex at level $n+1$ belonging to a trivial tree (that is, a tree with coalescence number 0).

The same reasoning shows that, when $d = \frac{q}{2}$, a forest in $\mathcal{F}_{n,q}$ with coalescence degree d, and coalescence sequence $\mathbf{r}$, does not contain a trivial tree of height $(n+1)$ if and only if it is the forest $\mathbf{f}_{\mathbf{r}} := T_0^{r_0} U_0^{r_0} \ldots T_n^{r_n} U_n^{r_n}$, where we write T_k for the unique tree of coalescence degree 1 with a coalescence at level k and its two leaves at level $(n+1)$, and where we write U_k for the trivial tree of height k.

We conclude this series of remarks by noting that, if $\mathbf{f} \in \mathcal{F}_{n,q}$ can be written as the product (or disjoint union) of a forest $\mathbf{g}$ in $\mathcal{F}_{n,q-1}$ with U_{n+1}, the trivial tree of height $(n+1)$, then, for any $F \in \mathcal{B}_b^{sym}(E_n^q)$ we have, by definition of the measures $\Delta^{\mathbf{f}}$

$$\Delta_{n,q}^{\mathbf{f}}(F) = \Delta_{n,q}^{\mathbf{g}}(D_{1_{q-1}} \otimes \gamma_n)(F)$$

which is equal to 0 if $F \in \mathcal{B}_0^{sym}(E_n^q)$.

We are now in position to derive the forest Wick formula.

Theorem 3.15. *For any even integer $q \leq N$ and any symmetric function $F \in \mathcal{B}_0^{sym}(E_n^q)$, we have*

$$\forall k < q/2 \qquad \partial^k \mathbb{Q}_{n,q}(F) = 0 \quad \text{and} \quad \partial^{q/2}\mathbb{Q}_{n,q}(F) = \sum_{\mathbf{r}<\mathbf{q},|\mathbf{r}|=\frac{q}{2}} \frac{q!}{2^{q/2}\ \mathbf{r}!}\ \Delta_{n,q}^{\mathbf{f_r}} F \tag{3.7}$$

For odd integers $q \leq N$, the partial measure valued derivatives ∂^k are the null measure on $\mathcal{B}_0^{sym}(E_n^q)$, up to any order $k \leq \lfloor q/2 \rfloor$.

We close this section with a gaussian field interpretation of the Wick formula (3.7). We further assume that q is an even integer. We consider a collection of independent gaussian fields $(W_k)_{0\leq k\leq n}$ on the Banach spaces $(\mathcal{B}_b(E_k))_{0\leq k\leq n}$, with for any $(\varphi_k, \psi_k) \in \mathcal{B}_b(E_k)^2$, and $0 \leq k \leq n$

$$\mathbb{E}(W_k(\varphi_k)) = 0 \quad \text{and} \quad \mathbb{E}(W_k(\varphi_k)W_k(\psi_k)) = \gamma_k(\varphi_k\psi_k)$$

We also introduce the centered gaussian field V_n on $\mathcal{B}_b(E_n)$ defined for any $\varphi_n \in \mathcal{B}_b(E_n)$ by the following formula

$$V_n(\varphi_n) = \sum_{0\leq k\leq n} \sqrt{\gamma_k(1)}\ W_k(Q_{k,n}(\varphi_n))$$

Let $(\varphi_i)_{1\leq i\leq q} \in \mathcal{B}_b(E_n)^q$ be a collection of functions such that $\gamma_n(\varphi_i) = 0$, for any $1 \leq i \leq q$. For the tensor product function $F = \frac{1}{q!}\sum_{\sigma\in\mathcal{G}_q}(\varphi_{\sigma(1)} \otimes \ldots \otimes \varphi_{\sigma(q)})$, one can check that

$$\Delta_{n,q}^{\mathbf{f_r}}(F) \quad = \quad \frac{2^{q/2}\mathbf{r}!}{q!}\sum_{I\in\mathcal{I}}\prod_{0\leq k\leq n}\sum_{J_k\in\mathcal{I}_k}\left\{\gamma_k(1)^{r_k}\prod_{(i,j)\in J_k}\gamma_k(Q_{k,n}(\varphi_i)Q_{k,n}(\varphi_j))\right\}$$

In the above displayed formula, the first sum is over the set $\mathcal{I}$ of all partitions $I = (I_k)_{0\leq k\leq n}$ of $[q]$ into $(n+1)$ blocks with cardinality $|I_k| = (2r_k)$, and the second sum ranges over the set $\mathcal{I}_k$ of all partitions J_k of the sets I_k into pairs, with $0 \leq k \leq n$. By definition of the gaussian fields $(W_k)_{0\leq k\leq n}$, and due to the classical Wick formula, we find that

$$\begin{aligned}\frac{q!}{2^{q/2}\mathbf{r}!}\Delta_{n,q}^{\mathbf{f_r}}(F_n) &= \sum_{I\in\mathcal{I}}\prod_{0\leq k\leq n}\mathbb{E}\left(\prod_{i\in I_k}\sqrt{\gamma_k(1)}\ W_k(Q_{k,n}(\varphi_i))\right)\\ &= \mathbb{E}\left(\sum_{I\in\mathcal{I}}\prod_{0\leq k\leq n}\prod_{i\in I_k}\sqrt{\gamma_k(1)}\ W_k(Q_{k,n}(\varphi_i))\right)\end{aligned}$$

from which we arrive at the following formula

$$\sum_{\mathbf{r}<\mathbf{q},|\mathbf{r}|=\frac{q}{2}}\frac{q!}{2^{q/2}\ \mathbf{r}!}\ \Delta_{n,q}^{\mathbf{f_r}}(F_n) = \mathbb{E}\left(\sum_{\mathbf{r}:|2\mathbf{r}|=q}\sum_{I}\prod_{0\leq k\leq n}\prod_{i\in I_k}\sqrt{\gamma_k(1)}\ W_k(Q_{k,n}(\varphi_n^i))\right)$$

Recalling that all gaussian fields $(W_k)_{0\leq k\leq n}$ are independent and centered, we prove that

$$\partial^{q/2}\mathbb{Q}_{n,q}(F_n) = \mathbb{E}\left(\prod_{1\leq i\leq q}\left(\sum_{0\leq k\leq n}\sqrt{\gamma_k(1)}\ W_k(Q_{k,n}(\varphi_n^i))\right)\right) = \mathbb{E}\left(\prod_{1\leq i\leq q}V_n(\varphi_n^i)\right)$$

Written in a more synthetic way, we have proved the following formula

$$\partial^{q/2}\mathbb{Q}_{n,q}(F_n) = \mathbb{E}\left(V_n^{\otimes q}(F_n)\right)$$

This result can alternatively be derived combining the $\mathbb{L}_q$-mean error estimates presented in [4, Thm.7.4.2], with the multidimensional central limit theorems presented in [4, Prop.9.4.1]. More precisely, the q-dimensional particle random fields $\left(V_n^N(\varphi_n^i)\right)_{1\leq i\leq q} := \left(\sqrt{N}\gamma_n^N(\varphi_n^i)\right)_{1\leq i\leq q}$, converge in law to $(V_n(\varphi_n^i))_{1\leq i\leq q}$. By the continuous mapping theorem, combined with simple uniform integrability arguments, one checks that

$$N^{q/2}\ \mathbb{Q}_{n,q}^N(F_n) = \mathbb{E}\left((V_n^N)^{\otimes q}(F_n)\right) \overset{N\uparrow\infty}{\longrightarrow} \partial^{q/2}\mathbb{Q}_{n,q}(F_n) = \mathbb{E}\left(V_n^{\otimes q}(F_n)\right)$$

4 Extension to path-space models

In the present section, we extend our previous analysis to the statistical study of path spaces. Due to the mean-field definition of the trajectories, it will appear soon that distributions on the path space are parametrized by forests with a slightly more complex structure than the classical ones. Namely, we have to introduce colored trees and forests (black and white, with a particular structure deviced to reflect the geometry of paths). These new objects are introduced and studied in the first two subsections; applications to the path space and to propagation of chaos properties are postponed to the last three sections.

4.1 Colored trees, forests, and jungles

Definition 4.1. *A colored tree is a tree with colored vertices, with two distinguished colors, say black and white. Only black vertices may have an ingoing edge. That is, equivalently, all white vertices are leaves -notice that the converse is not true in general. A colored forest is a multiset of colored trees. The sets of colored trees and colored forests are denoted respectively by* $\overline{\mathcal{T}}$ *and* $\overline{\mathcal{F}}$.

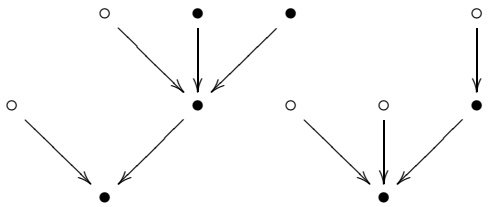

Figure 3: A colored forest

Notice that a colored forest can also be viewed as a commutative monomial over the set of colored trees. In this interpretation, the generating series techniques that we have developped to deal with the enumeration of forests will apply to colored forests. The computation of the corresponding series is left to the interested reader.

Most of the notions associated to trees and forests go over in a straightforward way to colored forests and colored trees. As a general rule, we will write a line over symbols associated to colored trees and colored forests. For instance, we write $\overline{v}_k(\overline{T}) = (v_k(\overline{T}), v'_k(\overline{T}))$ for the number $v_k(\overline{T})$, resp. $v'_k(\overline{T})$, of white, resp. black vertices in the colored tree $\overline{T}$, at each level $k \geq 0$.

We also let $\overline{\mathbf{V}}$ be the set of all sequences of pair of integers $\overline{\mathbf{p}} = (\overline{p}_k)_{k\geq 0} \in (\mathbb{N}^2)^{\mathbb{N}}$, with $\overline{p}_k = (p_k, p'_k)$ for every $k \geq 0$, satisfying the following property

$$\exists \overline{ht}(\overline{\mathbf{p}}) \in \mathbb{N} \quad \text{s.t.} \quad \inf_{k\leq \overline{ht}(\overline{\mathbf{p}})} p_k + p'_k > 0 \quad \text{and} \quad \sup_{k>\overline{ht}(\overline{\mathbf{p}})} p_k + p'_k = 0$$

For any $n \in \mathbb{N}$, we denote by $\overline{\mathbf{V}}_n \subset \overline{\mathbf{V}}$ the subset of sequences $\overline{\mathbf{p}}$ such that the height of $\overline{\mathbf{p}}$, $\overline{ht}(\overline{\mathbf{p}})$, is equal to n. Finally, we let $\overline{\mathcal{T}}_{\overline{\mathbf{p}}}$ be the set of colored trees $\overline{T}$ with $v_k(\overline{T}) = p_k$ white vertices, and $v'_k(\overline{T}) = p'_k$ black vertices, at each level $k \geq 0$. Since $\overline{T}$ is a colored tree, this implies that $p_0 = 0$ and $p'_0 = 1$, excepted if $\overline{ht}(\overline{\mathbf{p}}) = 0$. In that case, p_0 may also be equal to 1 (and then $p'_0 = 0$). Notice that the set of black and white vertices at each level k is in bijection with the disjoint union

$$[\overline{p}_k] = [(p_k, p'_k)] := [p_k] \amalg [p'_k]$$

of the sets $[p_k]$ and $[p'_k]$.

For $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$, let us call, by analogy with the uncolored case, planar colored trees $\overline{T}$ of type $\overline{\mathbf{p}}$ any sequence of maps $(\overline{\alpha}_0, \ldots, \overline{\alpha}_n)$, where

$$\overline{\alpha}_k = (\alpha_k \amalg \alpha'_k) \in [p'_k]^{[p_{k+1}]} \amalg [p'_k]^{[p'_{k+1}]} =: [p'_k]^{[(p_{k+1}, p'_{k+1})]}$$

is the disjoint union of the pair of weakly increasing maps $\alpha_k \in [p'_k]^{[p_{k+1}]}$, and $\alpha_k \in [p'_k]^{[p'_{k+1}]}$.

In the above definition, the parameter level k runs from 0, to n, and we have used the conventions $(\emptyset \amalg \emptyset) = \emptyset$, and $(\alpha \amalg \emptyset) = \alpha$. Writing $\mathcal{P}\overline{\mathcal{T}}$ for the set of planar colored trees and $\mathcal{P}\overline{\mathcal{T}}_{\overline{\mathbf{p}}}$ for the set of planar colored trees of type $\overline{\mathbf{p}}$, from the previous discussion, for any $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$ with $p_0 = 0,\ p'_0 = 1$, we have that

$$\begin{aligned} &\mathcal{P}\overline{\mathcal{T}}_{\overline{\mathbf{p}}} \\ &\simeq \left\{(\overline{\alpha}_0, \ldots, \overline{\alpha}_n) \in [1]^{[\overline{p}_1]} \times [p'_1]^{[\overline{p}_2]} \times \ldots \times [p'_n]^{[\overline{p}_{n+1}]} \ : \ \overline{\alpha}_k = (\alpha_k \amalg \alpha'_k) \text{ with } \alpha_k \uparrow \text{ and } \alpha'_k \uparrow \right\} \end{aligned}$$

A planar colored forest is, by definition, a sequence (or a non commutative monomial) of planar colored trees.

Since to any sequence $(\overline{\alpha}_k^1, ..., \overline{\alpha}_k^n)$, $\overline{\alpha}_k^i = (\alpha_k^i \coprod \alpha_k^{i'}) \in [p_k^{i'}]^{[(p_{k+1}^i, p_{k+1}^{i'})]}$ is associated a unique element $(\alpha_k^1 \coprod ... \coprod \alpha_k^n; \alpha_k^{1'} \coprod ... \coprod \alpha_k^{n'})$ of $[p_k^{1'}] \coprod ... \coprod [p_k^{n'}]^{(\coprod_{i=1}^n [p_{k+1}^i]) \coprod (\coprod_{i=1}^n [p_{k+1}^{i'}])}$, planar colored forests are in bijection with sequences of disjoint union of weakly increasing maps, in the sense that (with self-explaining notations)

$$\mathcal{P}\overline{\mathcal{F}}_{\overline{\mathbf{p}}} \simeq \mathcal{I}\overline{\mathcal{A}}_{\overline{\mathbf{p}}} := \left\{(\overline{a}_0, \ldots, \overline{a}_n) \in \overline{\mathcal{A}}_{\overline{\mathbf{p}}} \ : \ \overline{a}_k = (a_k \amalg a'_k) \text{ with } a_k \uparrow \text{ and } a'_k \uparrow\right\}$$

for any $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$, with

$$\overline{\mathcal{A}}_{\overline{\mathbf{p}}} := \left([p'_0]^{[\overline{p}_1]} \times [p'_1]^{[\overline{p}_2]} \times \ldots \times [p'_n]^{[\overline{p}_{n+1}]}\right)$$

Definition 4.2. *We associate with a sequence of mappings $\overline{\mathbf{a}} = (\overline{a}_0, \ldots, \overline{a}_n) \in \overline{\mathcal{A}}_{\overline{\mathbf{p}}}$, with $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$, a colored forest $\overline{F}(\overline{a})$ of height $(n+1)$. The vertex set of the colored forest $\overline{F}(\overline{a})$ is given by the disjoint union set $\left(\coprod_{k=0}^{n+1}[p_k]\right) \coprod \left(\coprod_{k=0}^{n+1}[p'_k]\right)$. The edges correspond to the pairs $(i, a_k(i))$, and $(i', a'_k(i'))$, with the pair of white and black colored vertices $(i, i') \in \left([p_{k+1}] \coprod [p'_{k+1}]\right)$. The sequence can be represented graphically by an entangled planar graph with colored vertices, or colored jungle , written $\overline{J}(\overline{a})$. Notice that edges between two levels are allowed to cross. At a given level k, all white vertices are drawn left to the black vertices. The set of colored jungles is written $\overline{\mathcal{J}}$.*

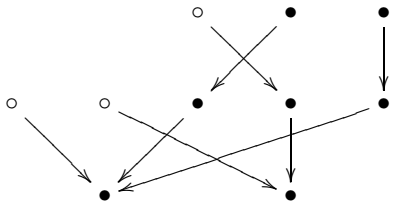

Figure 3: A colored jungle

Once again, the notation we use on colored forests extends to colored jungles in a self-explanatory way. For instance, $\overline{\mathcal{J}}_{\overline{\mathbf{p}}}$, $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$, stands for the set of colored jungles $\overline{\mathbf{j}}$ with $v_k(\overline{\mathbf{j}}) = p_k$ white vertices, and $v'_k(\overline{\mathbf{j}}) = p'_k$ black vertices, at each level $0 \leq k \leq (n+1)$. For any $\overline{\mathbf{p}} \in \overline{\mathbf{V}}_{n+1}$, we also have that $\overline{\mathcal{J}}_{\overline{\mathbf{p}}} \simeq \overline{\mathcal{A}}_{\overline{\mathbf{p}}}$.

Our next objective is to make more precise the algebraic picture of these objects. For any integers p, p', q, q', r, s, and any disjoint maps

$$\overline{a} = (a \amalg a') \in [r]^{[(p,p')]} \quad \text{and} \quad \overline{b} = (b \amalg b') \in [s]^{[(q,q')]}$$

we let $(\overline{a} \vee \overline{b})$ be the disjoint union of maps

$$(a \amalg a') \vee (b \amalg b') = (a \vee b) \amalg (a' \vee b') \in [r+s]^{[(p+p', q+q')]}$$

Notice that the property of being weakly increasing is preserved by this construction. More generally, for any $\overline{\mathbf{p}} \in \overline{\mathbf{V}}$, and $\overline{\mathbf{q}} \in \overline{\mathbf{V}}$, and any sequences $\overline{\mathbf{a}} = (\overline{a}_k)_{k \geq 0} \in \overline{\mathcal{A}}_{\overline{\mathbf{p}}}$, and $\overline{\mathbf{b}} = (\overline{b}_k)_{k \geq 0} \in \overline{\mathcal{A}}_{\overline{\mathbf{q}}}$, we set

$$\overline{\mathbf{a}} \vee \overline{\mathbf{b}} = (\overline{a}_k \vee \overline{b}_k)_{k \geq 0} \in \overline{\mathcal{A}}_{\overline{\mathbf{p}} + \overline{\mathbf{q}}} \tag{4.1}$$

In the above displayed formula, we have used the conventions $\overline{\mathbf{p}} + \overline{\mathbf{q}} = (p_k + q_k, p'_k + q'_k)_{k \geq 0}$, $\overline{a}_k = \emptyset$, and $\overline{b}_l = \emptyset$, for any $k > \overline{ht}(\overline{\mathbf{p}})$, and $l > \overline{ht}(\overline{\mathbf{q}})$. Notice that, if $\overline{\mathbf{a}} \in \mathcal{I}\overline{\mathcal{A}}_{\overline{\mathbf{p}}}$ and $\overline{\mathbf{b}} \in \mathcal{I}\overline{\mathcal{A}}_{\overline{\mathbf{q}}}$, then $\overline{\mathbf{a}} \vee \overline{\mathbf{b}} \in \mathcal{I}\overline{\mathcal{A}}_{\overline{\mathbf{p}} + \overline{\mathbf{q}}}$.

The main advantage of this algebraic structure lies in the following non commutative composition formula in the set of colored jungles

$$\overline{J}(\overline{\mathbf{a}})\overline{J}(\overline{\mathbf{b}}) = \overline{J}(\overline{\mathbf{a}} \vee \overline{\mathbf{b}}) \in \overline{\mathcal{J}}_{\overline{\mathbf{p}} + \overline{\mathbf{q}}}$$

Notice that any pair of permutations $(s, s') \in (\mathcal{G}_p \times \mathcal{G}_{p'})$ can be seen as a pair of mappings $(s, s') \in \left([p+p']^{[p]} \times [p+p']^{[p']}\right)$. In this interpretation, the disjoint union map

$$\overline{s} = (s \amalg s') \in [p+p']^{[(p,p')]}$$

can be seen as a permutation map of the set $[(p,p')] = ([p] \coprod [p'])$ that preserves the sets $[p]$, and $[p']$. The reverse assertion is also true. The set of these permutations is denoted by

$$\overline{\mathcal{G}}_{(p,p')} := \mathcal{G}_p \coprod \mathcal{G}_{p'}$$

We end this section with a technical lemma that permits to extend in a natural way the jungle automorphism group analysis developed in section 3 to their colored analogue.

Lemma 4.3. *For any collection of integers $p, p', q, q', r, r', s, s'$, for and sequence of permutation mappings $\overline{\rho} \in \overline{\mathcal{G}}_{(p,p')}$, $\overline{\theta} \in \overline{\mathcal{G}}_{(q,q')}$, $\overline{\sigma} \in \overline{\mathcal{G}}_{(r,r')}$, $\overline{\tau} \in \overline{\mathcal{G}}_{(s,s')}$, and for any pair of disjoint mappings $\overline{a} \in [r']^{[(p,p')]}$, and $\overline{b} \in [s']^{[(q,q')]}$ we have the composition formula*

$$(\overline{\sigma} \vee \overline{\tau}) \circ \left(\overline{a} \vee \overline{b}\right) \circ \left(\overline{\rho} \vee \overline{\theta}\right) = \left(\overline{\sigma} \circ \overline{a} \circ \overline{\rho}\right) \vee \left(\overline{\tau} \circ \overline{b} \circ \overline{\theta}\right)$$

In addition, for any $\overline{\lambda} \in \overline{\mathcal{G}}_{(p+q),(p'+q')}$ we have that

$$\left(\overline{a} \vee \overline{b}\right) \circ \overline{\lambda} \in \left([r']^{[(p,p')]} \vee [s']^{[(q,q')]}\right) \Longrightarrow \overline{\lambda} \in \left(\overline{\mathcal{G}}_{(p,p')} \vee \overline{\mathcal{G}}_{(q,q')}\right)$$

The two assertions are a direct consequence of the definition of the $\vee$ product. In the first equation above, it should be understood that $\overline{a} \in [r']^{[(p,p')]}$ is viewed as a map in $\overline{a} \in ([r] \coprod [r'])^{[(p,p')]}$ through the embedding of $[r']$ into $[r] \coprod [r']$. The same observation will be used without further notice in the forthcoming paragraph.

4.2 Automorphism groups on colored jungles

As for the expansion of measures in section 2, we will need a definition of equivalence classes of colored jungles, a parametrization of the classes and a computation of their cardinals. Thanks to the algebraic model of colored trees designed in section 4.1, most of the constructions in section 2, and section 3 can be extended without further work to colored jungles. Since all constructions can be mimicked, we only outline the main ideas of the generalization from jungles to colored jungles.

We let $\overline{\mathbf{p}} \in \overline{V}_{n+1}$ be a sequence of $(n+1)$ integer pairs $\overline{p}_k = (p_k, p'_k)$, with $0 \leq k \leq (n+1)$. We associate with $\overline{\mathbf{p}}$ the product permutation group

$$\overline{\mathcal{G}}_{\overline{\mathbf{p}}} = \overline{\mathcal{G}}_{\overline{p}_0} \times \overline{\mathcal{G}}_{\overline{p}_1} \times \ldots \times \overline{\mathcal{G}}_{\overline{p}_{n+1}}$$

This group acts naturally on sequences of maps $\overline{\mathbf{a}} = (\overline{a}_0, \overline{a}_1, ..., \overline{a}_n) \in \overline{\mathcal{A}}_{\overline{\mathbf{p}}}$, and on jungles $\overline{J}(\overline{\mathbf{a}})$ in $\overline{\mathcal{J}}_{\overline{\mathbf{p}}}$ by permutation of the colored vertices at each level. More formally, for any $\overline{\mathbf{s}} = (\overline{s}_0, ..., \overline{s}_{n+1}) \in \overline{\mathcal{G}}_{\overline{\mathbf{p}}}$ this pair of actions is given by

$$\overline{\mathbf{s}}(\overline{\mathbf{a}}) := (\overline{s}_0 \overline{a}_0 \overline{s}_1^{-1}, ..., \overline{s}_n \overline{a}_n \overline{s}_{n+1}^{-1}) \quad \text{and} \quad \overline{\mathbf{s}}\overline{J}(\overline{\mathbf{a}}) := \overline{J}(\overline{\mathbf{s}}(\overline{\mathbf{a}}))$$

Two colored jungles in the same orbit under the action of $\overline{\mathcal{G}}_{\overline{\mathbf{p}}}$ have the same underlying colored forest. Conversely, if two colored jungles in $\overline{\mathcal{J}}_{\overline{\mathbf{p}}}$ have the same underlying colored forests, they differ only by a permutation of the vertices of their colored graphs that preserves the colors of the vertices, and therefore are in the same orbit under the action of $\overline{\mathcal{G}}_{\overline{\mathbf{p}}}$. In other terms, we have the following lemma.

Lemma 4.4. *Equivalence classes of colored jungles in $\overline{\mathcal{J}}_{\overline{\mathbf{p}}}$ under the action of the permutation groups $\overline{\mathcal{G}}_{\overline{\mathbf{p}}}$ are in bijection with colored forests in $\overline{\mathcal{F}}_{\overline{\mathbf{p}}}$*

Notice the following corollary of the lemma: if $\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{p}}}$, with $\overline{\mathbf{p}} \in \overline{V}_{n+1}$, is a colored forest, then we can define unambiguously

$$|\overline{\mathbf{f}}| = (|\overline{\mathbf{f}}|_k)_{k \geq 0} = (|\overline{a}_0|, ..., |\overline{a}_n|, (0,0), (0,0), \ldots)$$

for any choice $(\overline{a}_0, ..., \overline{a}_n)$ of a representative of $\overline{\mathbf{f}}$ in $\overline{\mathcal{J}}_{\overline{\mathbf{p}}}$.

As for usual trees and forests, we write B for the map from colored trees to colored forests defined by removing the root of a colored tree, and we write also, as in section 3.3 , B for the induced map from the set of colored forests into itself.

Let us conclude this section by enumerating the number $\#(\overline{\mathbf{f}})$ of colored jungles associated to a given colored forest $\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{p}}}$, with $\overline{\mathbf{p}} \in \overline{V}_{n+1}$. The process is as in section 3.3, and the result follows ultimately from the class formula

$$\#(\overline{\mathbf{f}}) = \frac{\overline{\mathbf{p}}!}{|Stab(\overline{\mathbf{f}})|}$$

In the above displayed formula, $\overline{\mathbf{p}}!$ stands for the multi-index factorial $\overline{\mathbf{p}}! = \left(p_0!p'_0!...p_{n+1}!p'_{n+1}!\right)$, and we have written abusively $|Stab(\overline{\mathbf{f}})|$, for the cardinal of the stabilizer in $\overline{\mathcal{G}}_{\overline{\mathbf{p}}}$ of any representative $\overline{\mathbf{f}}'$ of $\overline{\mathbf{f}}$, where $\overline{\mathbf{f}}$ is viewed as an equivalence class of colored jungles.

Let us assume that $\overline{\mathbf{f}}$ can written, as a monomial over the set of colored trees as

$$\overline{\mathbf{f}} = \prod_{i=1}^{k} \overline{T_i}^{m_i}$$

where the T_is are pairwise distincts and $m_1 \geq ... \geq m_k$. We write $\overline{T_i}'$ for a set of representatives of the $\overline{T_i}$ viewed as equivalence classes of colored jungles.

As in section 3.4, we shall write $S(B^{-1}(\overline{\mathbf{f}}))$ for the unordered k-uplet $(m_1, ..., m_k)$, called the symmetry multiset of the colored tree $B^{-1}(\overline{\mathbf{f}})$, and extend the notation to colored forests so that $\mathbf{s}(\overline{\mathbf{f}})$ is the disjoint union of the $S(\overline{T_i})$. The proof in section 3.4 goes then over without changes, excepted for the replacement of forests, planar forests and planar trees by their colored analogues. For instance, we have the recursion formula

$$|Stab(\overline{\mathbf{f}})| = \prod_{i=1}^{k} \left(m_i! \ |Stab\ B(\overline{T'_i})|^{m_i} \right)$$

Expliciting the recursion in the Proposition, and following the wreath product analysis designed in section 3.6, we deduce the following theorem.

Theorem 4.5. *The number of colored jungles with underlying colored forest $\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{p}}}$, with $\overline{\mathbf{p}} \in \overline{V}_{n+1}$, is given by*

$$\#(\overline{\mathbf{f}}) = \frac{\overline{\mathbf{p}}!}{\prod_{i=-1}^{n} \mathbf{s}(B^i(\overline{\mathbf{f}}))!}$$

where we use the usual multi-index notation to define $\mathbf{s}(B^i(\overline{\mathbf{f}}))!$. In addition, we have the wreath product representation formula

$$Stab(\overline{\mathbf{f}}') \sim \mathcal{G}_{m_1} < Stab\ B(\overline{T}'_1) > \vee \ldots \vee \mathcal{G}_{m_k} < Stab\ B(\overline{T}'_k) >$$

4.3 Feynman-Kac semigroups

In this section, the parameter $n \geq 0$ represents a fixed time horizon. For any sequence of integers $\mathbf{q} = (q_0, \ldots, q_n) \in \mathbb{N}^{n+1}$, and any $-1 \leq m \leq n$, we set

$$q'_m := q_{m+1} + q'_{m+1} = \sum_{m<k\leq n} q_k \tag{4.2}$$

Notice that $q'_m = |[q_{m+1}]_{\amalg}[q'_{m+1}]|$. We associate with $\mathbf{q}$ the unnormalized Feynman-Kac measures

$$\gamma_n^{\mathbf{q}} = \gamma_0^{\otimes q_0} \otimes \cdots \otimes \gamma_n^{\otimes q_n} \in \mathcal{M}(\mathbf{E}_n^{\mathbf{q}}) \quad \text{where} \quad \mathbf{E}_n^{\mathbf{q}} = E_0^{q_0} \times \cdots \times E_n^{q_n}$$

is equipped with the tensor product sigma field. Points in $\mathbf{E}_n^{\mathbf{q}}$ are indexed as follows

$$\left((x_0^1, \ldots, x_0^{q_0}), \ldots, (x_n^1, \ldots, x_n^{q_n})\right)$$

Definition 4.6. *We let $(\gamma_p^{\mathbf{q}})_{0\leq p\leq n}$, and $(\mathbf{Q}_p^{\mathbf{q}})_{0\leq p\leq n}$, be the collection of measures and integral operators defined by*

$$\begin{aligned} \gamma_p^{\mathbf{q}} &= \left(\gamma_0^{\otimes q_0} \otimes \ldots \otimes \gamma_{p-1}^{\otimes q_{p-1}}\right) \otimes \gamma_p^{\otimes (q_p+q'_p)} \ \left(= \gamma_p^{(q_0,\ldots,q_{p-1},q_p+q'_p)}\right) \\ \mathbf{Q}_p^{\mathbf{q}} &= \left(D_{1_{q_0}} \otimes \cdots \otimes D_{1_{q_{p-2}}}\right) \otimes Q_p^{(q_{p-1},q'_{p-1})} \end{aligned}$$

with the operators $Q_p^{(q_{p-1},q'_{p-1})}$ from $E_{p-1}^{q_{p-1}+q'_{p-1}}$ into $E_{p-1}^{q_{p-1}} \times E_p^{q'_{p-1}}$ defined by the tensor product formula

$$Q_p^{(q_{p-1},q'_{p-1})} = D_{1_{q_{p-1}}} \otimes Q_p^{\otimes q'_{p-1}}$$

Notice that, for any $0 \leq p \leq n$, $\gamma_p^{\mathbf{q}}$ is a positive measure on the product space

$$\mathbf{E}_p^{\mathbf{q}} =_{\text{def.}} E_0^{q_0} \times \cdots \times E_{p-1}^{q_{p-1}} \times E_p^{q_p+q_p'}$$

and $\mathbf{Q}_p^{\mathbf{q}}$ is a positive integral operator from $\mathbf{E}_{p-1}^{\mathbf{q}}$ into $\mathbf{E}_p^{\mathbf{q}}$.

Lemma 4.7. *For any sequence of integers $\mathbf{q} \in \mathbb{N}^{n+1}$, and any $0 \leq p \leq n$, we have*

$$\gamma_n^{\mathbf{q}} = \gamma_p^{\mathbf{q}} \mathbf{Q}_{p,n}^{\mathbf{q}}$$

with the semigroup $(\mathbf{Q}_{p_1,p_2}^{\mathbf{q}})_{1 \leq p_1 \leq p_2 \leq n}$ defined by

$$\mathbf{Q}_{p_1,p_2}^{\mathbf{q}} = \mathbf{Q}_{p_1+1}^{\mathbf{q}} \left(\mathbf{Q}_{p_1+1,p_2}^{\mathbf{q}}\right) \quad \left(= \mathbf{Q}_{p_1+1}^{\mathbf{q}} \ldots \mathbf{Q}_{p_2-1}^{\mathbf{q}} \mathbf{Q}_{p_2}^{\mathbf{q}}\right)$$

and the convention $\mathbf{Q}_{p_1,p_1}^{\mathbf{q}} = \mathrm{Id}$, the identity operator, for $p_1 = p_2$.

Proof. We start with observing that

$$\begin{aligned} \gamma_n^{\mathbf{q}} &= \gamma_0^{\otimes q_0} \otimes \gamma_1^{\otimes q_1} \otimes \cdots \otimes \gamma_{n-1}^{\otimes q_{n-1}} \otimes \left(\gamma_{n-1}^{\otimes q_n} Q_n^{\otimes q_n}\right) \\ &= \gamma_0^{\otimes q_0} \otimes \gamma_1^{\otimes q_1} \otimes \cdots \otimes \gamma_{n-2}^{\otimes q_{n-2}} \otimes \left(\gamma_{n-1}^{\otimes (q_{n-1}+q_n)} (D_{1_{q_{n-1}}} \otimes Q_n^{\otimes q_n})\right) \end{aligned}$$

Therefore, we find that $\gamma_n^{\mathbf{q}} = \gamma_{n-1}^{\mathbf{q}} \mathbf{Q}_n^{\mathbf{q}}$. Using a simple induction, the lemma follows. ∎

4.4 Unnormalized particle measures

In this section, we derive a functional representation and a Laurent expansion of particle tensor product measures on the path space similar to the ones obtained in Section 2.2.

In the further development of this section, the time horizon $n \geq 0$ is a fixed parameter, and we let $\mathbf{q} = (q_0, \ldots, q_n) \in \mathbb{N}^{n+1}$ be a given sequence of $(n+1)$ integers. Before we turn to the study of the particle tensor product measures

$$\gamma_n^{\mathbf{q},N} = (\gamma_0^N)^{\otimes q_0} \otimes (\gamma_1^N)^{\otimes q_1} \otimes \cdots \otimes (\gamma_n^N)^{\otimes q_n}$$

and the associated path space measures $\overline{\mathbb{Q}}_{n,\mathbf{q}}^N \in \mathcal{M}(\mathbf{E}_n^{\mathbf{q}})$ defined by

$$\overline{\mathbb{Q}}_{n,\mathbf{q}}^N : \; F \in \mathcal{B}_b(\mathbf{E}_n^{\mathbf{q}}) \longmapsto \overline{\mathbb{Q}}_{n,\mathbf{q}}^N(F) = \mathbb{E}\left(\gamma_n^{\mathbf{q},N}(F)\right) \in \mathbb{R}$$

let us introduce some useful definitions. We associate with $\mathbf{q}$ the pair of integer sequences $\overline{\mathbf{q}}$, and $\mathbf{q}'$ defined by

$$\overline{\mathbf{q}} := (q_k, q_k')_{-1 \leq k \leq n} \quad \text{and} \quad \mathbf{q}' := (q_k')_{-1 \leq k \leq n}$$

with the convention $q_{-1} = 0$, and the integer sequence $\mathbf{q}'$ introduced in (4.2). Notice that $q_n' = 0$, $|\mathbf{q}| = q_0 + q_0'$, $\overline{ht}(\overline{\mathbf{q}}) = (n+1)$, and

$$|\mathbf{q}'| = \sum_{-1 \leq k \leq n} q_k' = \sum_{0 \leq k \leq n} (q_k + q_k') = \sum_{0 \leq k \leq n} (k+1) q_k$$

Definition 4.8. *We let $\Delta_{n,\mathbf{q}}$ be the nonnegative measure valued functional on colored jungles $\overline{J}(\overline{\mathbf{a}}) \in \overline{\mathcal{J}}_{\overline{\mathbf{q}}}$, with $\overline{\mathbf{a}} = (\overline{a}_0, \overline{a}_1, ..., \overline{a}_n) \in \overline{\mathcal{A}}_{\overline{\mathbf{q}}}$, and defined by*

$$\Delta_{n,\mathbf{q}} : \overline{J}(\overline{\mathbf{a}}) \longmapsto \Delta_{n,\mathbf{q}}^{\overline{\mathbf{a}}} := \eta_0^{\otimes |\mathbf{q}|} \mathbf{D}_{0,\overline{a}_0}^{\mathbf{q}} \mathbf{Q}_1^{\mathbf{q}} \mathbf{D}_{1,\overline{a}_1}^{\mathbf{q}} \ldots \mathbf{Q}_n^{\mathbf{q}} \mathbf{D}_{n,\overline{a}_n}^{\mathbf{q}} \in \mathcal{M}(\mathbf{E}_n^{\mathbf{q}})$$

with the Markov transitions $\mathbf{D}_{k,\overline{a}_k}^{\mathbf{q}}$, resp. $D_{\overline{a}_k}$, with $\overline{a}_k = (a_k \amalg a_k') \in [q_{k-1}']^{[q_k] \amalg [q_k']}$, from $\mathbf{E}_k^{\mathbf{q}}$, resp. $E_k^{q_k+q_k'}$, into itself, and defined by

$$\mathbf{D}_{k,\overline{a}_k}^{\mathbf{q}} = D_{1_{q_0}} \otimes \ldots \otimes D_{1_{q_{k-1}}} \otimes D_{\overline{a}_k}$$

and for any function $F \in \mathcal{B}_b(E_k^{q_k+q_k'})$ by

$$D_{\overline{a}_k}(F)((x_k^1, \ldots, x_k^{q_k}), (y_k^1, \ldots, y_k^{q_k'})) = F((x_k^{a_k(1)}, \ldots, x_k^{a_k(q_k)}), (y_k^{a_k'(1)}, \ldots, y_k^{a_k'(q_k')}))$$

Definition 4.9. *We write $\mathcal{B}_b^{sym}(\mathbf{E}_n^{\mathbf{q}})$ for the elements F of $\mathcal{B}_b(\mathbf{E}_n^{\mathbf{q}})$ that are symmetric in the first q_0 variables, the next q_1 variables,..., and the last q_n variables. That is we have that*

$$\forall(\sigma_0,\ldots,\sigma_n)\in(\mathcal{G}_{q_0}\times\ldots\times\mathcal{G}_{q_n})\qquad F=(D_{\sigma_0}\otimes\ldots\otimes D_{\sigma_n})\,F$$

We consider the subset $\mathcal{B}_0^{sym}(\mathbf{E}_n^{\mathbf{q}})\subset\mathcal{B}_b^{sym}(\mathbf{E}_n^{\mathbf{q}})$ of all functions F such that

$$\left(D_{1_{q_0}}\otimes\ldots\otimes D_{1_{q_{p-1}}}\otimes\left(D_{1_{q_{p-1}}}\otimes\gamma_p\right)\otimes D_{1_{q_{p+1}}}\otimes\ldots\otimes D_{1_{q_n}}\right)F=0$$

for any $0\leq p\leq n$.

Notice that $\mathcal{B}_0^{sym}(\mathbf{E}_n^{\mathbf{q}})$ contains the tensor product

$$(\mathcal{B}_0^{sym}(E_0^{q_0})\otimes\ldots\otimes\mathcal{B}_0^{sym}(E_n^{q_n}))=\left\{F_0\otimes\ldots\otimes F_n\ :\ \forall 0\leq p\leq n\quad F_p\in\mathcal{B}_0^{sym}(E_p^{q_p})\right\}$$

of the sets $\mathcal{B}_0^{sym}(E_p^{q_p})$ introduced in the beginning of section 3.9, with $0\leq p\leq n$. For the same reasons as in equation 1.15, we also have the symmetry invariance property

$$\forall F\in\mathcal{B}_b^{sym}(\mathbf{E}_n^{\mathbf{q}})\quad\forall\overline{\mathbf{s}}\in\overline{\mathcal{G}}_{\overline{\mathbf{q}}}\quad\forall\overline{\mathbf{f}}\in\overline{J}_{\overline{\mathbf{q}}}\qquad\Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}(F)=\Delta_{n,\mathbf{q}}^{\overline{\mathbf{s}}(\overline{\mathbf{f}})}(F)$$

By construction, we also have that

$$\forall F\in\mathcal{B}_0^{sym}(\mathbf{E}_n^{\mathbf{q}})\qquad\Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}(F)=0$$

as soon as the colored forest $\overline{\mathbf{f}}$ contains at least one trivial colored tree with a white leaf.

Recall that a tree, or a colored tree is said to be trivial if its coalescence sequence is the null sequence of integers. Notice also that for any $\mathbf{l}<\mathbf{q}'$, the set of colored forests with exactly $l_k(\in[0,q'_{k-1}[)$ coalescent edges at level k, is given by

$$\overline{\mathcal{F}}_{\overline{\mathbf{q}}}[\mathbf{l}]:=\left\{\overline{\mathbf{f}}\in\overline{\mathcal{F}}_{\overline{\mathbf{q}}}\ :\ |\overline{\mathbf{f}}|=\mathbf{q}'-\mathbf{l}\right\}$$

Thus, for any $\mathbf{r}<\mathbf{q}'$, the set of colored forests $\overline{\mathcal{F}}_{\overline{\mathbf{q}}}(\mathbf{r})$ with less than r_k coalescent edges at level k is given by

$$\overline{\mathcal{F}}_{\overline{\mathbf{q}}}(\mathbf{r}):=\cup_{\mathbf{l}\leq\mathbf{r}}\overline{\mathcal{F}}_{\overline{\mathbf{q}}}[\mathbf{l}]$$

The coalescence degree of a colored forest $\overline{\mathbf{f}}\in\overline{\mathcal{F}}_{\overline{\mathbf{q}}}[\mathbf{l}]$, is the sum $|\mathbf{q}'-\mathbf{l}|$ of the coalescence orders of its vertices.

Notice that a colored forest with coalescence degree d has at most d non trivial colored trees. In addition, a colored forest with a coalescence degree d, has at most $(2d)$ leaves belonging to non trivial trees. Since a colored forest in $\overline{\mathcal{F}}_{\overline{\mathbf{q}}}$ has exactly $|\mathbf{q}|$ white leaves, if $|\mathbf{q}|>(2d)$, then it contains at least one trivial colored tree with a white leaf.

Next, we discuss the situation where $|\mathbf{q}|$ is an even integer, and we characterize the subset of forests (in $\overline{\mathcal{F}}_{\overline{\mathbf{q}}}$), with a coalescence degree $d=|\mathbf{q}|/2$, without any trivial colored tree with a white leaf. This charaterization follows the same lines of arguments as the ones presented on page 31. For any integers $0\leq k\leq l\leq m\leq n$, we let $\overline{T}_{k,l,m}$ be the unique colored tree, with a single coalescence at level k, and white leaves at the levels $(l+1)$, and $(m+1)$. Notice that colored forests with $(2r_k)$ white leaves, no black leaves, and coalescent degree r_k, with r_k different pairs of coalescent edges at level k are necessarily of the form

$$\overline{\mathbf{f}}_k^{(\mathbf{t}_k)}=\prod_{k\leq l\leq m\leq n}\overline{T}_{k,l,m}^{\,t_{k,l,m}}$$

for some families of integers $\mathbf{t}_k=(t_{k,l,m})_{k\leq l\leq m\leq n}$ such that $|\mathbf{t}_k|=\sum_{k,l,m}t_{k,l,m}=r_k$. Therefore, a colored forest in $\overline{\mathcal{F}}_{\overline{\mathbf{q}}}$, with coalescence degree $|\mathbf{q}|/2$, a coalescence sequence $\mathbf{r}$ such that $|\mathbf{r}|=|\mathbf{q}|/2$, without a trivial tree with a white leaf, has necessarily the following form

$$\overline{\mathbf{f}}^{(\mathbf{t})}=\overline{\mathbf{f}}_0^{(\mathbf{t}_0)}U_0^{r_0}\ldots\overline{\mathbf{f}}_n^{(\mathbf{t}_n)}U_n^{r_n}$$

for some sequences of integers $\mathbf{t}=(\mathbf{t}_k)_{0\leq k\leq n}$, such that $|\mathbf{t}_k|=r_k$, for any $0\leq k\leq n$. In the above displayed formula, U_k denotes the unique trivial tree with a single black leaf at level k, with $0\leq k\leq n$. We write $\mathbf{t}!$ for $\prod_{k\leq l\leq m\leq n}t_{k,l,m}!$.

We are now ready to extend the Laurent expansions, and the Wick formula presented in theorem 2.7, and theorem 3.15 to particle models in path spaces.

Theorem 4.10. *For any $\mathbf{q} = (q_0, \ldots, q_n)$, with $|\mathbf{q}| \leq N$, and any $F \in \mathcal{B}_b^{sym}(\mathbf{E}_n^{\mathbf{q}})$, we have the Laurent expansion*

$$\overline{\mathbb{Q}}_{n,\mathbf{q}}^N(F) = (\gamma_n^{\mathbf{q}} + \sum_{1 \leq k \leq |(\mathbf{q}'-\mathbf{1})_+|} \frac{1}{N^k}\, \partial^k \overline{\mathbb{Q}}_{n,\mathbf{q}})(F) \tag{4.3}$$

with the signed measures $\partial^k \overline{\mathbb{Q}}_{n,\mathbf{q}}$ defined by

$$\partial^k \overline{\mathbb{Q}}_{n,\mathbf{q}}(F) = \sum_{\mathbf{r} < \mathbf{q}',\ |\mathbf{r}|=k} \ \sum_{\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{q}}}(\mathbf{r})} \frac{s(|\overline{\mathbf{f}}|, \mathbf{q}' - \mathbf{r})\ \#(\overline{\mathbf{f}})}{(\mathbf{q}')_{|\overline{\mathbf{f}}|}}\ \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}(F)$$

In addition, for any even integer $|\mathbf{q}| \leq N$, and any symmetric function $F \in \mathcal{B}_0^{sym}(\mathbf{E}_n^{\mathbf{q}})$, we have

$$\forall k < |\mathbf{q}|/2 \qquad \partial^k \overline{\mathbb{Q}}_{n,\mathbf{q}}(F) = 0 \quad \text{and} \quad \partial^{|\mathbf{q}|/2} \overline{\mathbb{Q}}_{n,\mathbf{q}}(F) = \sum_{\mathbf{r} < \mathbf{q}', |\mathbf{r}| = |\mathbf{q}|/\mathbf{2}} \ \sum_{\|\mathbf{t}\| = \mathbf{r}} \frac{\mathbf{q}!}{2^{\delta(\mathbf{t})}\mathbf{t}!} \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}^{(\mathbf{t})}} F$$

with the integer sequence $\|\mathbf{t}\| = (|\mathbf{t}_k|)_{0 \leq k \leq n}$, and the sum of the diagonal terms $\delta(\mathbf{t}) = \sum_{0 \leq k \leq l \leq n} t_{k,l,l}$. For odd integers $|\mathbf{q}| \leq N$, the partial measure valued derivatives ∂^k are the null measure on $\mathcal{B}_0^{sym}(\mathbf{E}_n^{\mathbf{q}})$, up to any order $k \leq \lfloor |\mathbf{q}|/2 \rfloor$.

Before getting into the details of the proof, we mention that the Wick formula stated above has a natural interpretation in terms of the gaussian fields $(V_k)_{0 \leq k \leq n}$ introduced in section 3.9. Following the discussion given in that section, we get that

$$\partial^{|\mathbf{q}|/2} \overline{\mathbb{Q}}_{n,\mathbf{q}}(F) = \mathbb{E}\left(\left(V_0^{\otimes q_0} \otimes \ldots \otimes V_n^{\otimes q_n}\right)(F)\right)$$

for any tensor product function F of the following form $F = F_{0,q_0} \otimes \ldots \otimes F_{n,q_n}$, with $F_{k,q_k} = \frac{1}{q_k!} \sum_{\sigma_k \in \mathcal{G}_{q_k}} (\varphi_k^{\sigma_k(1)} \otimes \ldots \otimes \varphi_k^{\sigma_k(q_k)})$, $\varphi_k^i \in \mathcal{B}_b(E_k)$, and $\gamma_k(\varphi_k^i) = 0$, for any $1 \leq i \leq q_k$, and any $0 \leq k \leq n$.

Proof. By definition of the particle model, and arguing as in the proof of (2.1) we find that

$$\begin{aligned}
\mathbb{E}\left(\left[(\gamma_{n-1}^N)^{\otimes q_{n-1}} \otimes (\gamma_n^N)^{\otimes q_n}\right](\varphi) \,\middle|\, \xi_{n-1}^{(N)}\right) &= \left[(\gamma_{n-1}^N)^{\otimes q_{n-1}} \otimes \left\{(\gamma_{n-1}^N)^{\otimes q_{n-1}} Q_n^{\otimes q_n} D_{L_{q_n}^N}\right\}\right](\varphi) \\
&= \left[(\gamma_{n-1}^N)^{\otimes (q_{n-1}+q_n)}\right] Q_n^{(q_{n-1},q_n)}\left(D_{1_{q_{n-1}} \amalg L_{q_n}^N}\right)(\varphi) \\
&= \left[(\gamma_{n-1}^N)^{\otimes (q_{n-1}+q'_{n-1})}\right] Q_n^{(q_{n-1},q'_{n-1})}\left(D_{1_{q_{n-1}} \amalg L_{q_n}^N}\right)(\varphi)
\end{aligned}$$

for any $\varphi \in \mathcal{B}_b^{sym}(E_{n-1}^{q_{n-1}} \times E_n^{q_n})$. This yields that

$$\mathbb{E}\left(\gamma_n^{\mathbf{q},N}(F) \,\middle|\, \xi_{n-1}^{(N)}\right) = \gamma_{n-1}^{\mathbf{q},N} \mathbf{Q}_n^{\mathbf{q}} \mathbf{D}_{n,L_{q_n}^N}^{\mathbf{q}}(F) = \gamma_{n-1}^{\mathbf{q},N} \mathbf{Q}_n^{\mathbf{q}} \mathbf{D}_{n,L_{q_n+q'_n}^N}^{\mathbf{q}}(F)$$

from which we readily conclude that

$$\mathbb{E}\left(\gamma_n^{\mathbf{q},N}(F)\right) = \mathbb{E}\left(\gamma_{n-1}^{\mathbf{q},N}\left[\mathbf{Q}_n^{\mathbf{q}} \mathbf{D}_{n,L_{q_n+q'_n}^N}^{\mathbf{q}}(F)\right]\right)$$

A simple induction yields that

$$\begin{aligned}
\mathbb{E}\left(\gamma_n^{\mathbf{q},N}(F)\right) &= \mathbb{E}\left(\eta_0^{\otimes q_0+q'_0} \mathbf{D}_{0,L_{q_0+q'_0}^N}^{\mathbf{q}} \mathbf{Q}_1^{\mathbf{q}} \mathbf{D}_{1,L_{q_1+q'_1}^N}^{\mathbf{q}} \ldots \mathbf{Q}_n^{\mathbf{q}} \mathbf{D}_{n,L_{q_n+q'_n}^N}^{\mathbf{q}}(F)\right) \\
&= \frac{1}{N^{|\mathbf{q}'|}} \sum_{\overline{\mathbf{a}} \in \overline{\mathcal{A}}_{\overline{\mathbf{q}}}} \frac{(\mathbf{N})_{|\overline{\mathbf{a}}|}}{(\mathbf{q}')_{|\overline{\mathbf{a}}|}}\, \Delta_{n,\mathbf{q}}^{\overline{\mathbf{a}}}(F) = \frac{1}{N^{|\mathbf{q}'|}} \sum_{\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{q}}}} \frac{(\mathbf{N})_{\overline{\mathbf{f}}}}{(\mathbf{q}')_{\overline{\mathbf{f}}}}\, \#(\overline{\mathbf{f}})\, \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}(F)
\end{aligned}$$

from which we find the following formula

$$\overline{\mathbb{Q}}_{n,\mathbf{q}}^N(F) = \frac{1}{N^{|\mathbf{q}'|}} \sum_{\mathbf{1} \leq \mathbf{p} \leq \mathbf{q}'} \frac{(\mathbf{N})_{|\mathbf{p}|}}{(\mathbf{q}')_{|\mathbf{p}|}} \sum_{\overline{\mathbf{f}} \in \overline{\mathcal{F}}_{\overline{\mathbf{q}}} : |\overline{\mathbf{f}}| = \mathbf{p}} \#(\overline{\mathbf{f}})\, \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}(F)$$

Using the Stirling formula (2.3), we readily check that

$$\overline{\mathbb{Q}}_{n,\mathbf{q}}^N = \sum_{\mathbf{1}\leq \mathbf{l}\leq \mathbf{q}'} \sum_{\mathbf{l}\leq \mathbf{p}\leq \mathbf{q}'} s(\mathbf{p},\mathbf{l}) \frac{1}{N^{|\mathbf{q}'-\mathbf{l}|}} \frac{1}{(\mathbf{q}')_{|\mathbf{p}|}} \sum_{\overline{\mathbf{f}}\in \overline{\mathcal{F}}_{\overline{\mathbf{q}}}:|\overline{\mathbf{f}}|=\mathbf{p}} \#(\overline{\mathbf{f}})\ \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}$$

From previous computations, we conclude that

$$\overline{\mathbb{Q}}_{n,\mathbf{q}}^N = \sum_{\mathbf{r}<\mathbf{q}'} \frac{1}{N^{|\mathbf{r}|}} \sum_{\overline{\mathbf{f}}\in \overline{\mathcal{F}}_{\overline{\mathbf{q}}}(\mathbf{r})} s(|\overline{\mathbf{f}}|,\mathbf{q}'-\mathbf{r}) \frac{1}{(\mathbf{q}')_{|\overline{\mathbf{f}}|}}\ \#(\overline{\mathbf{f}})\ \Delta_{n,\mathbf{q}}^{\overline{\mathbf{f}}}$$

Finally, we notice that $\overline{\mathcal{F}}_{\overline{\mathbf{q}}}(\mathbf{0})$ reduces to the single class of all sequences of bijections in $\overline{\mathcal{A}}_{\overline{\mathbf{q}}}$. The end the proof of the first assertion is now clear. To end the proof of the theorem, we notice that (with $\mathbf{t}$ as in the expansion of $\partial^{|\mathbf{q}|/2}\overline{\mathbb{Q}}_{n,\mathbf{q}}(F)$)

$$\#(\overline{\mathbf{f}}^{(\mathbf{t})}) = \frac{\overline{\mathbf{q}}!}{\prod_{0\leq k\leq n}\left[\ r_k!\ \left(\prod_{k\leq l\leq m\leq n} t_{k,l,m}!\right)\ \left(\prod_{k\leq l\leq n} 2^{t_{k,l,l}}\right)\right]}$$

Since we have $\mathbf{s}(|\overline{\mathbf{f}}^{(\mathbf{t})}|,\mathbf{q}'-\mathbf{r})=1$, and $(\mathbf{q}')_{|\overline{\mathbf{f}}^{(\mathbf{t})}|}=\prod_{-1\leq k<n}(q'_k)_{q'_k-r_{k+1}}$, for any $\mathbf{r}=(r_k)_{0\leq k\leq n}$, we conclude that

$$\frac{\mathbf{s}(|\overline{\mathbf{f}}^{(\mathbf{t})}|,\mathbf{q}'-\mathbf{r})}{(\mathbf{q}')_{|\overline{\mathbf{f}}^{(\mathbf{t})}|}}\ \#(\overline{\mathbf{f}}^{(\mathbf{t})}) = \frac{\mathbf{q}!}{2^{\delta(\mathbf{t})}\ \mathbf{t}!}$$

with $\mathbf{t}!=\prod_{0\leq k\leq l\leq m\leq n} t_{k,l,m}!$. The end of the proof of the theorem is now straightforward. ∎

4.5 Propagations of chaos type expansions

This section is essentially concerned with applications of the differential forest expansion machinery developed earlier to propagations of chaos properties of interacting particle models. In order to state, and prove the main results of this section, we need to introduce some notation. We shall work throughout with a fixed time horizon $n\geq 0$, and a constant particle block size $q\leq N$. We let $\mathbb{N}_q^{n+1}$ be the set of integer sequences $\mathbf{p}=(p_k)_{0\leq k\leq n}\in\mathbb{N}^{n+1}$ such that $|\mathbf{p}|=q$. We associate to a given $\mathbf{p}=(p_k)_{0\leq k\leq n}\in\mathbb{N}^{n+1}$ the pair of integer sequences $\mathbf{p}'$ and $\mathbf{p}+q$ defined by

$$\forall 0\leq k\leq n \qquad p'_k := \sum_{k<l\leq n} p_l \quad \text{and} \quad \mathbf{p}+q := (p_0,\ldots,p_{n-1},p_n+q)$$

We also denote by $(\overline{G}_k)_{0\leq k\leq n}$, and $\overline{G}_n^{\otimes\mathbf{p}}$ the collection of functions defined by

$$\overline{G}_k := \frac{1}{\gamma_k(G_k)}\times(\eta_k(G_k)-G_k) \quad \text{and} \quad \overline{G}_n^{\otimes\mathbf{p}} := \overline{G}_0^{\otimes p_0}\otimes\ldots\otimes\overline{G}_n^{\otimes p_n}$$

Finally, we consider the integral operators $S_n^{(\mathbf{p}+q)}$ from $\mathbf{E}_n^{\mathbf{p}+q}$ into E_{n+1}^q, defined for any function $F_{n+1}\in\mathcal{B}_b(E_{n+1}^q)$ by the following formula

$$S_n^{(\mathbf{p}+q)}(F_{n+1}) := \overline{G}_0^{\otimes p_0}\otimes\ldots\otimes\overline{G}_{n-1}^{\otimes p_{n-1}}\otimes\left(\overline{G}_n^{\otimes p_n}\otimes Q_{n+1}^{\otimes q}\overline{F}_{n+1}\right)_{\text{sym}}$$

with

$$\overline{F}_n := \frac{1}{\gamma_n(1)^q}\left(F_n-\eta_n^{\otimes q}(F_n)\right)$$

Proposition 4.11. *For any $q\leq N$, and any $n\geq 0$, we have the polynomial decompositions*

$$E_{q,n}^N := \mathbb{E}\left(\left(1-\gamma_n^N(G_n)/\gamma_n(G_n)\right)^q\right) = \sum_{q/2\leq k\leq (n+1)(q-1)} \frac{1}{N^k}\ \partial^k E_{q,n}$$

The derivatives of order $q/2\leq k\leq (n+1)(q-1)$ are given by the following formula

$$\partial^k E_{q,n} = \sum_{\mathbf{p}\in\mathbb{N}_q^{n+1}} \sum_{k\leq |(\mathbf{p}'-1)_+|} \frac{q!}{\mathbf{p}!}\ \partial^k\overline{\mathbb{Q}}_{n,\mathbf{p}}\left(\overline{G}_n^{\otimes\mathbf{p}}\right)$$

Proof. We first check the following decomposition

$$1-\gamma_n^N(G_n)/\gamma_n(G_n)=\sum_{0\leq p\leq n}\gamma_p^N(\overline{G}_p) \tag{4.4}$$

To prove this formula, we simple notice that

$$\begin{aligned}\gamma_n^N(G_n)-\gamma_n(G_n) &= \prod_{0\leq p\leq n}\eta_p^N(G_p)-\prod_{0\leq p\leq n}\eta_p(G_p)\\ &= \gamma_n^N(G_n-\eta_n(G_n))+\left[\gamma_{n-1}^N(G_{n-1})-\gamma_{n-1}(G_{n-1})\right]\times\eta_n(G_n)\\ &= \ldots\\ &= \sum_{0\leq p\leq n}\gamma_p^N(G_p-\eta_p(G_p))\times\left[\prod_{p+1\leq k\leq n}\eta_k(G_k)\right]\end{aligned}$$

with the convention $\prod_{\emptyset}=1$. Finally, combining the multinomial decomposition

$$\left(1-\gamma_n^N(G_n)/\gamma_n(G_n)\right)^q = \sum_{|\mathbf{p}|=q}\frac{q!}{\mathbf{p}!}\,\gamma_n^{\mathbf{p},N}\left(\overline{G}_n^{\mathbf{p}}\right) \tag{4.5}$$

with the Wick expansion stated in Thm.4.10 we conclude that

$$E_{q,n}^N=\sum_{\mathbf{p}\in\mathbb{N}_q^{n+1}}\frac{q!}{\mathbf{p}!}\,\overline{\mathbb{Q}}_{n,\mathbf{p}}^N\left(\overline{G}_n^{\mathbf{p}}\right)=\sum_{\mathbf{p}\in\mathbb{N}_q^{n+1}}\sum_{q/2\leq k\leq|(\mathbf{p}'-1)_+|}\frac{1}{N^k}\,\frac{q!}{\mathbf{p}!}\,\partial^k\overline{\mathbb{Q}}_{n,\mathbf{p}}\left(\overline{G}_n^{\mathbf{p}}\right)$$

This completes the proof of the proposition. ∎

Theorem 4.12. *For any $n\geq 0$, the sequence of probability measures $(\mathbb{P}_{n+1,q}^N)_{N\geq q}$ is differentiable up to any order with $\partial^0\mathbb{P}_{n+1,q}=\eta_{n+1}^{\otimes q}$, and the partial derivatives given by the following formula*

$$\partial^k\mathbb{P}_{n+1,q}=\sum_{\mathbf{p}\in\cup_{0\leq l<2k}\mathbb{N}_l^{n+1}}\frac{((q-1)+|\mathbf{p}|)!}{(q-1)!\;\mathbf{p}!}\;\partial^k\overline{\mathbb{Q}}_{n,\mathbf{p}+q}S_n^{\mathbf{p}+q}$$

At any order $N\geq q$, we have the exact formula

$$\mathbb{P}_{n+1,q}^N=\eta_{n+1}^{\otimes q}+\sum_{1\leq k<\lfloor(N-q)/2\rfloor}\frac{1}{N^k}\;\partial^k\mathbb{P}_{n+1,q}+\frac{1}{N^{\lfloor(N-q)/2\rfloor}}\;\partial^{\lfloor(N-q)/2\rfloor}\mathbb{P}_{n+1,q}^N$$

Proof. We let F_{n+1} be a bounded measurable symmetric function on $\mathbf{E}_{n+1}^q$, such that $\eta_{n+1}^{\otimes q}(F_{n+1})=0$. By definition of the particle model, we have that

$$\begin{aligned}\mathbb{E}\left(\left(\eta_{n+1}^N\right)^{\odot q}(F_{n+1})\mid\xi_n^{(N)}\right) &= \gamma_n^N(G_n)^{-q}\times\left(\gamma_n^N\right)^{\otimes q}Q_{n+1}^{\otimes q}(F_{n+1})\\ &= \left(1-u_n^N\right)^{-q}\times\left(\gamma_n^N\right)^{\otimes q}\;Q_{n+1}^{\otimes q}\left(\overline{F}_{n+1}\right)\end{aligned}$$

with the sequence of random variables $u_n^N:=\left(1-\gamma_n^N(G_n)/\gamma_n(G_n)\right)$. On the other hand, for any $u\in\mathbb{R}-\{1\}$, and $m\geq 1$, we have the more of less well known decomposition

$$\frac{1}{(1-u)^{q+1}}=\sum_{0\leq k\leq m}(q+k)_k\;\frac{u^k}{k!}+u^m\sum_{1\leq k\leq q+1}\left(\begin{array}{c}(q+1)+m\\ k+m\end{array}\right)\left(\frac{u}{1-u}\right)^k$$

with $(q+k)_k=(q+k)!/q!$. The proof of this formula is essentially based on the fact that for any $n\geq 0$, and $u\neq 1$, we have

$$f(u)=\frac{1}{1-u}=\sum_{0\leq k\leq n}u^k\;+\frac{u^{n+1}}{1-u}\text{and}\frac{\partial^n f}{\partial u^n}=n!\;f^{n+1}$$

This implies that for any $m > n$, we have

$$\frac{\partial^n f}{\partial u^n} = \frac{n!}{(1-u)^{n+1}} = \sum_{n \leq k \leq m} (k)_n \; u^{k-n} + \frac{\partial^n}{\partial u^n} \left(\frac{u^{m+1}}{1-u} \right)$$

Recalling the Leibniz binomial derivation formula

$$\frac{\partial^n}{\partial u^n}(fg) = \sum_{0 \leq k \leq n} \left(\begin{array}{c} n \\ k \end{array} \right) \; \frac{\partial^{n-k}}{\partial u^{n-k}}(f) \; \frac{\partial^k}{\partial u^k}(g)$$

we find that

$$\begin{aligned} \frac{1}{n!} \; \frac{\partial^n}{\partial u^n} \left(\frac{u^{m+1}}{1-u} \right) &= \sum_{0 \leq k \leq n} \left(\begin{array}{c} n \\ k \end{array} \right) \; (m+1)_k \; u^{(m+1)-k} \; \frac{(n-k)!}{n!} \; \frac{1}{(1-u)^{(n-k)+1}} \\ &= u^{m-n} \sum_{0 \leq k \leq n} \left(\begin{array}{c} m+1 \\ k \end{array} \right) \; \frac{u^{(n-k)+1}}{(1-u)^{(n-k)+1}} \end{aligned}$$

and

$$\frac{1}{n!} \; \frac{\partial^n}{\partial u^n} \left(\frac{u^{m+1}}{1-u} \right) = u^{m-n} \sum_{0 \leq k \leq n} \left(\begin{array}{c} m+1 \\ n-k \end{array} \right) \; \frac{u^{k+1}}{(1-u)^{k+1}}$$

Thus,

$$\frac{1}{(1-u)^{n+1}} = \sum_{0 \leq k \leq (m-n)} \frac{(k+n)!}{n!} \; \frac{u^k}{k!} + u^{m-n} \sum_{0 \leq k \leq n} \left(\begin{array}{c} m+1 \\ n-k \end{array} \right) \; \frac{u^{k+1}}{(1-u)^{k+1}}$$

We prove the desired decomposition with replacing m by $(m+q)$, and n by q. Using the above decomposition we find that

$$\begin{aligned} &\mathbb{E}\left(\left(\eta_{n+1}^N\right)^{\odot q}(F_{n+1}) \mid \xi_n^{(N)} \right) \\ &= \sum_{0 \leq k \leq m} ((q-1)+k)_k \; \frac{1}{k!} \left(1 - \frac{\gamma_n^N(G_n)}{\gamma_n(G_n)} \right)^k \; \left(\gamma_n^N\right)^{\otimes q} (Q_{n+1}^{\otimes q} \overline{F}_{n+1}) + R_{m,n}^{q,N}(F_{n+1}) \end{aligned}$$

with the remainder term

$$\begin{aligned} &R_{m,n}^{q,N}(F_{n+1}) \\ &= \left(1 - \tfrac{\gamma_n^N(G_n)}{\gamma_n(G_n)} \right)^{m+1} \left(\gamma_n^N\right)^{\otimes q} (Q_{n+1}^{\otimes q} \overline{F}_{n+1}) \; \textstyle\sum_{1 \leq k \leq q} \; \left(\begin{array}{c} q+m \\ k+m \end{array} \right) \; \left(1 - \tfrac{\gamma_n^N(G_n)}{\gamma_n(G_n)} \right)^{k-1} / \left(\tfrac{\gamma_n^N(G_n)}{\gamma_n(G_n)} \right)^k \end{aligned}$$

We also have from proposition 4.11

$$\sup_{N \geq m+1} \sqrt{N} \mathbb{E}\left(\left| \gamma_n^N(G_n) - \gamma_n(G_n) \right|^{m+1} \right)^{\frac{1}{m+1}} < \infty$$

Finally, using the multinomial decomposition (4.5), we conclude that

$$\begin{aligned} &\sum_{0 \leq k \leq m} ((q-1)+k)_k \; \frac{1}{k!} \left(1 - \frac{\gamma_n^N(G_n)}{\gamma_n(G_n)} \right)^k \; \left(\gamma_n^N\right)^{\otimes q} (Q_{n+1}^{\otimes q} \overline{F}_{n+1}) \\ &= \sum_{0 \leq k \leq m} ((q-1)+k)_k \sum_{\mathbf{p} \in \mathbb{N}_k^{n+1}} \frac{1}{\mathbf{p}!} \; \gamma_n^{(\mathbf{p}+q),N} \left(S_n^{(\mathbf{p}+q)}(F_{n+1}) \right) \end{aligned}$$

This yields, for any $m + q \leq N$, the functional expansion

$$\mathbb{P}_{n+1,q}^N = \sum_{|\mathbf{p}| \leq m} \frac{((q-1)+|\mathbf{p}|)!}{(q-1)! \; \mathbf{p}!} \; \overline{\mathbb{Q}}_{n,\mathbf{p}+q}^N S_n^{\mathbf{p}+q} + \overline{R}_{m,n}^{q,N}$$

with a remainder measure $\overline{R}^{q,N}_{m,n}$, such that $\sup_{N\geq 1} N^{(m+1)/2} \left\|\overline{R}^{q,N}_{m,n}\right\|_{\mathrm{tv}} < \infty$. The last assertion follows from the regularity hypothesis 1.1 on the potential functions. This implies that for any $k < (m+1)/2$ we have

$$\partial^k \mathbb{P}_{n+1,q} = \sum_{|\mathbf{p}|\leq m} \frac{((q-1)+|\mathbf{p}|)!}{(q-1)!\ \mathbf{p}!}\ \partial^k \overline{\mathbb{Q}}_{n,\mathbf{p}+q} S_n^{\mathbf{p}+q}$$

Notice that k-th order derivative measure $\partial^k \overline{\mathbb{Q}}_{n,\mathbf{p}+q}$ only involves colored forests with less than k coalescent branches, from the original root, up to the final level. If $|\mathbf{p}| \geq (2k)$, then these colored forests contain at least one elementary tree with a white leaf. By definition of the operator $S_n^{\mathbf{p}+q}$ we find that

$$\partial^k \overline{\mathbb{Q}}_{n,\mathbf{p}+q} S_n^{\mathbf{p}+q}(F_{n+1}) = 0$$

This yields that

$$\mathbb{P}^N_{n+1,q} = \eta_{n+1}^{\otimes q} + \sum_{1\leq k<((N-q)+1)/2} \frac{1}{N^k}\ \partial^k \mathbb{P}_{n+1,q} + \mathbb{R}^N_{n+1,q}$$

with a remainder measure $\mathbb{R}^N_{n+1,q}$ such that $\sup_{N\geq 1} N^{((N-q)+1)/2} \left\|\mathbb{R}^N_{n+1,q}\right\|_{\mathrm{tv}} < \infty$. We end of the proof of the second assertion of the theorem using the fact that

$$\lfloor (N-q)/2 \rfloor \leq (N-q)/2 \leq ((N-q)+1)/2$$

The Theorem follows. ∎

We end this section, with a series of some direct consequences of the above theorem:

- The differential forest expansions presented in theorem 4.12 allow us to deduce precise strong propagation of chaos estimates. For instance, for any $N \geq (q+7)$ (so that $(\lfloor (N-q)/2 \rfloor - 1) \geq 2$) we have

$$\sup_{N\geq q+7} N^2 \left\|\mathbb{P}^N_{n+1,q} - \eta_{n+1}^{\otimes q} - \frac{1}{N}\ \partial^1 \mathbb{P}_{n+1,q}\right\|_{\mathrm{tv}} < \infty$$

  with a first order partial derivative given by the formula

$$\partial^1 \mathbb{P}_{n+1,q}(F_{n+1}) = \partial^1 \mathbb{Q}_{n,q} Q_{n+1}^{\otimes q}(\overline{F}_{n+1}) + q \sum_{0\leq m\leq n} \partial^1 \overline{\mathbb{Q}}_{n,\mathbf{q}_m} S_n^{\mathbf{q}_m}(F_{n+1})$$

  with the sequence of integers $\mathbf{q}_m = (1_m(k) + q1_n(k))_{0\leq k\leq n}$. The first term in the right hand side in the above displayed formula has been treated in corollary 3.14, on page 30, and we have that

$$\begin{aligned}
&\partial^1 \mathbb{Q}_{n,q} Q_{n+1}^{\otimes q}(\overline{F}_{n+1}) \\
&= \frac{q(q-1)}{2} \sum_{0\leq k\leq n} \Delta_{n,q}^{\mathbf{f}_{1,k}} Q_{n+1}^{\otimes q}(\overline{F}_{n+1}) \\
&= \frac{q(q-1)}{2} \sum_{0\leq k\leq n} \gamma_k(1) \int_{E_k^{q-1}} \gamma_k^{(q-1)}(d(x^2,\ldots,x^q))\ Q_{k,n+1}^{\otimes q}(\overline{F}_{n+1})(x^2,x^2,x^3,\ldots,x^q)
\end{aligned}$$

  Each of the terms $\partial^1 \overline{\mathbb{Q}}_{n,\mathbf{q}_m} S_n^{\mathbf{q}_m}(F_{n+1})$ only involves the colored forests

$$\forall 0 \leq k \leq m \qquad \overline{\mathbf{f}}_{k,m} := U_k T_{k,m,n} U_{n+1}^{q-1}$$

  associated with the trees $T_{k,m,n}$, and U_l, introduced on page 38. After some elementary manipulations, we find that

$$\begin{aligned}
&\partial^1 \overline{\mathbb{Q}}_{n,\mathbf{q}_m} S_n^{\mathbf{q}_m}(F_{n+1}) \\
&= q \sum_{0\leq k\leq m} \Delta_{n,\mathbf{q}_m}^{\overline{\mathbf{f}}_{k,m}} S_n^{\mathbf{q}_m}(F_{n+1}) \\
&= q \sum_{0\leq k\leq m} \gamma_k(1) \int_{E_k^q} \gamma_k^{\otimes q}(d(x^1,\ldots,x^q))\ Q_{k,m}(\overline{G}_m)(x^1)\ Q_{k,n+1}^{\otimes q}(\overline{F}_{n+1})(x^1,\ldots,x^q)
\end{aligned}$$

- A Wick formula derives from theorem 4.10. More precisely, for any $F \in \mathcal{B}_0^{\mathrm{sym}}(E_{n+1}^q)$, the partial derivatives $\partial^k \mathbb{P}_{n+1,q}(F)$ are null up to any order $k \leq \lfloor q/2 \rfloor$, and for any even integer q, we have

$$\partial^{q/2}\mathbb{P}_{n+1,q}(F) = \gamma_{n+1}(1)^{-q}\ \partial^{q/2}\mathbb{Q}_{n,q}Q_{n+1}^{\otimes q}(F)$$

as soon as $N \geq 2(q+2)$ (so that $(\lfloor (N-q)/2 \rfloor - 1) \geq q/2$). To prove this claim, we notice that for any $k < q/2$, and any $F \in \mathcal{B}_0^{\mathrm{sym}}(E_{n+1}^q)$, we have

$$\forall \mathbf{p} \in \cup_{0 \leq l < 2k} \mathbb{N}_l^{n+1} \qquad k < q/2 \leq (|\mathbf{p}|+q)/2 \quad \text{and therefore} \quad \partial^k \overline{\mathbb{Q}}_{n,\mathbf{p}+q} S_n^{\mathbf{p}+q}(F) = 0$$

This yields that $\partial^k \mathbb{P}_{n+1,q}(F) = 0$, for any even integer q, and any $k < q/2$. In the case $k = q/2$, we have $k = q/2 = (|\mathbf{p}|+q)/2$ if, and only if, $\mathbf{p}$ coincide with the null sequence $\mathbf{0}$.

- We let $\|\mu\|_{\mathcal{B}_{0,1}^{\mathrm{sym}}} = \sup_{F \in \mathcal{B}_{0,1}^{\mathrm{sym}}} |\mu(F)|$ be the Zolotarev seminorm on $\mathcal{M}(E_{n+1}^q)$ associated with the collection of functions

$$\mathcal{B}_{0,1}^{\mathrm{sym}} := \left\{ F \in \mathcal{B}_0^{\mathrm{sym}}(E_{n+1}^q) \ :\ \|F\| \leq 1 \right\}$$

For any even integer q such that $(q+2) \leq N/2$, we have

$$\sup_{N \geq q+7} N^{1+q/2} \left\| \mathbb{P}_{n+1,q}^N - \frac{1}{N^{q/2}}\ \frac{1}{\gamma_{n+1}(1)^q}\ \partial^{q/2}\mathbb{Q}_{n,q}Q_{n+1}^{\otimes q} \right\|_{\mathcal{B}_{0,1}^{\mathrm{sym}}} < \infty$$

- Combining the Wick formula stated above with the Borel-Cantelli lemma, we obtain for all $q \geq 4$ the almost sure convergence result

$$\lim_{N \to \infty} (\eta_n^N)^{\odot q}(F) = 0 \quad \text{p.s.}$$

for any bounded symmetric function $F \in \mathcal{B}_0^{\mathrm{sym}}(E_n^q)$. This result is an extension of the law of large numbers for U-statistics obtained by W. Hoeffding [7] for independent and identically distributed random variables to interacting particle models.

- We mention that the same lines of arguments used in the proof of theorem 4.12 show that the sequence of probability measures

$$\tilde{\mathbb{P}}_{n+1,q}^N \ :\ F \in \mathcal{B}_b(E_{n+1}^q) \mapsto \tilde{\mathbb{P}}_{n+1,q}^N(F) := \mathbb{E}\left(\left(\eta_{n+1}^N\right)^{\otimes q}(F) \right)$$

is differentiable up to order $\lfloor (N-q)/2 \rfloor$, with $\partial^0 \tilde{\mathbb{P}}_{n+1,q} = \eta_{n+1}^{\otimes q}$, and the partial derivatives given for any $1 \leq k < \lfloor (N-q)/2 \rfloor$ by the following formula

$$\partial^k \tilde{\mathbb{P}}_{n+1,q}(F) = \sum_{\mathbf{p} \in \cup_{0 \leq l < 2k} \mathbb{N}_l^{n+1}} \frac{((q-1)+|\mathbf{p}|)!}{(q-1)!\ \mathbf{p}!}\ \partial^k \overline{\mathbb{Q}}_{n+1,(\mathbf{p},q)} \left(\overline{G}_n^{\mathbf{p}} \otimes \overline{F} \right)$$

In the same way, for any $F \in \mathcal{B}_0^{\mathrm{sym}}(E_{n+1}^q)$, and $N \geq 2(q+2)$, the partial derivatives $\partial^k \mathbb{Q}_{n+1,q}(F)$ are null up to any order $k < q/2$, and for any even integer q, we have that

$$\partial^{q/2}\tilde{\mathbb{P}}_{n+1,q}(F) = \gamma_{n+1}(1)^{-q}\ \partial^{q/2}\mathbb{Q}_{n+1,q}(F)$$

- Finally, the Wick formula stated above allows to deduce sharp $\mathbb{L}_q$-mean error bound. To see this claim, we simply observe that

$$F = (f - \eta_{n+1}(f))^{\otimes q} \quad \text{with} \quad f \in \mathcal{B}_b(E_{n+1}) \Longrightarrow \tilde{\mathbb{P}}_{n+1,q}^N(F) = \mathbb{E}(\left[\eta_{n+1}^N(f) - \eta_{n+1}(f)\right]^q)$$

and for any even integer q such that $(q+2) \leq N/2$, we have the inequality

$$\sup_{N \geq q+7} N^{1+q/2} \left\| \tilde{\mathbb{P}}_{n+1,q}^N - \frac{1}{N^{q/2}}\ \frac{1}{\gamma_{n+1}(1)^q}\ \partial^{q/2}\mathbb{Q}_{n+1,q} \right\|_{\mathcal{B}_{0,1}^{\mathrm{sym}}} < \infty$$

References

[1] Ch. Brouder. Runge-Kutta methods and renormalization. *Eur. Phys. J. C.* 12, 521-534 (2000).

[2] L. Comtet, *Analyse combinatoire*, Tome second, Presses universitaires de France (1970).

[3] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Commun. Math. Phys.* 199, No.1, 203-242 (1998).

[4] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems, with applications*, Springer Verlag New York, Series: Probability and its Applications (2004).

[5] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Pratice*. Statistics for Engineering and Information Science. Springer New York, 2001.

[6] A. Guionnet and E. Segala Maurel, *Combinatorial aspects of matrix models*, arXiv:math.PR/0503064 v2, June 2005.

[7] W. Hoeffding, The strong law of large numbers for U-statistics. *Inst. Statist. Univ. of North Carolina, Mimeo report no 302* (1961).

[8] P. Hoffman. τ-Rings and wreath product representations. Springer Verlag, *Lect. Notes in Math.* 746 (1979).

[9] S. D. Jacka, G. O. Roberts, On strong forms of weak convergence, *Stoch. Proc. & Appl.* 67, 41-53 (1997).