On particle Gibbs samplers

P. Del Moral[a], R. Kohn[b] and F. Patras[c]

[a]School of Mathematics and Statistics, University of New South Wales, High Street, Kensington NSW 2052, Australia.
E-mail: p.del-moral@unsw.edu.au
[b]School of Economics, University of New South Wales, Sydney NSW 2052, Australia. E-mail: r.kohn@unsw.edu.au
[c]Mathématiques, Université de Nice et CNRS, Parc Valrose, 06108 Nice Cedex 2, France. E-mail: patras@unice.fr

Abstract. This article analyses a new class of advanced particle Markov chain Monte Carlo algorithms recently introduced by Andrieu, Doucet and Holenstein (*J. R. Stat. Soc. Ser. B Stat. Methodol.* **72** (2010) 1–269). We present a natural interpretation of these methods in terms of well known unbiasedness properties of Feynman–Kac particle measures, and a new duality with Feynman–Kac models.

This perspective sheds new light on the foundations and the mathematical analysis of this class of methods. A key consequence is their equivalence with the Gibbs sampling of a many-body Feynman–Kac target distribution. Our approach also presents a new stochastic differential calculus based on geometric combinatorial techniques. These techniques allow us to derive non-asymptotic Taylor type series for the semigroup of a class of particle Markov chain Monte Carlo models around their invariant measures with respect to the population size of the auxiliary particle sampler.

These results provide sharp quantitative estimates of the convergence rate to equilibrium of the models with respect to the time horizon and the size of the systems. We illustrate the direct implication of these results with sharp estimates of the contraction coefficient and the Lyapunov exponent of the corresponding samplers, and explicit and non-asymptotic $\mathbb{L}_p$-mean error decompositions of the law of the random states around the limiting invariant measure. The abstract framework developed in the article also allows the design of natural extensions to island (also called SMC^2) type particle methodologies.

We illustrate this general framework and results in the context of nonlinear filtering, hidden Markov chain problems with fixed unknown parameters, and Feynman–Kac path-integration models arising in computational quantum physics and molecular chemistry.

Résumé. Cet article analyse une classe de méthodes de Monte Carlo avancées de type particulaire introduites par Andrieu, Doucet, et Holenstein (*J. R. Stat. Soc. Ser. B Stat. Methodol.* **72** (2010) 1–269). Nous présentons une interprétation naturelle de ces méthodes en termes de mesures de Feynman–Kac particulaires non biaisées classiques et d'une nouvelle formule de dualité entre modèles de Feynman–Kac.

Ce cadre d'étude apporte un nouvel éclairage sur les fondations et l'analyse mathématique de ces méthodes. Une conséquence importante est l'équivalence de ces dernières avec la méthode d'échantillonnage de Gibbs d'une distribution de Feynman–Kac multi-corps. Notre étude développe aussi un nouveau calcul différentiel stochastique fondé sur des techniques géométriques et combinatoires. Ces techniques permettent d'obtenir des développements non asymptotiques des semigroupes de modèles de Monte Carlo par Chaînes de Markov particulaires autour de leur mesure invariante, en fonction de la taille des systèmes de particules en interaction auxiliaires.

Cette analyse conduit à des estimations quantitatives précises de la convergence à l'équilibre de ces modèles par rapport à l'horizon temporel et la taille des systèmes. Nous illustrons ces résultats avec quelques implications directes, notamment l'estimation précise des coefficients de contraction et des exposants de Lyapunov de ces algorithmes de simulation, ainsi que l'estimation fine de l'erreur en norme $\mathbb{L}_p$ entre la loi des états aléatoires de ces chaînes de Markov et leur mesure d'équilibre. Le cadre abstrait de l'article permet d'élaborer et d'étendre de façon naturelle ces méthodes à des classes d'algorithmes fondés sur des évolutions d'îlots particulaires (aussi connus sous le nom SMC^2).

Nous montrons enfin comment ce cadre général et les résultats de l'article s'appliquent à l'étude de problèmes de filtrage non linéaire, l'estimation de paramètres fixes dans des modèles de chaînes de Markov cachées, et dans des problèmes d'integration trajectorielle rencontrés en physique quantique et en chimie moléculaire.

1. Introduction

In the last two decades particle simulation and Sequential Monte Carlo (SMC) methodologies (see e.g., [16,20,29]) have become one of the most active contact points between Bayesian statistical inference and applied probability. In contrast to conventional Markov chain Monte Carlo methodologies (MCMC), particle methods are not based on sampling long runs of a judiciously chosen Markov chain with a prescribed target probability measure, but on the mean field particle simulation of nonlinear Markov chain models.

The seminal article [2] by Andrieu, Doucet and Holenstein introduced a new way to combine MCMC with SMC methods. A variant of the method, where ancestors are resampled in a forward pass, was developed by Lindsten, Schön and Jordan in [46], and Lindsten and Schön [47]. This new class of Monte Carlo samplers is called particle Markov chain Monte Carlo methods (PMCMC). These emerging particle sampling technologies are particularly important in signal processing and in Bayesian statistics where they are used to estimate posterior distributions of the unknown parameters when the likelihood functions are unknown or are computationally intractable. These methods have recently attracted considerable attention in a variety of application domains, including statistical machine leaning [5,39,46,57], finance and econometrics [15,28,35,48,51], biology [36,44,52], computer science [37], environmental statistics [31, 32,50], social networks analysis [34], signal processing [47,49], forecasting and data assimilation [43,45,56].

The convergence analysis of the PMCMC methods was initiated in a series of articles [3,12,42,46,47]. The ϕ-irreducibility and aperiodicity of PMCMC models was already discussed in the pioneering article [2]. Uniform ergodicity results with quantitative estimates of the convergence properties of particle Gibbs methods are presented by Chopin and Singh in [12], using a coupling technology of ancestral particle paths. More refined contraction estimates are obtained by Andrieu, Lee and Vihola [3] using a conditional type analysis of the normalizing particle constants, and in Lindsten, Douc and Moulines [42] which provides similar quantitative estimates using lower bound estimates of PMCMC transition probabilities based on the stability of Feynman–Kac semigroups. In all of these studies, the validity of PMCMC samplers is assessed by interpreting these models as a traditional MCMC sampler on a sophisticated and extended state space. The more recent article [11] also proposes a first order analysis of the bias estimates of forward and backward particle approximations of Feynman–Kac models. For a time horizon n and particle systems of size N, these first order expansions allow us to consider the bias estimates (both upper and lower bounds) for possibly unbounded functions; these estimates are of order n/N. As a by-product of their analysis, the authors provide convergence rates to equilibrium of order n/N for forward and backward type particle Gibbs samplers.

Our article is concerned with an alternative probabilistic foundation of PMCMC methodology. It is known that Metropolis–Hasting type PMCMC methods reduce to standard Metropolis–Hastings with a target Feynman–Kac distribution that encapsulates the distribution of the auxiliary particle model (see, for instance, the article [1], and Section 2.2.5 in [25], in the context of Hidden Markov Chain problems with fixed unknown parameters).

The first part of our article shows that the PMCMC sampler based on the sampling of a particle model with a frozen trajectory (refered to as a *particle Gibbs sampler* or *PG sampler* from now on, see Section 3.1) reduces to a standard Gibbs sampler described in Section 3.2. The proof of this result is based on a new duality relation (Theorem 3.1) that can be seen as an extension of known unbiasedness properties of Feynman–Kac models to their many-body version (henceforth, the many-body Feynman–Kac models refers to the extension of the usual Feynman–Kac models to collective motions of particles; see Section 2.6 for precise definitions).

This natural point of view considerably simplifies the design and the convergence analysis of PG samplers, see e.g., Section 4.4. Finally, the new formulation also allows the design of new and natural classes of PMCMC samplers based on island type models and particle Gibbs methodology, see Section 6.1.

The last part of the article is concerned with the propagation of chaos properties of PG samplers. We design explicit Taylor type expansions of the law of a finite block of particles in terms of the population size of the particle model (Theorem 5.21). These expansions are naturally parametrized by decorated ("infected") forests. Their accuracy at any

order is related naturally to the number of coalescent edges and the number of infections. To the best of our knowledge, these propagation of chaos series are the first results of this type for this class of PMCMC methods.

The convergence rate n/N to equilibrium for PG samplers follows from the propagation of chaos expansion as soon as the Feynman–Kac semigroup of the marginal target measures forgets its initial value (see, for example, Theorem 3.2, and the regularity condition (2.6)). We note that similar decay rates to equilibrium with linear scaling are also derived in [3,42], but under stronger mixing conditions on the reference Markov chain, and using different lower bound estimation techniques on the minorizing constant for the particle Gibbs kernel. Using partially heuristic arguments, the linear scaling of the number of particles with respect to the time horizon is also discussed in [30,51].

These expansions also provide Taylor decompositions of the semigroup of PG chains *around their invariant target measures* with respect to the precision parameter $1/N$, where N represents the size of the particle system (see, e.g., Theorem 4.14). Subject to some stability properties on Feynman–Kac semigroups, a direct consequence of these Taylor series expansions is that the Dobrushin contraction coefficient is of order n/N, where n stands for the length of the trajectories (Proposition 4.11).

The article [11] shows that the linear dependence of the convergence rate with respect to the time horizon n is not surprising: the same type of linear scaling arises in the propagation of chaos and the fluctuation analysis of Feynman–Kac particle models on path spaces (see, for instance, Corollary 8.5.1 in [16], as well as Corollary 15.2.5 and Theorem 16.5.1 in [20]; see also Theorem 3.3 in [11]).

The impact of the Taylor series expansions developed in the present article is also illustrated with sharp and non-asymptotic expansions of the Dobrushin contraction coefficient of iterated PG transitions. We also provide an explicit decomposition of the $\mathbb{L}_p$-distance between the law of the random states of a class of PMCMC methods around the limiting invariant measure. These results can also be used to estimate the bias and the variance of the random states of the occupation measures of the particle systems. The duality phenomenon allows us to apply these Taylor expansions to the original Feynman–Kac particle models.

The article is organized as follows. Section 2 reviews some well known results on Feynman–Kac models and their mean field particle interpretation, including path space models and backward particle Markov chain measures. We illustrate the results with applications to nonlinear filtering and physics. Section 2.6 introduces many-body Feynman–Kac models aimed at describing the collective motion of particles in the usual Feynman–Kac models. These models are particularly well suited to the analysis of PMCMC samplers.

Section 3 briefly describes particle Gibbs models and states some of the main results of the article.

Section 4 studies these models in depth and proves that PG samplers reduce to Gibbs samplers of a many-body Feynman–Kac target measure. Section 4.1 provides a transport equation and a new duality relation between many-body Feynman–Kac models and Feynman–Kac particle models with a frozen trajectory. Section 4.2 considers historical particle methods and their dual frozen particle methods. For instance, we show that the conditional distribution of *the ancestral lines* of the Feynman–Kac particle model with respect to the complete history of its populations coincides with the backward particle method. Section 4.3 studies the behaviour (reversibility, Dobrushin coefficient, etc.) of the Markov transitions of two PG samplers (the one based on the selection of an ancestral line in a genealogical tree and the one based on a backward sampling method). Section 4.4 describes the Taylor expansions of PG transitions around their invariant measures. We also derive some important consequences of these expansions, including quantitative estimates of the stability properties of these methods, and sharp estimates of the bias and variance of the random states of the PG Markov chain.

Section 5 considers the propagation of chaos properties of these models. Section 5.1 collects some preliminary combinatorial results on tensor products of empirical measures. Section 5.2 considers non-asymptotic Taylor series of q-tensor products of unnormalized particle measures. Section 5.3 discusses the propagation of chaos properties and related Taylor expansions of frozen particle models. Section 5.4 describes the Taylor series decompositions in terms of infected and coalescent forest expansions.

Section 6 presents a new class of island PMCMC samplers, and discusses some extensions and open questions.

2. From mean field to many-body Feynman–Kac models

This section first collects some basic notation used in this article and then recalls the definition and main properties of Feynman–Kac measures on their usual state and path spaces. The last paragraph introduces a particular Feynman–Kac model that is well-suited to the mathematical analysis of PMCMC samplers (see also [55]). We refer to [16] for further details on Feynman–Kac models and their standard properties.

2.1. *Notation*

Given some measurable space S, we denote by $\mathcal{M}(S)$, $\mathcal{P}(S)$ and $\mathcal{B}(S)$, respectively, the set of finite signed measures on S, the convex subset of probability measures, and the Banach space of bounded measurable functions equipped with the uniform norm $\|f\| = \sup_{x \in S} |f(x)|$.

The total variation norm on measures $\mu \in \mathcal{M}(S)$ is defined by

$$\|\mu\|_{\mathrm{tv}} := \sup_{f \in \mathcal{B}(S): \|f\| \leq 1} |\mu(f)| \quad \text{with the Lebesgue integral } \mu(f) := \int \mu(dx) f(x).$$

We also denote by δ_a the Dirac measure at some state a, so that $\delta_a(f) = f(a)$. We say that $\nu \leq \mu$ if $\nu(f) \leq \mu(f)$ for any non-negative function f.

A bounded integral operator $Q(x, dy)$ between the measurable spaces S and S' maps any $f \in \mathcal{B}(S')$ to the measurable function $Q(f) \in \mathcal{B}(S)$ defined by

$$Q(f)(x) := \int Q(x, dy) f(y).$$

The operator Q generates a dual operator $\mu \in \mathcal{M}(S) \mapsto \mu Q \in \mathcal{M}(S')$ by the dual formula $(\mu Q)(f) = \mu(Q(f))$.

When a bounded integral operator M from a state space S into a possibly different state space S' has a constant mass, that is, when $M(1)(x) = M(1)(y)$ for any $x, y \in S$, the operator $\mu \mapsto \mu M$ maps the set $\mathcal{M}_0(S)$ of measures μ on S with null mass $\mu(1) = 0$ into $\mathcal{M}_0(S')$. In this situation, we let $\beta(M)$ be the Dobrushin coefficient of a bounded integral operator M defined by the formula

$$\beta(M) := \sup \{ \mathrm{osc}(M(f)); \ f \ \text{s.t.} \ \mathrm{osc}(f) \leq 1 \},$$

where $\mathrm{osc}(f) := \sup_{x,y} |f(x) - f(y)|$ stands for the oscillation of some function.

When M is a Markov transition operator, $\beta(M)$ coincides with the Dobrushin contraction parameter, also known as the Dobrushin ergodic coefficient, which is defined by

$$\beta(M) = \sup_{\mu,\nu} \left(\|\mu M - \nu M\|_{\mathrm{tv}} / \|\mu - \nu\|_{\mathrm{tv}} \right) = 2^{-1} \sup_{x,y} \|M(x, \cdot) - M(y, \cdot)\|_{\mathrm{tv}}.$$

The q-tensor product of Q is the integral operator defined for any $f \in \mathcal{B}(S^q)$ by

$$Q^{\otimes q}(f)(x^1, \ldots, x^q) := \int \left\{ \prod_{1 \leq i \leq q} Q(x^i, dy^i) \right\} f(y^1, \ldots, y^q).$$

We also denote by $Q_1 Q_2$ the composition of two operators defined by

$$(Q_1 Q_2)(x, dz) := \int Q_1(x, dy) Q_2(y, dz).$$

The Boltzmann–Gibbs transformation $\Psi_G : \eta \in \mathcal{P}(S) \mapsto \Psi_G(\eta) \in \mathcal{P}(S)$ associated with some positive function G on some state space S is defined by

$$\Psi_G(\eta)(dx) := \frac{1}{\eta(G)} G(x) \eta(dx).$$

We also denote by $\#(E)$ the cardinality of a finite set and we use the standard conventions $(\sup_{\varnothing}, \inf_{\varnothing}) = (-\infty, +\infty)$, and $(\sum_{\varnothing}, \prod_{\varnothing}) = (0, 1)$.

Finally, we recall the notion of the differential for sequences of measures introduced in [21]. We let μ^N be a uniformly bounded sequence of measures on some measurable state space S in the sense that $\sup_{N \geq 1} \|\mu^N\|_{\mathrm{tv}} < \infty$.

The sequence μ^N is said to converge strongly to some measure μ, as $N \uparrow \infty$, if $\lim_{N \uparrow \infty} \mu^N(f) = \mu(f)$ for any bounded measurable function f. In this case, the discrete derivative of μ^N is defined by

$$\partial \mu^N := N(\mu^N - \mu).$$

We say that μ^N is differentiable whenever $\partial \mu^N$ is uniformly bounded and strongly converges to some signed measure $d^{(1)}\mu$ as $N \uparrow \infty$. When $\partial \mu^N$ is differentiable, with a discrete derivative written as $\partial^{(2)}\mu^N$, we can define its derivative, denoted by $d^{(2)}\mu$, etc. A mapping $N \mapsto \mu^N$ that is differentiable up to some order l can be written as

$$\mu^N = \sum_{0 \le k \le l} \frac{1}{N^k} d^{(k)}\mu + \frac{1}{N^{l+1}} \partial^{(l+1)}\mu^N,$$

with the convention $d^{(0)}\mu = \mu$. We can easily extend these definitions to a sequence of integral operators Q^N and a sequence of functions f^N, where we denote the corresponding differentials by $d^{(l)}Q$ and $d^{(l)}f$.

2.2. *Feynman–Kac measures and their mean field interpretation*

We now consider a collection of non-negative bounded potential functions G_n on some measurable state spaces S_n, with $n \in \mathbb{N}$. We also let X_n be a Markov chain on S_n with initial distribution $\eta_0 \in \mathcal{P}(S_0)$ and some Markov transitions M_n from S_{n-1} into S_n. The Feynman–Kac measures (η_n, γ_n) associated with the parameters (G_n, M_n) are defined for any $f_n \in \mathcal{B}(S_n)$ by $\eta_n(f_n) := \gamma_n(f_n)/\gamma_n(1)$ with

$$\gamma_n(f_n) = \mathbb{E}(f_n(X_n)Z_n(X)) \quad \text{and} \quad Z_n(X) = \prod_{0 \le p < n} G_p(X_p). \tag{2.1}$$

We implicitly assume that $\gamma_n(1) > 0$. The evolution equations associated with these measures are

$$\gamma_{n+1} = \gamma_n Q_{n+1} \quad \text{and} \quad \eta_{n+1} = \Phi_{n+1}(\eta_n) := \Psi_{G_n}(\eta_n)M_{n+1}, \tag{2.2}$$

with the integral operators

$$Q_{n+1}(x_n, dx_{n+1}) = G_n(x_n)M_{n+1}(x_n, dx_{n+1}).$$

We will often assume below that $X_n = (X'_0, \ldots, X'_n)$ is the chain of trajectories associated with an auxiliary Markov chain X'_n with Markov transitions M'_n from S'_{n-1} to S'_n (so that $S_n = S'_0 \times \cdots \times S'_n$). In this context we will always assume that $G_n(X_n)$ only depends on the terminal state of the trajectory $(G_n(X_n) = G'_n(X'_n))$. All the quantities, measure, operators, etc. associated with the Markov chain X'_n are written with a prime. For example, η'_n is the Feynman–Kac measure defined recursively by

$$\eta'_{n+1} = \Psi_{G'_n}(\eta'_n)M'_{n+1}, \tag{2.3}$$

or, equivalently, as the nth marginal of η_n.

The unnormalized measures γ_n can be expressed in terms of the normalized ones using the well known product formula

$$\gamma_n(f_n) = \eta_n(f_n) \prod_{0 \le p < n} \eta_p(G_p).$$

We also recall the semigroup decompositions

$$\forall 0 \le p \le n \quad \gamma_n = \gamma_p Q_{p,n} \quad \text{and} \quad \eta_n = \eta_p \overline{Q}_{p,n}$$

with the integral operators $Q_{p,n} = Q_{p+1} \cdots Q_n$, and the normalized semigroups

$$\overline{Q}_{p,n}(f_n)(x_p) = Q_{p,n}(f_n)(x_p)/\eta_p Q_{p,n}(1) = (\overline{Q}_{p+1} \cdots \overline{Q}_n)(f_n)(x_p).$$

In the above equation, $\overline{Q}_{p+1}$ stands for the collection of integral operators defined as Q_{p+1} by replacing G_p with the normalized potential functions $\overline{G}_p = G_p/\eta_p(G_p)$. We will also often assume that

$$(G) \qquad \overline{g}_n = \sup_{0 \leq p \leq q \leq n} \left\| \overline{Q}_{p,q}(1) \right\| < \infty. \tag{2.4}$$

This condition is clearly met when $\tilde{g}_n =: \sup_{x,y}(G_n(x)/G_n(y)) < \infty$.

The mean field particle interpretation of the measures (η_n, γ_n) starts with N independent random variables $\xi_0 := (\xi_0^i)_{1 \leq i \leq N} \in S_0^N$ with common law η_0. The simplest way to evolve the population of N individuals (also known as samples, particle, or walkers) $\xi_n := (\xi_n^i)_{1 \leq i \leq N} \in S_n^N$ is to consider N conditionally independent individuals $\xi_{n+1} := (\xi_{n+1}^i)_{1 \leq i \leq N} \in S_{n+1}^N$ that have a common distribution

$$\Phi_{n+1}\big(m(\xi_n)\big) \quad \text{with } m(\xi_n) := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}. \tag{2.5}$$

The resulting particle model is a genetic type particle model with a selection and a mutation transition dictated by the potential function G_n and the Markov transition M_n.

Loosely speaking, the model works recursively as follows: If we start from a sample $\xi_0^{(N)}$ at $t = 0$ of the initial distribution η_0 (so that $m(\xi_0) \simeq_{N \uparrow \infty} \eta_0$), and assume that $m(\xi_n) \simeq_{N \uparrow \infty} \eta_n$, then the population at time $(n + 1)$ is formed with N "almost" independent samples with respect to η_{n+1} so that $m(\xi_{n+1}) \simeq_{N \uparrow \infty} \eta_{n+1}$. See [16] for details.

If X_n is the Markov chain of the trajectories of an auxiliary Markov chain X_n', then ξ_n is a set of trajectories, $\xi_n^i = (\xi_{0,n}^i, \ldots, \xi_{n,n}^i) \in S_n = (S_0' \times \cdots \times S_n')$. These trajectories will also be referred to as *ancestral lines* (of the $\xi_{n,n}^i$) as each trajectory ξ_n^i encodes the genealogy of the terminal particle $\xi_{n,n}^i$. The sequence $(\xi_0, \ldots, \xi_n)$ of all the ξ_p, from $p = 0$ to $p = n$, contains all the informations about the genealogies of the $\xi_{p,p}^i$, $p \leq n, i \leq N$. It will also be referred to as the *genealogical tree of the Feynman–Kac particle interpretation* ξ_n' *of the measures* (η_n', γ_n').

2.3. *Regularity conditions*

It is necessary to introduce regularity conditions to control the convergence of the mean field approximations of Feynman–Kac measures. We introduce a regularity hypothesis that (besides (G)) will be frequently used later in the article. It is well-suited, as is clear from its definition, to the study of ergodic properties of Feynman–Kac models, see e.g., [20] for further details.

This section assumes that X_n is the chain of trajectories of an auxiliary Markov chain X_n' and states the conditions for the latter chain.

We let $\eta_{p,n}'^{,X}$ be the solution of the equations (2.3) starting at the Dirac measure δ_x at time $p \leq n$, for some state $x \in S_p'$. The regularity condition, denoted as (H), is

$$(H) \qquad \sup_{p \geq 0} \sum_{n \geq p} \big(\tilde{g}_n' - 1\big)\beta_{p,n}' < \infty, \tag{2.6}$$

with

$$\beta_{p,n}' := \sup_{x,y} \left\| \eta_{p,n}'^{,x} - \eta_{p,n}'^{,y} \right\|_{\mathrm{tv}} \quad \text{and} \quad \tilde{g}_n' = \sup_{x,y}\big(G_n'(x)/G_n'(y)\big) < \infty,$$

where $\|\cdot\|_{\mathrm{tv}}$ is the total variation norm defined in Section 2.1.

Condition (H) is satisfied if $\sup_{n \geq 0} \tilde{g}_n' < \infty$ and $\beta_{p,n}' \leq ae^{-\lambda(n-p)}$, for some finite constants $0 < a, \lambda < \infty$. It is related to the stability properties of the limiting Feynman–Kac measures and ensures that local errors do not propagate with respect to the time horizon. For time homogeneous models with $\sup_{x,y}(G'(x)/G'(y)) < \infty$ (where $G' := G_n'$ and similarly for other symbols without lower indices), condition (H) is satisfied as soon as the Markov transition kernel M' of the chain X_n' satisfies the minorisation condition $M'^{,m}(x, dz) \geq \varepsilon M'^{,m}(y, dz)$ for some $m \geq 1$ and $\varepsilon > 0$, and for any $x, y \in S'$ (see, e.g., [18,27] for applications to nonlinear filtering).

In addition, $\sup_{n \geq 0} \overline{g}_n := \overline{g} < \infty$, where $\overline{g}_n$ is as in Equation (2.4), as soon as condition (H) is satisfied. Chapter 12.2.1 in [20] proves this result. In particular, (H) implies (G).

2.4. *Some illustrations*

Feynman–Kac measures appear in numerous scientific fields including signal processing, statistics, mathematical finance, rare event analysis, chemistry and statistical physics; see [8–10,16,18,20,24] and [29]. Their mean field interpretation depends on the application domain. We briefly outline two examples to illustrate.

2.4.1. *Nonlinear filtering*

Let $(X'_n, Y_n)_{n \geq 0}$ be a Markov chain on some product state space $(E_1 \times E_2)$ whose transition mechanism takes the form

$$\mathbb{P}\big((X'_n, Y_n) \in d(x, y) | (X'_{n-1}, Y_{n-1})\big) = K_n\big(X'_{n-1}, dx\big) g_n(y, x) \nu_n(dy), \tag{2.7}$$

where $(\nu_n)_{n \geq 0}$ is a sequence of positive measures on E_2, $(K_n)_{n \geq 0}$ is a sequence of Markov kernels from E_1 into itself, and $(g_n(\cdot, x))_{n \geq 0}$ is a sequence of conditional density functions on E_2. These filtering models are often described by a partially observed dynamic random system. For instance, when $E_1 = E_2 = \mathbb{R}$, these systems may take the common form

$$X'_n = a\big(X'_{n-1}\big) + W_n \quad \text{and} \quad Y_n = b\big(X'_n\big) + V_n, \tag{2.8}$$

with a sequence of independent random variables W_n and V_n, and some bounded functions a, b on $\mathbb{R}$. When the random variables V_n have some density $h_n(v)$ with respect to the Lebesgue measure dv on $\mathbb{R}$, (2.7) holds with $g_n(y, x) = h_n(y - b(x))$. The aim of filtering is to infer the trajectories $X_n = (X'_0, \ldots, X'_n)$ of the hidden Markov process X'_n given a series of observations Y_k from the origin $k = 0$, up to the current time $k = n$. Choosing

$$S_n = E_1, \qquad G_k(x_k) = G'_k\big(x'_k\big) := g_k\big(y_k, x'_k\big) \quad \text{and} \quad M_k = K_k \tag{2.9}$$

for any $x_k = (x'_0, \ldots, x'_k) \in S_n = \mathbb{R}^{n+1}$, it is easily checked that

$$\eta_n = \mathrm{Law}(X_n | Y_k = y_k, 0 \leq k < n) = \mathrm{Law}\big((X'_0, \ldots, X'_n) | Y_k = y_k, 0 \leq k < n\big).$$

2.4.2. *Hidden Markov chain models*

We further assume that the signal perturbations W_n in (2.8) are Gaussian centered random variables with an unknown variance θ. We take the prior distribution for θ as an inverse gamma distribution

$$\lambda(d\theta) := \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} 1_{\mathbb{R}_+(\theta)} \, d\theta,$$

with α and β its shape and scale parameters, where $\Gamma(\alpha)$ is the gamma function. The Hidden Markov chain problem consists of computing the conditional distributions

$$\overline{\eta}_n = \mathrm{Law}\big((\Theta, X_n) | Y_k = y_k, 0 \leq k < n\big). \tag{2.10}$$

A natural strategy when the priors on the parameters are conjugate is to use a Gibbs sampler. We end this section with a brief discussion of these MCMC methodologies. Section 4.5 provides further details on these statistical learning models and their mean field particle Gibbs samplers.

It is readily checked that the conditional distribution $\lambda_{n,X}$ of Θ given the historical process $X_n = x = (x'_0, \ldots, x'_n)$ (and the sequence of observations $Y_k = y_k, 0 \leq k < n$) is again an inverse-gamma distribution with shape and scale parameters

$$\alpha_n := \alpha + n/2 \quad \text{and} \quad \beta_n(x) := \beta + \frac{1}{2} \sum_{1 \leq k \leq n} \big(x'_k - a_k\big(x'_{k-1}\big)\big)^2.$$

It is straightforward to sample from this conditional distribution. On the other hand, arguing as in (2.9), we first observe that the conditional distribution $\eta_{\theta,n}$ of X_n given $\Theta = \theta$, and $Y_k = y_k$, $0 \leq k < n$ are defined by the Feynman–Kac model (2.1) with a reference Markov chain X_n with Markov transitions $M_{n,\theta}$ and the potential functions

$$G_{k,\theta}(x_k) = G'_{k,\theta}(x'_k) := \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}(y_k - b_k(x'_k))^2\right)$$

that depends on the parameter θ. We let $\mathbb{K}_{n,\theta}$ be a collection of Markov transitions with invariant measure $\eta_{n,\theta}$ on the path space S_n,

$$\overline{\mathbb{K}}_n\big((\theta, dx), d(\overline{\theta}, d\overline{x})\big) = \lambda_{n,X}(d\overline{\theta})\mathbb{K}_{n,\overline{\theta}}(x, d\overline{x}) \tag{2.11}$$

and denote by $\nu_n = \mathrm{Law}(X_n | Y_k = y_k, 0 \leq k < n)$ and $\lambda_n = \mathrm{Law}(\Theta | Y_k = y_k, 0 \leq k < n)$. We readily check that $\overline{\eta}_n = \overline{\eta}_n \overline{\mathbb{K}}_n$ is $\overline{\mathbb{K}}_n$ invariant; i.e.,

$$\int \overline{\eta}_n\big(d(\theta, x)\big)\overline{\mathbb{K}}_n\big((\theta, dx), d(\overline{\theta}, d\overline{x})\big) = \int \nu_n(dx)\lambda_{n,X}(d\overline{\theta})\mathbb{K}_{n,\overline{\theta}}(x, d\overline{x})$$

$$= \lambda_n(d\overline{\theta})\int \eta_{n,\overline{\theta}}(dx)\mathbb{K}_{n,\overline{\theta}}(x, d\overline{x}) = \overline{\eta}_n\big(d(\overline{\theta}, \overline{x})\big).$$

For more general hidden Markov chain problems with general priors, we can replace the distribution $\lambda_{n,X}$ in (2.11) by any collection of Markov chain transition $P_{n,X}(\theta, d\overline{\theta})$ s.t. $\lambda_{n,X} = \lambda_{n,X}P_{n,X}$, for any path x.

2.4.3. *Physics and chemistry*

Feynman–Kac models are widely in physics and chemistry to describe molecular systems. The Feynman–Kac measure η_n can be interpreted as the solution of a discrete-time approximation of an imaginary time Schrödinger equation. If we set I as the identity operator, then the Markov kernel $M \simeq_{\Delta t \downarrow 0} I + L\Delta t$ of the chain X_n corresponds to the discretization of a continuous-time stochastic process X_t with infinitesimal generator L, $G_n = e^{-V\Delta t}$, where V is a potential energy, and $t_{n+1} - t_n := \Delta t \ll 1$, is a discretization time-step associated with some time mesh $t_n = n\lfloor t/n \rfloor$. Replacing the chain X_n in the formulas by the random state of the discrete time approximation model X_{t_n}, we obtain

$$\eta_{t_n}(f) \propto \mathbb{E}\left(f(X_{t_0}, \ldots, X_{t_n})\exp\left\{-\sum_{0 \leq t_k < t_n} V(X_{t_k})(t_{k+1} - t_k)\right\}\right).$$

The marginal γ_t with respect to the terminal time t of the above measures is often defined, in a weak sense, by the imaginary time Schrödinger equation

$$\frac{d}{dt}\gamma_t(f) = \gamma_t\big(L^V(f)\big) \quad \text{with } L^V(f) = L(f) - Vf.$$

For a more thorough discussion of these continuous-time models and their applications in chemistry and physics, see [6,7,26,38,40,41,54], the recent monograph [20], and the references therein.

2.5. *Path space models*

To illustrate the generality of the Feynman–Kac models discussed above, we develop more systematically the links between a given state space and its path space. We replace the 5-tuple $(G_n, M_n, Q_n, S_n, X_n)$ by its path-space analog $(\mathbf{G}_n, \mathbf{M}_n, \mathbf{Q}_n, \mathbf{S}_n, \mathbf{X}_n)$. That is, in the constructions of the previous paragraph, each item of the first 5-tuple is going to be replaced by its path space analog: $\mathbf{X}_n$ is the historical process associated with X_n,

$$\mathbf{X}_n := (X_0, \ldots, X_n) \in \mathbf{S}_n := (S_0 \times \cdots \times S_n). \tag{2.12}$$

We write $\mathbf{M}_n$ for the Markov transition of $\mathbf{X}_n$. The functions $\mathbf{G}_n$ on $\mathbf{S}_n$ only depend on the last coordinate and are defined by $\mathbf{G}_n(\mathbf{X}_n) := G_n(X_n)$.

When X_n is already the path $X_n = (X'_0, \ldots, X'_n)$ of an auxiliary Markov chain, $\mathbf{X}_n$ is a path of paths:

$$\mathbf{X}_n = (X_0, \ldots, X_n) = \big((X'_0), (X'_0, X'_1), \ldots, (X'_0, \ldots, X'_n)\big).$$

This situation occurs later in the article.

In general, dotted symbols, where appropriate, (e.g., $M'_n, \eta'_n, \ldots$) refer to the auxiliary Markov chain X'_n, a bold symbol denotes an element, function, measure, etc. on the path space of X_n – as in the present paragraph. In particular, we let $(\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n, \boldsymbol{\xi}_n)$ be the Feynman–Kac measures and the particle model defined as $(\gamma_n, \eta_n, \xi_n)$, by replacing $(G_n, M_n, Q_n, S_n, X_n)$ by $(\mathbf{G}_n, \mathbf{M}_n, \mathbf{Q}_n, \mathbf{S}_n, \mathbf{X}_n)$. The two measures on the state space $\mathbf{S}_n$ are given for any $\mathbf{f}_n \in \mathcal{B}(\mathbf{S}_n)$ by $\boldsymbol{\eta}_n(\mathbf{f}_n) := \boldsymbol{\gamma}_n(\mathbf{f}_n)/\boldsymbol{\gamma}_n(\mathbf{1})$, with

$$\boldsymbol{\gamma}_n(\mathbf{f}_n) = \mathbb{E}\big(\mathbf{f}_n(\mathbf{X}_n) Z_n(X)\big). \tag{2.13}$$

The same observations hold as for the relationships between X_n and its auxiliary Markov chain X'_n: by construction, (γ_n, η_n) are the S_n marginals of the measures $(\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n)$. The mean field particle interpretation $\boldsymbol{\xi}_n = (\boldsymbol{\xi}_n^i)_{1 \leq i \leq N}$ of the path-space Feynman–Kac measures $(\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n)$ is defined in terms of path-valued particles

$$\forall 1 \leq i \leq N \quad \boldsymbol{\xi}_n^i = \big(\xi_{0,n}^i, \xi_{1,n}^i, \ldots, \xi_{n,n}^i\big) \in \mathbf{S}_n := (S_0 \times \cdots \times S_n). \tag{2.14}$$

The ith particle $\boldsymbol{\xi}_n^i$ can be seen as the line of ancestors $\xi_{p,n}^i$ of the ith individual $\xi_n^i := \xi_{n,n}^i$ at time n, at every ancestral level $0 \leq p \leq n$, with $1 \leq i \leq N$. The path space model $\boldsymbol{\xi}_n$ clearly coincides with the genealogical tree of the Feynman–Kac particle interpretation $\xi_n = (\xi_n^i)_{1 \leq i \leq N}$ of the measures (η_n, γ_n) defined in (2.5).

To distinguish as clearly as possible the three levels of Feynman–Kac models we adopt the following terminology.

Definition 2.1. *The 3-tuple $(\eta_n, \gamma_n, \xi_n)$ is called the Feynman–Kac particle model associated with the potential functions G_n and the Markov transitions M_n on the state spaces S_n. When $X_n = (X'_0, \ldots, X'_n)$, the 3-tuple $(\eta'_n, \gamma'_n, \xi'_n)$ is called the auxiliary particle model associated with $(\eta_n, \gamma_n, \xi_n)$. The path space model $(\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n, \boldsymbol{\xi}_n)$ is called the historical version of $(\gamma_n, \eta_n, \xi_n)$.*

Whenever the integral operators Q_n have some densities H_n with respect to some reference distributions υ_n on S_n, the path space measure $\boldsymbol{\eta}_n$ can be expressed in terms of the marginal measures $(\eta_p)_{0 \leq p \leq n}$ using the well known backward formula (see for instance [22])

$$\boldsymbol{\eta}_n(\mathbf{d}x_n) = \eta_n(dx_n) \prod_{1 \leq k \leq n} \mathbb{L}_{k, \eta_{k-1}}(x_k, dx_{k-1}) \tag{2.15}$$

with the collection of Markov transitions $\mathbb{L}_{n+1, \eta_n}$ from S_{n+1} into S_n defined by

$$\mathbb{L}_{n+1, \eta_n}(x_{n+1}, dx_n) = \eta_n(dx_n) H_{n+1}(x_n, x_{n+1})/\eta_n\big(H_{n+1}(\cdot, x_{n+1})\big). \tag{2.16}$$

In (2.15), $\mathbf{x}_n$ stands for the trajectory $\mathbf{x}_n = (x_0, \ldots, x_n) \in \mathbf{S}_n := (S_0 \times \cdots \times S_n)$. Finally, we recall that the mean field approximations of the Feynman–Kac measure $\boldsymbol{\gamma}_n$ are unbiased (see e.g., [16]):

Proposition 2.2. *Two unbiased estimators of $\boldsymbol{\gamma}_n$ are defined by*

$$\boldsymbol{\gamma}_n^{(N,i)} := \mathcal{Z}_n(\xi) \times \boldsymbol{\eta}_n^{(N,i)}, \quad i = 1, 2 \text{ with } \mathcal{Z}_n(\xi) := \prod_{0 \leq p < n} m(\xi_p)(G_p) \tag{2.17}$$

and the two random measures $(\boldsymbol{\eta}_n^{(N,1)}, \boldsymbol{\eta}_n^{(N,2)})$ on $\mathbf{S}_n$ are defined by

$$\boldsymbol{\eta}_n^{(N,1)}(\mathbf{d}x_n) := m(\boldsymbol{\xi}_n)(\mathbf{d}x_n) \quad and \quad \boldsymbol{\eta}_n^{(N,2)}(\mathbf{d}x_n) := m(\xi_n)(dx_n) \prod_{1 \leq k \leq n} \mathbb{L}_{k, m(\xi_{k-1})}(x_k, dx_{k-1}).$$

2.6. *Many-body Feynman–Kac models*

This section introduces a key ingredient of the developments to follow: the idea of many-body Feynman–Kac models. As the name "many-body" suggests, these Feynman–Kac models encode properly the collective motion under mean field constraints of the system of particles associated with a standard Feynman–Kac particle system.

In practice, we consider the *set* of particles of a standard Feynman–Kac particle model as a proper Markov chain. This simple idea allows us to recast the unbiasedness properties of mean field approximations of the Feynman–Kac measures, and is also be the key to the duality results below.

We fix the size N of the particle model, and set $\mathcal{S}_n := S_n^{[N]}$ for the Nth symmetric power of S_n: $S_n^{[N]} := (S_n \times \cdots \times S_n)/\Sigma_N = S_n^N/\Sigma_N$, where we write Σ_N for the symmetric group of order N. The image in $\mathcal{S}_n$ of an ordered sequence $(x_1, \ldots, x_n) \in S_n^N$ will sometimes be written with the set-theoretical notation $\{x_1, \ldots, x_n\}$ to emphasize that the order of the x_i does not matter, although we will also often identify $(x_1, \ldots, x_n)$ with its image in $S_n^{[N]}$ without further notice when no confusion can arise.

For example, with this slight abuse of notation, we note for later use that the particle model ξ_n can be viewed as a $\mathcal{S}_n$-valued Markov chain (since the distribution of the $\xi_n^i, i = 1, \ldots, N$, is Σ_N-invariant) we will have, for a function f on $\mathcal{S}_n$,

$$f(\xi_n) := f\big(\{\xi_n^1, \ldots, \xi_n^N\}\big) =: f\big(\xi_n^1, \ldots, \xi_n^N\big).$$

In general, it will be always implicitly assumed from now on that the functions on S_n^N (in practice, functions of population of particles of the models) are symmetric. Due to the symmetry properties of Feynman–Kac models, this does not result in a loss of generality.

Later in this section we use calligraphic letters such as x_n and $y_n = \{y_n^i\}_{1 \le i \le N}$ to denote states in the product spaces $\mathcal{S}_n = S_n^{[N]}$, and slanted roman letters such as x_n, y_n, z_n to denote states in S_n. The path sequences in the product spaces $\boldsymbol{\mathcal{S}}_n := \prod_{0 \le p \le n} \mathcal{S}_p$ and $\mathbf{S}_n := \prod_{0 \le p \le n} S_p$ are denoted by bold letters such as $\boldsymbol{x}_n = (x_p)_{0 \le p \le n} \in \boldsymbol{\mathcal{S}}_n$ and $\mathbf{x}_n = (x_p)_{0 \le p \le n} \in \mathbf{S}_n$.

We write $\mathcal{M}_n$ for the Markov transitions of the particle model $\mathcal{X}_n := \xi_n$ viewed now as a *Markov chain* on $\mathcal{S}_n$, and introduce the potential functions $\mathcal{G}_n(\mathcal{X}_n) = m(\mathcal{X}_n)(G_n)$. We let (Π_n, Γ_n) be the Feynman–Kac measures on $\mathcal{S}_n$ defined for any $\mathcal{F}_n \in \mathcal{B}(\mathcal{S}_n)$ by $\Pi_n(\mathcal{F}_n) := \Gamma_n(\mathcal{F}_n)/\Gamma_n(1)$, with

$$\Gamma_n(\mathcal{F}_n) = \mathbb{E}\big(\mathcal{F}_n(\mathcal{X}_n)\mathcal{Z}_n(\mathcal{X})\big) \quad \text{and} \quad \mathcal{Z}_n(\mathcal{X}) = \prod_{0 \le p < n} \mathcal{G}_p(\mathcal{X}_p). \tag{2.18}$$

Equivalently, the Feynman–Kac model on $\mathcal{S}_n$ can be turned into a Feyman–Kac model on $\mathcal{P}(S_n)$, by replacing in (2.18) the chain $\mathcal{X}_n$ by the occupation measure $m(\mathcal{X}_n)$.

We note that the unbiasedness properties of $\gamma_n^{(N,1)}(1)$ ensure that $\Gamma_n(1) = \gamma_n(1)$. Using (2.2), it is readily checked that

$$\Gamma_{n+1} = \Gamma_n \mathcal{Q}_{n+1} \quad \text{and} \quad \Pi_{n+1} := \Psi_{\mathcal{G}_n}(\Pi_n)\mathcal{M}_{n+1}, \tag{2.19}$$

with the integral operators

$$\mathcal{Q}_{n+1}(x_n, dx_{n+1}) = \mathcal{G}_n(x_n)\mathcal{M}_{n+1}(x_n, dx_{n+1}).$$

We denote by $(\boldsymbol{\Pi}_n, \boldsymbol{\Gamma}_n)$ the Feynman–Kac measures associated with the historical process $\boldsymbol{\mathcal{X}}_n = (\mathcal{X}_0, \ldots, \mathcal{X}_n)$, and the potential functions $\boldsymbol{\mathcal{G}}_n(\boldsymbol{\mathcal{X}}_n) := \mathcal{G}_n(\mathcal{X}_n)$ on the path space $\boldsymbol{\mathcal{S}}_n$. These measures are defined for any $\boldsymbol{\mathcal{F}}_n \in \mathcal{B}(\boldsymbol{\mathcal{S}}_n)$ by $\boldsymbol{\Pi}_n(\boldsymbol{\mathcal{F}}_n) := \boldsymbol{\Gamma}_n(\boldsymbol{\mathcal{F}}_n)/\boldsymbol{\Gamma}_n(\mathbf{1})$, with

$$\boldsymbol{\Gamma}_n(\boldsymbol{\mathcal{F}}_n) = \mathbb{E}\big(\boldsymbol{\mathcal{F}}_n(\boldsymbol{\mathcal{X}}_n)\mathcal{Z}_n(\mathcal{X})\big). \tag{2.20}$$

Whenever the integral operators Q_n have some densities H_n with respect to some reference distributions υ_n on S_n, given $\boldsymbol{\mathcal{X}}_n$ we let $\mathbb{X}_n^\flat$ be a random path with conditional distributions

$$\mathcal{K}_n^\flat(\boldsymbol{\mathcal{X}}_n, \mathbf{dx}_n) := m(\mathcal{X}_n)(dx_n) \prod_{1 \le k \le n} \mathbb{L}_{k,m(\mathcal{X}_{k-1})}(x_k, dx_{k-1}). \tag{2.21}$$

In (2.21), $\boldsymbol{x}_n$ stands for the path $\boldsymbol{x}_n = (x_p)_{0 \le p \le n} \in \mathbf{S}_n$, and $\mathbb{L}_{k,m(\chi_{k-1})}$ are the Markov transitions defined in (2.16). We also denote by $\mathbb{X}_n$ a random variable with conditional distribution given χ_n defined by

$$\mathcal{K}_n(\chi_n, dx_n) = m(\chi_n)(dx_n). \tag{2.22}$$

Lemma 2.3. *The unbiasedness property of the measure* $\boldsymbol{\gamma}_n^{(N,2)}$ *is equivalent to*

$$\mathbb{E}\big(\mathbf{f}_n\big(\mathbb{X}_n^\flat\big)\mathcal{Z}_n(\chi)\big) = \mathbb{E}\big(\mathbf{f}_n(\mathbf{X}_n)Z_n(X)\big), \tag{2.23}$$

for any $\mathbf{f}_n \in \mathcal{B}(\mathbf{S}_n)$. *The unbiasedness property of the measure* $\boldsymbol{\gamma}_n^{(N,1)}$ *implies that for any* $f_n \in \mathcal{B}(S_n)$,

$$\mathbb{E}\big(f_n(\mathbb{X}_n)\mathcal{Z}_n(\chi)\big) = \mathbb{E}\big(f_n(X_n)Z_n(X)\big). \tag{2.24}$$

We emphasize that (2.24) holds for general Feynman–Kac models, i.e. without any regularity conditions on Q_n.

Definition 2.4. *The measures* (Π_n, Γ_n) *and their path space versions* $(\boldsymbol{\Pi}_n, \boldsymbol{\Gamma}_n)$ *are called the many body Feynman–Kac measures associated with the particle interpretation* (2.5) *of the measures* (η_n, γ_n).

From an abstract point of view, in view of (2.24), all of these measures are essentially equivalent to the abstract Feynman–Kac model introduced in (2.1).

3. Particle Gibbs samplers: Presentation and statement of results

In this section we will always assume that X_n is the historical path of an auxiliary Markov chain X_n'; that is, $X_n = (X_0', \ldots, X_n') \in S_n = (S_0' \times \cdots \times S_n')$, for each $n \ge 0$.

3.1. *Frozen trajectories and the PG sampler*

We first describe the *freezing process*. The notation for Markov chains, state spaces, etc. is the same as in the previous section. We now fix the size of the particle system N as well as the time horizon $n \ge 0$.

Let some random path $X := (X_k)_{k \ge 0}$ be fixed. We let $\mathcal{X}_n$ be the N-particle Markov chain defined as χ_n in Section 2.6 except that we enforce it to contain one frozen state X_n at each time step. More precisely, we start with $(N-1)$ independent random variables $(\xi_0^i)_{2 \le i \le N} \in S_0^{N-1}$ with common law η_0, and set: $\mathcal{X}_0 := \{X_0, \xi_0^2, \ldots, \xi_0^N\}$. In general, given $\mathcal{X}_k = \{X_k, \xi_k^2, \ldots, \xi_k^N\}, 0 \le k < n$, we consider $(N-1)$ conditionally independent individuals $(\xi_{k+1}^i)_{2 \le i \le N} \in S_{k+1}^{N-1}$ having common distribution

$$\Phi_{k+1}\big(m(\mathcal{X}_k)\big) \quad \text{with } m(\mathcal{X}_k) := \frac{1}{N}\left(\delta_{X_k} + \sum_{2 \le i \le N} \delta_{\xi_k^i}\right), \tag{3.1}$$

and set $\mathcal{X}_{k+1} := \{X_{k+1}, \xi_{k+1}^2, \ldots, \xi_{k+1}^N\}$.

The key idea underlying PMCMC methods is that this freezing technique allows us to define a new Markov transition on the space of trajectories that (under some reasonable assumptions) has the targeted Feynman–Kac measure η_n as its invariant distribution. The standard MCMC method can then be used: a new Markov chain can be constructed iteratively using the new Markov transition to sample the invariant distribution η_n.

We now define the new Markov chain $\mathbb{Y}_p^{(N)} \in S_n, p \ge 0$. We start from a trajectory $x = (x_p)_{0 \le p \le n} \in \mathbf{S}_n, x_p = (x_0', \ldots, x_p') \in S_p = (S_0' \times \cdots \times S_p')$. We set $\mathbb{Y}_0^{(N)} := x_n \in S_n$ and run the particle model $(\mathcal{X}_k)_{0 \le k \le n}$ defined as above with the frozen trajectory x, up to a given time horizon n. We write, as usual, $\boldsymbol{\mathcal{X}}_n = (\mathcal{X}_0, \ldots, \mathcal{X}_n) \in \boldsymbol{\mathcal{S}}_n$.

For $\mathcal{X}_n = \{y^1, \ldots, y^N\} \in S_n^{[N]}$, we let $\mathbb{X}_n$ be a randomly chosen element in S_n with the uniform distribution $\frac{1}{N}\sum_{1 \le i \le N} \delta_{y^i}$, and set

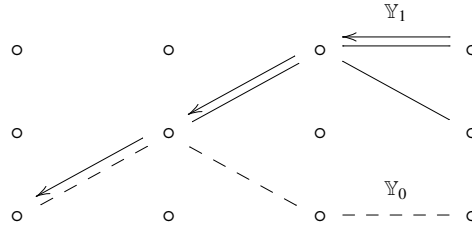$$\mathbb{Y}_1^{(N)} \sim \text{Law}(\mathbb{X}_n | \mathcal{X}_n).$$

Fig. 1. A realization of the transition $\mathbb{Y}_0 \rightsquigarrow \mathbb{Y}_1$ for $N = 3$ particles and a time horizon $n = 3$. The dotted and plain lines account together for the three paths in $\mathcal{X}_3$, the dotted line represents $\mathbb{Y}_0$, and the sequence of arrows $\mathbb{Y}_1$.

Specifically, at the terminal time n, we select uniformly in $\mathcal{X}_n$, viewed as a set of N particles in S_n (or, equivalently, N elements in the path space $(S'_0 \times \cdots \times S'_n)$) one element $\mathbb{Y}_1^{(N)} = y$. More generally, we let $(\mathbb{Y}_k^{(N)})_{k \geq 0}$ be a time-homogeneous Markov chain on S_n with transition probabilities

$$\mathbb{K}_n^{(N)}(f_n)(x) := \mathbb{E}\big(f_n\big(\mathbb{Y}_1^{(N)}\big)|\mathbb{Y}_0^{(N)} = x\big). \tag{3.2}$$

To simplify the notation, we suppress from now on the index $(\cdot)^{(N)}$ and write $\mathbb{Y}_k$ and $\mathbb{K}_n$, instead of $\mathbb{Y}_k^{(N)}$ and $\mathbb{K}_n^{(N)}$. A similar convention holds for the self-avoiding Gibbs sampler (whose transition operator is written as $\tilde{\mathbb{K}}_n$).

Figure 1 illustrates a realization of the transition $\mathbb{Y}_0 \rightsquigarrow \mathbb{Y}_1$.

For later use, it is convenient to summarize this sampling process graphically: although the chain runs ultimately on S_n, we incorporate the space $\mathcal{S}_n$ to the diagram and view the process as running on $(S_n \times \mathcal{S}_n)$. We obtain,

$$\left\{ \begin{array}{l} \mathbb{Y}_n = x \\ \mathcal{X}_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{Y}}_n = \overline{x} \sim (\mathbb{X}_n | \mathcal{X}_n = x) \\ \mathcal{X}_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{Y}}_n = \overline{x} \\ \overline{\mathcal{X}}_n = \overline{x} \sim (\mathcal{X}_n | X_n = \overline{x}) \end{array} \right\}.$$

The fundamental result of the particle Gibbs methodology is that the "marginal" chain obtained by forgetting the auxiliary particle system $\mathcal{X}_n$, i.e. the Markov chain whose Markov transition maps $\mathbb{Y}_n$ to $\overline{\mathbb{Y}}_n$, simulates the Feynman–Kac measure η_n (in the MCMC sense: iterating the chain p times yields an empirical measure that converges to η_n when $p \to \infty$). The goal of our article is to build on previous results in the literature and analyse these phenomena from the perspective of Feynman–Kac models. As we observed in the Introduction and the references there, in spite of a wide range of applications, the literature on the *theoretical* properties of these models is still very limited.

3.2. *The many-body Gibbs sampler*

We now return to the classical Feynman–Kac particle models, in their many-body version. Let $\chi_k := \{\chi_k^i\}_{1 \leq i \leq N}$ be the population after the kth mutation (initially we start with N independent copies of X_0) and write $\boldsymbol{\chi}_n = (\chi_k)_{0 \leq k \leq n} \in \mathcal{S}_n := \prod_{0 \leq k \leq n} \mathcal{S}_k$. With a slight abuse of the notation used in the previous section, we also let $\mathbb{X}_n$ be a random variable in S_n with the uniform distribution $\frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\chi_n^i}$.

Finally, we consider the probability distributions π_n on $(S_n \times \mathcal{S}_n)$ given for any bounded measurable function f_n on $(S_n \times \mathcal{S}_n)$ by the formula

$$\pi_n(f_n) \propto \mathbb{E}\big(f_n(\mathbb{X}_n, \boldsymbol{\chi}_n) \mathcal{Z}_n(\chi)\big)$$

with

$$\mathcal{Z}_n(\chi) = \prod_{0 \leq k < n} \mathcal{G}_k(\chi_k) \quad \text{with } \mathcal{G}_k(\chi_k) = \frac{1}{N} \sum_{1 \leq i \leq N} G_k(\chi_k^i).$$

The transition probabilities of the Gibbs sampler of the multivariate distribution π_n on the product space $(S_n \times \mathcal{S}_n)$ are described by the synthetic diagram

$$\left\{ \begin{array}{l} \mathbb{X}_n = x \\ \boldsymbol{\chi}_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{X}}_n = \overline{x} \sim (\mathbb{X}_n | \boldsymbol{\chi}_n = x) \\ \boldsymbol{\chi}_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{X}}_n = \overline{x} \\ \overline{\boldsymbol{\chi}}_n = \overline{x} \sim (\boldsymbol{\chi}_n | \mathbb{X}_n = \overline{x}) \end{array} \right\}.$$

The notation $(\mathbb{X}_n | \pmb{\chi}_n)$ and $(\pmb{\chi}_n | \mathbb{X}_n)$ is shorthand for the π_n-conditional distributions of $\mathbb{X}_n$ given $\pmb{\chi}_n$, and $\pmb{\chi}_n$ given $\mathbb{X}_n$.

The first transition of the Gibbs sampler reduces to uniformly sampling an ancestral line.

3.3. *First steps for obtaining duality and convergence properties*

Since the theoretical analysis of PMCMC models under strong regularity assumptions on the potential functions is already a challenging task from the Feynman–Kac perspective which we wish to develop, we decided not to obtain the most general form of the results. That is, to avoid an unnecessary technical discussion, we will often assume that the potential functions G_n are upper and lower bounded by some finite positive constant. However, in view of the existing results on the convergence of classical Feynman–Kac particle models, we expect the boundedness assumption on the potential functions can relaxed at the cost of an increased length and complexity of the proofs.

We note, for example, that the algebraic polynomial developments presented in this article involve the positivity of certain integral operators (indexed by infected and coalescent forests). In this context, these rather strong regularity properties of the potential functions can be relaxed so that these integral operators still have a finite norm. For instance, Lemma 5.18 allows us to extend the analysis to bounded potential functions which are not necessarily lower bounded. Besides, the extension of the results presented in this article to more general models, including indicator type functions and unbounded potential functions, can also be analyzed using the methodologies developed in [16] (see for instance Sections 2.3, 2.4, 3.5.2, and Section 7.2.2).

Our article expresses the rate of convergence to equilibrium of the PG chain $\mathbb{Y}_k$ presented above in terms of powers $(c(n)/N)^k$ of some ratio depending on some constant $c(n)$ that, in turn, depends on the time horizon n of the target Feynman–Kac measures (2.2). The non-asymptotic estimates derived in the article are valid for any bounded potential functions G_k and for any Markov chain X_k such that $\eta_n(G_n) > 0$. The constant $c(n)$ is expressed in terms of the norm of the potential functions and the quantities $\eta_k(G_k)$ (see, for instance the statement of Theorem 4.14).

Without any additional regularity conditions $c(n)$ grows exponentially with respect to the parameter n. One of the main purposes of the article is to show that the map $n \mapsto c(n)$ grows linearly with respect to the time n as soon as the Feynman–Kac model satisfies some natural stability conditions.

Next, we present a key duality formula that shows that the second transition of the many body Gibbs sampler amounts to sampling a Feynman–Kac particle model with a frozen ancestral line.

Theorem 3.1. *For any bounded measurable function f_n on the product space $(S_n \times \pmb{\mathcal{S}}_n)$, we have the duality formula*

$$\mathbb{E}\big(f_n(\mathbb{X}_n, \pmb{\chi}_n) \mathcal{Z}_n(\pmb{\chi})\big) = \mathbb{E}\big(f_n(X_n, \pmb{\mathcal{X}}_n) Z_n(X)\big), \tag{3.3}$$

where $\pmb{\mathcal{X}}_n = (\mathcal{X}_k)_{0 \leq k \leq n}$ stands for the historical version of a many body Feynman–Kac particle model with a given frozen trajectory X_n.

Consequently, the particle Gibbs chain $\mathbb{Y}_k$ with ancestral sampling defined in (3.2) coincides with the first coordinate of the many body Gibbs sampler of the target distribution π_n introduced in 3.2.

Section 4 proves the first assertion (see, e.g., Theorem 4.3 and Corollary 4.4 for Feynman–Kac models (2.2) on general state spaces S_n, and Section 4.2 for models associated with an historical process). This result shows that the chain $\mathbb{Y}_k$ reduce to the first coordinate of a pair of Gibbs samplers with target measure π_n. By (3.3), we conclude that $\mathbb{Y}_k$ is reversible with respect to the target measure η_n.

Besides stating this duality property, the following theorem gives a flavor for our forthcoming developments by also stating various consequences of stronger results when they are applied to the study of convergence properties of the particle Gibbs sampler.

Theorem 3.2. *The operators $\mathbb{K}_n$ are differentiable at any order $l \geq 1$ with $d^{(0)}\mathbb{K}_n(f_n) = \eta_n(f_n)$. If we further assume that the regularity condition* (H) *stated in* (2.6) *is satisfied, then for any bounded measurable function f on the path space S_n s.t. $\|f_n\| \leq 1$, the non-asymptotic estimates*

$$\forall 1 \leq k \leq l \quad \big\|d^{(k)}\mathbb{K}_n(f)\big\| \leq \big(cnk^2\big)^k \quad and \quad \big\|\partial^{(l+1)}\mathbb{K}_n(f)\big\| \leq \big(cn(l+1)^2\big)^{l+1}$$

as soon as $N > cn(l + 1)^2$, for some finite constant $c < \infty$.

Hence, there exists some finite constant $c < \infty$ such that for any $m \geq 1$, $x \in S_n$, and any $N > cn$,

$$\left| \mathbb{E}\big(f(\mathbb{Y}_m)|\mathbb{Y}_0 = x\big) - \eta_n(f) \right| \leq (cn/N)^m \quad (\longrightarrow_{\min(N,m)\to\infty} 0), \tag{3.4}$$

and for any $p \geq 1$, we have a sharp non-asymptotic estimate of the $\mathbb{L}_p(\eta_n)$-mean error norms

$$\left| \left\| \mathbb{K}_n^m(f) - \eta_n(f) \right\|_{\mathbb{L}_p(\eta_n)} - N^{-m} \left\| [d^{(1)}\mathbb{K}_n]^m(f) \right\|_{\mathbb{L}_p(\eta_n)} \right| \leq (cn/N)^{m+1},$$

with the mth iterates $\mathbb{K}_n^m = \mathbb{K}_n^{m-1}\mathbb{K}_n$ and $[d^{(1)}\mathbb{K}_n]^m := [d^{(1)}\mathbb{K}_n]^{m-1}d^{(1)}\mathbb{K}_n$ of the operators $\mathbb{K}_n$ and $d^{(1)}\mathbb{K}_n$.

This theorem is a direct consequence of the more general Theorem 4.14, describing non asymptotic Taylor series and their behaviour under regularity conditions.

The kth derivative integral operators $d^{(k)}\mathbb{K}_n$ are described in Section 5.4 in terms of Feynman–Kac semigroups parametrized by coalescent and decorated forests of length n with $(2k + 1)$ edges, and less than k coalescences and infections.

We note that (3.4) can be used to estimate the Lyapunov exponent of the distribution semigroup of the PMCMC chain $\mathbb{Y}_m$; i.e.,

$$\liminf_{m\to\infty} -\frac{1}{m} \log \left\| \text{Law}(\mathbb{Y}_m) - \eta_n \right\|_{\text{tv}} \geq \log \big(N/(cn)\big).$$

Sections 4.4 and 4.5 describe more precisely these non-asymptotic Taylor expansions, including a series of illustrations of the impact of these results on the estimation of the variance and the Dobrushin contraction coefficient of these models. See, e.g., Theorem 4.14 and the discussion that follows.

3.4. *Backward sampling*

We end this section with a discussion of PG models based on backward particle samplers (see [2]).

We further assume that the integral operators $G'_k(x)M'_{k+1}(x, dy)$ have a density $H'_{k+1}(x, y)$ with respect to some reference measure. In this context, and in order to apply the results of Section 2.5, we will view X_n as the historical version of X'_n and introduce some extra notation: given the many body model associated with X'_n, $\mathcal{X}'_k := \{\mathcal{X}'^i_k\}_{1\leq i\leq N} \in \mathcal{S}_k$, the historical process associated with these populations is written as

$$\boldsymbol{\mathcal{X}}'_n := (\mathcal{X}'_k)_{0\leq k\leq n} \in \boldsymbol{\mathcal{S}}'_n := \prod_{0\leq k\leq n} \mathcal{S}'_k.$$

We will call $\boldsymbol{\mathcal{X}}'_n$ the complete population of the Feynman–Kac particle model associated with X'_n.

Here, the Feynman–Kac measure η_n defined in (2.2) can be interpreted as the distribution of a nonlinear backward Markov chain model (see, for instance, (2.15) and (2.17) in Section 2.5). This backward particle model is a Markov chain running backward in time with the state spaces $\{\mathcal{X}'^i_k; 1 \leq i \leq N\}$ at each level k. The initial state of the chain takes the value $\mathcal{X}'^i_n$ with probability $1/N$, with $1 \leq i \leq N$. Then, at each level $0 \leq k < n$, the (conditional) probability of going from state $\mathcal{X}'^i_{k+1}$ to state $\mathcal{X}'^j_k$ is proportional to $H'_{k+1}(\mathcal{X}'^j_k, \mathcal{X}'^i_{k+1})$. Consistent with Section 2.6, we denote by $\mathbb{X}^\flat_n$ such a backward randomly chosen ancestral line.

Running the Feynman–Kac particle model with a given frozen path $\mathbb{Y}^\flat_0 := x = (x'_p)_{0\leq p\leq n} \in S_n$ up to a given time horizon n, we let $\mathbb{Y}^\flat_1 := y$ be an ancestral line randomly chosen with the backward Markov chain model discussed above (the initial value of the chain is one of the states at time n of the Feynman–Kac particle model with the frozen trajectory, including x'_n). Iterating this transition, we define a Markov chain $(\mathbb{Y}^\flat_k)_{k\geq 0}$ on S_n. Section 4.3 provides a more detailed description of the Markov transition of this backward particle MCMC model that we will call the *backward particle Gibbs sampler* (BPG sampler).

Figure 2 illustrates a realization of the transitions $\mathbb{Y}^\flat_0 \rightsquigarrow \mathbb{Y}^\flat_1$.
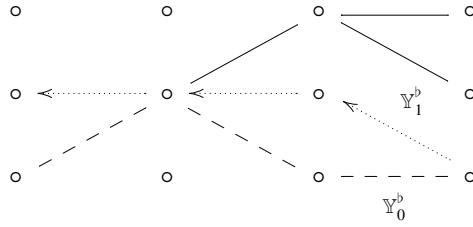
Fig. 2. Illustration of a realization of the transitions $\mathbb{Y}_0^\flat \rightsquigarrow \mathbb{Y}_1^\flat$ for $N = 3$ particles and a time horizon $n = 3$. The nodes stand for all the particles generated by the run of the model. The dotted line $\mathbb{Y}_0^\flat$ together with the plain lines account for the three paths generated by the run at time $n = 3$. $\mathbb{Y}_1^\flat$ is the ancestral line randomly chosen according to the backward model.
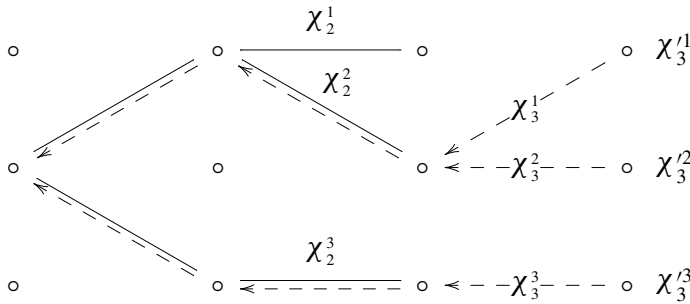

Fig. 3. A realization of the transitions $\chi_2 \rightsquigarrow \chi_3$ given χ_3' for $N = 3$ particles. Plain lines represent χ_2, dotted line χ_3. The law of the conditional transition is described in Theorem 4.8.

In the context of BPG models, one of our main results can be stated as follows. Here, $\mathbb{X}_n$ is the randomly chosen ancestral line defined in (2.22), and the many-body model χ_n' associated to X_n' is viewed in practice as a marginalization of the Feynman–Kac particle model $\xi_n \in S_n^N$ associated to X_n (so that, if $\xi_p = (\xi_p^i)_{i \leq N}$ and $\xi_p^i = (\xi_{0,p}^i, \ldots, \xi_{p,p}^i) \in S_p = (S_0' \times \cdots \times S_p')$, $p \leq n$, $\chi_p' = \{\xi_{p,p}^i\}_{i \leq N}$).

Theorem 3.3. *Given the complete populations χ_n', the ancestral lines $\xi_n^i = (\xi_{k,n}^i)_{0 \leq k \leq n}$ are copies of the backward trajectory $\mathbb{X}_n^\flat$ starting at the terminal state $\chi_n'^i = \xi_{n,n}^i$, with $1 \leq i \leq N$. That is,*

$$\mathrm{Law}\big(\mathbb{X}_n | \chi_n'\big) = \mathrm{Law}\big(\mathbb{X}_n^\flat | \chi_n'\big). \tag{3.5}$$

Consequently, the backward particle Gibbs sampler coincides with the Gibbs sampler targeting the $(S_n \times \boldsymbol{S}_n')$-marginal distribution π_n' of the measure π_n defined in (3.3).

Section 4.2 proves the equivalence formula (3.5) between the ancestral and the backward samplers. Theorem 4.8 in the same section also provides an interpretation of the conditional behaviour of $(\chi_k)_{0 \leq k \leq n}$ given the complete populations χ_n' in terms of a Markov chain with elementary transitions defined by backward ancestor sampling. Figure 3 illustrates a realization of the transitions $\chi_2 \rightsquigarrow \chi_3$ given χ_3'. Combining (3.3) and (3.5), we prove the duality formula

$$\pi_n'(f_n) \propto \mathbb{E}\big(f_n\big(\mathbb{X}_n^\flat, \chi_n'\big) \mathcal{Z}_n'(\chi')\big) = \mathbb{E}\big(f_n\big(X_n, \mathcal{X}_n'\big) Z_n'(X')\big) \tag{3.6}$$

for any bounded measurable function f_n on the product space (symmetric on the product spaces $S_k'^N$), where $\mathcal{X}_n' = (\mathcal{X}_k')_{0 \leq k \leq n}$ stands for the populations of the Feynman–Kac particle model with a given frozen trajectory X_n and $\mathcal{Z}_n', Z_n'$ are defined as usual.

The theorem shows that the Markov chain $\mathbb{Y}_k^\flat$ reduces to the first coordinate of a Gibbs sampler with target measure π_n'; by (3.6) and (3.3) we conclude that $\mathbb{Y}_k^\flat$ is reversible with respect to the target measure η_n.

Finally, by the duality formula (3.6), the transition probabilities of the Gibbs sampler of the multivariate distribution π'_n on the product space $(S_n \times \mathcal{S}'_n)$ are described by the synthetic diagram

$$\left\{ \begin{array}{l} \mathbb{X}^\flat_n = x \\ \pmb{\chi}'_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{X}}^\flat_n = \overline{x} \sim (\mathbb{X}^\flat_n | \pmb{\chi}'_n = x) \\ \pmb{\chi}'_n = x \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \overline{\mathbb{X}}^\flat_n = \overline{x} \\ \overline{\pmb{\chi}}'_n = \overline{x} \sim (\pmb{\chi}'_n | X_n = \overline{x}) \end{array} \right\}.$$

Above, $(\mathbb{X}^\flat_n | \pmb{\chi}'_n)$ and $(\pmb{\chi}'_n | X_n)$ is a shorthand notation for the π'_n-conditional distributions of $\mathbb{X}^\flat_n$ given $\pmb{\chi}'_n$, and $\pmb{\chi}'_n$ given $\mathbb{X}^\flat_n$.

4. Conditional particle Markov chain models

The purpose of this section is to explain PMCMC and, in particular, particle Gibbs sampling from the point of view of many body Feynman–Kac models. It contains most of the proofs of the results stated in the previous section in relation to the duality phenomena. Section 4.5 contains a number of remarks and implications of our work. These developments are only sketched out to avoid making our article excessively long.

4.1. *Transport equation for the many-body Feynman–Kac models*

We start the section with a pivotal duality formula between the Feynman–Kac integral operator Q_n and its many-body version $\mathcal{Q}_n$.

Lemma 4.1. *We have the duality formula between integral operators on $\mathcal{S}_n \times S_n$*

$$\mathcal{Q}_n(x_{n-1}, dx_n) m(x_n)(dx_n) = \big(m(x_{n-1}) Q_n\big)(dx_n) \mathcal{M}_{x_n,n}(x_{n-1}, dx_n) \tag{4.1}$$

and

$$\eta_0^{\otimes N}(dx_0) m(x_0)(dx_0) = \eta_0(dx_0) \mu_{x_0}(dx_0),$$

with the collection of Markov transitions

$$\mathcal{M}_{x_n,n}(x_{n-1}, dx_n) = \frac{1}{N} \left[\sum_{i=0}^{N-1} \Phi_n\big(m(x_{n-1})\big)^{\otimes(i)} \otimes \delta_{x_n} \otimes \Phi_n\big(m(x_{n-1})\big)^{\otimes(N-i-1)} \right](dx_n)$$

and the distribution

$$\mu_{x_0} := \frac{1}{N} \sum_{i=0}^{N-1} \big(\eta_0^{\otimes(i)} \otimes \delta_{x_0} \otimes \eta_0^{\otimes(N-i-1)} \big).$$

Proof. To prove (4.1) we use the symmetry properties of the Markov transitions $\mathcal{M}_n$ to check that for any function $H_n \in \mathcal{B}(S_n \times \mathcal{S}_n)$ (extended by right composition with the canonical projection from S_n^N to $\mathcal{S}_n$ to a function still written H_n in $\mathcal{B}(S_n \times S_n^N)$),

$$\int \mathcal{Q}_n(x_{n-1}, dx_n) m(x_n)(dz_n) H_n(z_n, x_n)$$

$$= \mathcal{G}_{n-1}(x_{n-1}) \int \Phi_n\big(m(x_{n-1})\big)^{\otimes N}(dx_n) H_n\big(x_n^1, x_n\big)$$

$$= m(x_{n-1})(G_{n-1}) \int \Phi_n\big(m(x_{n-1})\big)\big(dx_n^1\big) \big[\delta_{x_n^1} \otimes \Phi_n\big(m(x_{n-1})\big)^{\otimes(N-1)}\big](dy_n) H_n\big(x_n^1, y_n\big).$$

We complete the proof by using

$$m(x_{n-1})(G_{n-1})\Phi_n\big(m(x_{n-1})\big)(dx_n^1) = \big(m(x_{n-1})Q_n\big)(dx_n^1). \qquad \square$$

The following is a formal definition of Feynman–Kac particle models with a frozen trajectory introduced in the previous section.

Definition 4.2. *Given a random path* $(X_n)_{n\geq 0}$, *let* $\mathcal{X}_n = \{\mathcal{X}_n^i\}_{i=1,\dots,N} \in \mathcal{S}_n$ *be the Markov chain with the transitions* $\mathcal{M}_{X_n,n}$, *and the initial distribution* μ_{X_0} *introduced in Lemma* 4.1. *We denote by* $\mathbb{M}_n(\mathbf{X}_n, \mathbf{d}\mathbf{x}_n)$ *the conditional distributions of the random path* $\boldsymbol{\mathcal{X}}_n = (\mathcal{X}_p)_{0\leq p\leq n}$ *on* $\mathcal{S}_n$. *The process* $\mathcal{X}_n$ *is called the dual many body model associated with the Feynman–Kac particle model* χ_n *and the frozen path* $\mathbf{X}_n$.

The justification of the "duality" terminology between the processes $\mathcal{X}_n$ and χ_n is discussed at the end of the section. The Feynman–Kac measures $(\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n)$ and their many body version $(\boldsymbol{\Gamma}_n, \boldsymbol{\Pi}_n)$ are connected by the following duality theorem which can be seen as an extended version of the unbiasedness properties (2.24).

Theorem 4.3. *For any* $\mathbf{F}_n \in \mathcal{B}(S_n \times \mathcal{S}_n)$, *we have the duality formula*

$$\mathbb{E}\big(\mathbf{F}_n(\mathbb{X}_n, \boldsymbol{\mathcal{X}}_n)\mathcal{Z}_n(\chi)\big) = \mathbb{E}\big(\mathbf{F}_n(X_n, \boldsymbol{\mathcal{X}}_n)Z_n(X)\big). \tag{4.2}$$

If the integral operators Q_n *have some densities* H_n *with respect to some reference distributions* υ_n, *for any* $\mathbf{F}_n \in \mathcal{B}(\mathbf{S}_n \times \boldsymbol{\mathcal{S}}_n)$, *then we also have the duality formula*

$$\mathbb{E}\big(\mathbf{F}_n\big(\mathbb{X}_n^\flat, \boldsymbol{\mathcal{X}}_n\big)\mathcal{Z}_n(\chi)\big) = \mathbb{E}\big(\mathbf{F}_n(\mathbf{X}_n, \boldsymbol{\mathcal{X}}_n)Z_n(X)\big). \tag{4.3}$$

Proof. The proof of (4.2) is a direct consequence of (4.1). Indeed, using this formula, we find that

$$\mathcal{Q}_n(x_{n-1}, dx_n)m(x_n)(dz_n) = \big[m(x_{n-1})Q_n\big](dz_n)\mathcal{M}_{z_n,n}(x_{n-1}, dx_n)$$
$$= \int m(x_{n-1})(dz_{n-1})Q_n(z_{n-1}, dz_n)\mathcal{M}_{z_n,n}(x_{n-1}, dx_n)$$

and therefore

$$\mathcal{Q}_{n-1}(x_{n-2}, dx_{n-1})\mathcal{Q}_n(x_{n-1}, dx_n)m(x_n)(dz_n)$$
$$= \int m(x_{n-2})(dz_{n-2})Q_{n-1}(z_{n-2}, dz_{n-1})Q_n(z_{n-1}, dz_n)$$
$$\times \mathcal{M}_{z_{n-1},n-1}(x_{n-2}, dx_{n-1})\mathcal{M}_{z_n,n}(x_{n-1}, dx_n).$$

Iterating backwards in time, we obtain

$$\boldsymbol{\Gamma}_n(\mathbf{d}\mathbf{x}_n)m(x_n)(dz_n) = \int \eta_0(dz_0)Q_1(z_0, dz_1)\cdots Q_n(z_{n-1}, dz_n)\mathbb{M}_n(\mathbf{z}_n, \mathbf{d}\mathbf{x}_n).$$

This completes the proof of the first assertion.

The proof of (4.3) is a also direct consequence of (4.1). Using this formula, we obtain

$$\boldsymbol{\Gamma}_n(\mathbf{d}\mathbf{x}_n) \prod_{0\leq p\leq n} m(x_p)(dx_p)$$
$$= \mathcal{Z}_n(x)\eta_0^{\otimes N}(dx_0)m(x_0)(dx_0)\bigg\{ \prod_{1\leq p\leq n} \mathcal{M}_p(x_{p-1}, dx_p)m(x_p)(dx_p)\bigg\}$$

$$= \mathcal{Z}_n(x)\left\{\eta_0(dx_0)\prod_{1\leq p\leq n}\Phi_p\big(m(x_{p-1})\big)(dx_p)\right\}\mathbb{M}_n(\mathbf{x}_n,\mathbf{d}x_n)$$

$$= \left\{\eta_0(dx_0)\prod_{1\leq p\leq n}m(x_{p-1})\big(H_p(\cdot,x_p)\big)\upsilon_p(dx_p)\right\}\mathbb{M}_n(\mathbf{x}_n,\mathbf{d}x_n)$$

with $\mathcal{Z}_n(x) := \prod_{0\leq p<n}m(x_p)(G_p)$. The last assertion follows from

$$m(x_{p-1})(G_{p-1})\Phi_p\big(m(x_{p-1})\big)(dx_p) = m(x_{p-1})\big(H_p(\cdot,z_p)\big)\upsilon_p(dx_p).$$

We also have that,

$$\mathcal{K}_n^\flat(\mathbf{x}_n,\mathbf{d}x_n) := m(x_n)(dx_n)\prod_{1\leq p\leq n}\frac{m(x_{p-1})(dx_{p-1})H_p(x_{p-1},x_p)}{m(x_{p-1})(H_p(\cdot,x_p))},$$

where $\mathbf{x}_n$ stands for the path $\mathbf{x}_n = (x_p)_{0\leq p\leq n}\in\mathbf{S}_n$. Recalling that

$$Q_p(x_{p-1},dx_p) = G_p(x_{p-1})M_p(x_{p-1},dx_p) = H_p(x_{p-1},x_p)\upsilon_p(dx_p),$$

we obtain

$$\boldsymbol{\Gamma}_n(\mathbf{d}x_n)\mathcal{K}_n^\flat(\mathbf{x}_n,\mathbf{d}x_n)$$

$$= \left\{\eta_0(dx_0)\prod_{1\leq p\leq n}Q_p(x_{p-1},dx_p)\right\}\mathbb{M}_n(\mathbf{x}_n,\mathbf{d}x_n) = \boldsymbol{\gamma}_n(\mathbf{d}x_n)\mathbb{M}_n(\mathbf{x}_n,\mathbf{d}x_n).$$

This completes the proof of (4.3) and the theorem.   □

The following corollary is a direct consequence of (2.24) and (4.3). It provides an interpretation of the conditional distribution of the dual process $\boldsymbol{\mathcal{X}}_n$ with respect to a given frozen trajectory as a conditional many body Feynman–Kac model with respect to a random path $\mathbb{X}_n^\flat$ sampled with the backward distribution (2.21).

Corollary 4.4. *For any* $\mathbf{F}_n\in\mathcal{B}(\boldsymbol{\mathcal{S}}_n)$, *and for* $(\boldsymbol{\eta}_n\otimes\eta_n)$-*almost every paths* $(\mathbf{x}_n,x_n)$,

$$\mathbb{E}\big(\mathbf{F}_n(\boldsymbol{\mathcal{X}}_n)|X_n=x_n\big)\propto\mathbb{E}\big(\mathbf{F}_n(\boldsymbol{\chi}_n)\mathcal{Z}_n(\chi)|\mathbb{X}_n=x_n\big). \tag{4.4}$$

For any $(f_n,\mathbf{f}_n)\in(\mathcal{B}(S_n)\times\mathcal{B}(\mathbf{S}_n))$, *and for* $\boldsymbol{\Pi}_n$-*almost every path* $\mathbf{x}_n$,

$$\mathbb{E}\big(f_n(\mathbb{X}_n)|\boldsymbol{\chi}_n=\mathbf{x}_n\big)\propto\mathbb{E}\big(f_n(X_n)Z_n(X)|\boldsymbol{\mathcal{X}}_n=\mathbf{x}_n\big).$$

In addition, if the integral operators Q_n *have some densities* H_n *with respect to some reference distributions* υ_n, *then*

$$\mathbb{E}\big(\mathbf{F}_n(\boldsymbol{\mathcal{X}}_n)|\mathbf{X}_n=\mathbf{x}_n\big)\propto\mathbb{E}\big(\mathbf{F}_n(\boldsymbol{\chi}_n)\mathcal{Z}_n(\chi)|\mathbb{X}_n^\flat=\mathbf{x}_n\big) \tag{4.5}$$

and

$$\mathbb{E}\big(\mathbf{f}_n\big(\mathbb{X}_n^\flat\big)|\boldsymbol{\chi}_n=\mathbf{x}_n\big)\propto\mathbb{E}\big(\mathbf{f}_n(\mathbf{X}_n)Z_n(X)|\boldsymbol{\mathcal{X}}_n=\mathbf{x}_n\big).$$

We end this section with an analytic description of the duality formulae (4.2) and (4.3) in terms of the conditional distributions $\mathbb{M}_n$ and $\mathcal{K}_n^\flat$ introduced in Definition 4.2 and in (2.21). Using (4.2), we have

$$\forall\mathbf{x}_n\in\mathbf{S}_n\quad\mathbb{M}_n(\mathbf{x}_n,\cdot)\ll\eta_n\mathbb{M}_n=\boldsymbol{\Pi}_n.$$

Thus, we can define the dual operator $\mathbb{M}_{n,\eta_n}^\star$ of $\mathbb{M}_n$ from $\mathbb{L}_1(\boldsymbol{\eta}_n)$ into $\mathbb{L}_1(\boldsymbol{\Pi}_n)$, given for any $\mathbf{f}_n \in \mathbb{L}_1(\boldsymbol{\eta}_n)$, by

$$\mathbb{M}_{n,\eta_n}^\star(\mathbf{f}_n) = \frac{\mathbf{d}(\boldsymbol{\eta}_{n,\mathbf{f}_n}\mathbb{M}_n)}{\mathbf{d}(\boldsymbol{\eta}_n\mathbb{M}_n)} = \frac{\mathbf{d}(\boldsymbol{\eta}_{n,\mathbf{f}_n}\mathbb{M}_n)}{\mathbf{d}\boldsymbol{\Pi}_n} \quad \text{with } \boldsymbol{\eta}_{n,\mathbf{f}_n}(\mathbf{d}x_n) := \boldsymbol{\eta}_n(\mathbf{d}x_n)\mathbf{f}_n(x_n).$$

In addition, for any conjugate integers $\frac{1}{p} + \frac{1}{q} = 1$, with $1 \leq p, q \leq \infty$, and any pair of functions $(\mathbf{f}_n, \mathbf{F}_n) \in (\mathbb{L}_p(\boldsymbol{\eta}_n) \times \mathbb{L}_q(\boldsymbol{\Pi}_n))$,

$$\boldsymbol{\Pi}_n\big(\mathbf{F}_n\mathbb{M}_{n,\eta_n}^\star(\mathbf{f}_n)\big) = \boldsymbol{\eta}_n\big(\mathbb{M}_n(\mathbf{F}_n)\mathbf{f}_n\big). \tag{4.6}$$

These constructions show that formula (4.3) holds for general models, i.e. even if the integral operators Q_n do not have a density, where $\mathbb{X}_n^\flat$ is a random path with conditional distribution $\mathbb{M}_{n,\eta_n}^\star(\boldsymbol{\mathcal{X}}_n, \cdot)$ given the historical process $\boldsymbol{\mathcal{X}}_n$.

For a more detailed discussion on dual Markov transitions, see [17,53]. Conversely,

$$\forall x_n \in \boldsymbol{\mathcal{S}}_n, \quad \mathcal{K}_n^\flat(x_n, \cdot) \ll \boldsymbol{\Pi}_n\mathcal{K}_n^\flat = \boldsymbol{\eta}_n.$$

Thus (4.3) also implies that $\mathbb{M}_n$ coincides with the dual operator $\mathcal{K}_{n,\Pi_n}^{\flat\star}$ of $\mathcal{K}_n^\flat$ from $\mathbb{L}_1(\boldsymbol{\Pi}_n)$ into $\mathbb{L}_1(\boldsymbol{\eta}_n)$; i.e.,

$$(4.3) \quad \Longrightarrow \quad \boldsymbol{\Pi}_n\mathcal{K}_n^\flat = \boldsymbol{\eta}_n \quad \Longrightarrow \quad \boldsymbol{\eta}_n\big(\mathbf{f}_n\mathcal{K}_{n,\Pi_n}^{\flat\star}(F_n)\big) = \boldsymbol{\Pi}_n\big(\mathbf{F}_n\mathcal{K}_n^\flat(\mathbf{f}_n)\big),$$

with

$$\mathcal{K}_{n,\Pi_n}^{\flat\star}(\mathbf{z}_n, \mathbf{d}x_n) = \boldsymbol{\Pi}_n(\mathbf{d}x_n)\frac{\mathbf{d}\mathcal{K}_n^\flat(x_n, \cdot)}{d\boldsymbol{\Pi}_n\mathcal{K}_n^\flat}(\mathbf{z}_n) = \mathbb{M}_n(\mathbf{z}_n, \mathbf{d}x_n). \tag{4.7}$$

These results emphasize the duality between the random paths $\boldsymbol{\mathcal{X}}_n$ and $\mathbb{X}_n^\flat$ under the Feynman–Kac measures $\boldsymbol{\eta}_n$ and their many-body version $\boldsymbol{\Pi}_n$.

4.2. *Historical processes*

Suppose that $(\eta_n, \gamma_n, \boldsymbol{\mathcal{X}}_n, \mathcal{G}_n)$ is the historical version of an auxiliary Feynman–Kac model $(\gamma_n', \eta_n', \boldsymbol{\mathcal{X}}_n', \mathcal{G}_n')$ associated with some potential functions G_n' and some Markov chain X_n' with transitions M_n' on some state spaces S_n'.

In this situation, we recall that Φ_n' is the one step semigroup defined as Φ_n by replacing (G_n, M_n, S_n) by (G', M_n', S_n') in (2.2). In addition, $Z_n'(X') = Z_n(X)$ and $\mathcal{Z}_n'(\boldsymbol{\chi}') = \mathcal{Z}_n(\boldsymbol{\chi})$ the corresponding Radon–Nikodym derivatives defined in terms of $(G_n', \mathcal{G}_n', \boldsymbol{\chi}_n', X_n')$. Besides, when the integral operators Q_n' have some densities H_n' with respect to some reference distributions υ_n' on S_n', the measure η_n is expressed by a backward formula (2.15) with Markov transitions $\mathbb{L}_{n+1,\eta_n'}'$ defined as in (2.16) by replacing (η_n, H_n, S_n) by (η_n', H_n', S_n').

In this context, the Feynman–Kac models (η_n, γ_n) and (Π_n, Γ_n) defined in (2.1) and (2.18) are defined in terms of the historical process $X_n = \mathbf{X}_n' = (X_0', \ldots, X_n')$ of the chain X_n' and the ancestral lines $\boldsymbol{\chi}_n$ of the particle model $\boldsymbol{\chi}_n'$.

Notice that the dual process $\boldsymbol{\mathcal{X}}_n$ associated with the particle model $\boldsymbol{\chi}_n$ and the frozen path $\mathbf{X}_n = (X_0, \ldots, X_n)$ is a (many body) Markov chain on path space: at each time step, given $\boldsymbol{\mathcal{X}}_{n-1}$, we sample N random trajectories $\boldsymbol{\mathcal{X}}_n = (\boldsymbol{\mathcal{X}}_n^i)_{1 \leq i \leq N}$. One of them, say the first $\boldsymbol{\mathcal{X}}_n^1 = x_n$, takes the value of the frozen trajectory $X_n = x_n = (x_0', \ldots, x_n') \in S_n$. The others are (conditionally) independent random paths with common distribution $\Phi_n(m(\boldsymbol{\mathcal{X}}_{n-1}))$. This path-space chain can be interpreted as an evolution of a genealogical tree with a frozen ancestral line.

Definition 4.5. *We let $\mathbb{M}_n^\natural$ be the conditional expectation operator of the dual ancestral lines $\boldsymbol{\mathcal{X}}_n$ given the frozen path X_n, that is*

$$\mathbb{E}\big(F_n(\boldsymbol{\mathcal{X}}_n)|X_n\big) =: \mathbb{M}_n^\natural(F_n)(X_n) \tag{4.8}$$

for any function $F_n \in \mathcal{B}(\boldsymbol{\mathcal{S}}_n)$. We also denote by $(\mathbb{M}_n, \mathcal{K}_n^\flat)$ the Markov transitions defined as $(\mathbb{M}_n, \mathcal{K}_n^\flat)$ by replacing $(\mathbf{X}_n, \boldsymbol{\mathcal{X}}_n, \boldsymbol{\chi}_n)$ by the historical processes $(\mathbf{X}_n', \boldsymbol{\mathcal{X}}_n', \boldsymbol{\chi}_n')$ of the chains $(X_n', \boldsymbol{\mathcal{X}}_n', \boldsymbol{\chi}_n')$. We let $\mathbb{X}_n^\flat$ be a random path on S_n with conditional distributions $\mathcal{K}_n^\flat(\boldsymbol{\chi}_n', dx_n)$ with respect to the complete populations $\boldsymbol{\chi}_n' \in \boldsymbol{\mathcal{S}}_n' = \prod_{0 \leq k \leq n} S_k'$, with the product spaces $S_k' := S_k'^{[N]}$, for any $k \geq 0$.

In this context, the duality formulae stated in Theorem 4.3 take the following form.

Corollary 4.6. *For any* $\mathbf{F}_n \in \mathcal{B}(S_n \times \boldsymbol{\mathcal{S}}_n)$,

$$\mathbb{E}\big(\mathbf{F}_n(\mathbb{X}_n, \boldsymbol{\chi}_n) \mathcal{Z}'_n(\chi')\big) = \mathbb{E}\big(\mathbf{F}_n(X_n, \boldsymbol{\mathcal{X}}_n) Z'_n(X')\big). \tag{4.9}$$

In addition, when the integral operators Q'_n *have some densities* H'_n *with respect to some reference distributions* υ'_n *on* S'_n,

$$\mathbb{E}\big(F_n(\mathbb{X}_n^\flat, \boldsymbol{\chi}'_n) \mathcal{Z}'_n(\chi')\big) = \mathbb{E}\big(F_n(X_n, \boldsymbol{\mathcal{X}}'_n) Z'_n(X')\big) = \mathbb{E}\big(F_n(\mathbb{X}_n, \boldsymbol{\chi}'_n) \mathcal{Z}'_n(\chi')\big). \tag{4.10}$$

The formulae (4.9) and (4.10) are direct consequences of (4.2) and (4.3).

To further develop this section, we assume that the integral operators Q'_n have some densities H'_n with respect to some reference distributions υ'_n on S'_n.

The next proposition follows from (4.10).

Proposition 4.7.

$$\mathrm{Law}\big(\mathbb{X}_n^\flat, \boldsymbol{\chi}'_n\big) = \mathrm{Law}\big(\mathbb{X}_n, \boldsymbol{\chi}'_n\big) \quad \textit{and} \quad \mathrm{Law}\big(\mathbb{X}_n | \boldsymbol{\chi}'_n\big) = \mathrm{Law}\big(\mathbb{X}_n^\flat | \boldsymbol{\chi}'_n\big). \tag{4.11}$$

In particular, given the complete populations $\boldsymbol{\chi}'_n$, *the ancestral lines* $\chi_n^i = (\chi_{k,n}^i)_{0 \le k \le n}$ *are* N *copies of the backward ancestral line* $\mathbb{X}_n^\flat$ *starting at* $\chi_n^{\prime i} = \chi_{n,n}^i$, *with* $1 \le i \le N$.

The next result provides a more detailed description of the conditional distribution of the genealogical trees given the complete populations.

Theorem 4.8. *Given the complete populations* $\boldsymbol{\chi}'_n$, *the sequence of genealogical trees* $(\chi_k)_{0 \le k \le n}$ *is a Markov chain starting at* $\chi_0 = \chi'_0$. *The elementary transitions of the ancestral lines* $\chi_k \rightsquigarrow \chi_{k+1}$ *given the population* χ'_{k+1} *are defined for any* $f \in \mathcal{B}(\mathcal{S}_{k+1})$

$$\mathbb{E}\big(f(\chi_{k+1}) | \chi_k, \chi'_{k+1}\big)$$
$$\propto \int \bigg\{ \prod_{1 \le i \le N} m(\chi_k)\big(dx_k^i\big) H'_{k+1}\big(x_{k,k}^i, \chi_{k+1}^{\prime i}\big) \bigg\} f\big((x_k^l, \chi_{k+1}^{\prime l})_{1 \le l \le N}\big),$$

where $x_k^i := (x_{l,k}^i)_{0 \le l \le k}$ *stands for an ancestral line of length* k.

Proof. By construction, for any $f_1, f_2 \in \mathcal{B}(\mathcal{S}_{k+1})$,

$$\mathbb{E}\big(f_1(\chi_{k+1}) f_2(\chi_k, \chi'_{k+1}) | \chi_k\big)$$
$$\propto \int \bigg\{ \prod_{1 \le i \le N} m(\chi_k)\big(dx_k^i\big) Q'_{k+1}\big(x_{k,k}^i, dx_{k+1}^{\prime i}\big) \bigg\}$$
$$\times f_1\big((x_k^j, x_{k+1}^{\prime j})_{1 \le j \le N}\big) f_2\big(\chi_k, (x_{k+1}^{\prime j})_{1 \le j \le N}\big).$$

Using the result,

$$m(\chi_k)\big(dx_k^i\big) Q'_{k+1}\big(x_{k,k}^i, dx_{k+1}^{\prime i}\big)$$
$$= \frac{m(\chi_k)(dx_k^i) H'_{k+1}(x_{k,k}^i, x_{k+1}^{\prime i})}{m(\chi'_k)(H'_{k+1}(\cdot, x_{k+1}^{\prime i}))} \times m\big(\chi'_k\big)\big(H'_{k+1}(\cdot, x_{k+1}^{\prime i})\big) \upsilon_{k+1}\big(dx_{k+1}^{\prime i}\big)$$
$$\propto \frac{m(\chi_k)(dx_k^i) H'_{k+1}(x_{k,k}^i, x_{k+1}^{\prime i})}{m(\chi'_k)(H'_{k+1}(\cdot, x_{k+1}^{\prime i}))} \times \Phi'_{k+1}\big(m(\chi'_k)\big)\big(dx_{k+1}^{\prime i}\big),$$

we conclude that

$$
\mathbb{E}\big(f_1(\chi_{k+1})f_2\big(\chi_k,\chi'_{k+1}\big)|\chi_k\big)
$$
$$
= \int \left\{ \prod_{1 \le i \le N} \varPhi'_{k+1}\big(m(\chi'_k)\big)\big(dx'^i_{k+1}\big)\right\} f_2\big(\chi_k,(x'^j_{k+1})_{1 \le j \le N}\big)
$$
$$
\times \int \left\{ \prod_{1 \le i \le N} \frac{m(\chi_k)(dx^i_k)H'_{k+1}(x^i_{k,k},x'^i_{k+1})}{m(\chi'_k)(H'_{k+1}(\cdot,x'^i_{k+1}))} \right\} f_1\big((x^j_k,x'^j_{k+1})_{1 \le j \le N}\big). \qquad \square
$$

We end this section with an elementary proof of the left-hand side backward duality formula in (4.10). Theorem 4.8 implies that

$$
\mathrm{Law}\big(\mathbb{X}_n|\boldsymbol{\chi}'_n\big) = \mathrm{Law}\big(\mathbb{X}^\flat_n|\boldsymbol{\chi}'_n\big).
$$

Combining this result with the duality formula stated in the right-hand side of (4.10), we obtain

$$
\mathbb{E}\big(F_n\big(X_n,\boldsymbol{\mathcal{X}}'_n\big)Z'_n\big(X'\big)\big) = \mathbb{E}\big(\mathbb{E}\big(F_n\big(\mathbb{X}_n,\boldsymbol{\chi}'_n\big)|\boldsymbol{\chi}'_n\big)\mathcal{Z}'_n\big(\chi'\big)\big)
$$
$$
= \mathbb{E}\big(F_n\big(\mathbb{X}^\flat_n,\boldsymbol{\chi}'_n\big)\mathcal{Z}'_n\big(\chi'\big)\big).
$$

4.3. *Genealogy and backward sampling models*

We now again assume that (η_n,γ_n) is the historical version of an auxiliary Feynman–Kac model (γ'_n,η'_n).

Definition 4.9. *We consider the Markov transitions from S_n into itself defined by $\mathbb{K}_n := \mathbb{M}^\natural_n \mathcal{K}_n$, with the operators $(\mathbb{M}^\natural_n,\mathcal{K}_n)$ introduced in (4.8) and in (2.22). In other words, for any function f_n on S_n and any frozen trajectory $x_n \in S_n$, we have*

$$
\mathbb{K}_n(f_n)(x_n) = \mathbb{E}\big(m(\mathcal{X}_n)(f_n)|X_n = x_n\big).
$$

When the integral operators Q'_n have some densities H'_n with respect to some distributions υ'_n, we consider the Markov transition from S_n into itself defined by $\mathbb{K}^\flat_n := \mathbb{M}_n \mathcal{K}^\flat_n$, with the operators $(\mathbb{M}_n,\mathcal{K}^\flat_n)$ introduced in Section 4.2.

Proposition 4.10. *The Markov transitions $\mathbb{K}_n$ and $\mathbb{K}^\flat_n$ are reversible with respect to the probability measures η_n.*

The reversibility property follows directly because the PG and BPG samplers reduce to a standard Gibbs sampler of a many-body Feynman–Kac distribution.

Next, under some rather strong regularity conditions, we present an elementary proof of the ergodicity of the two PG transitions discussed above based on the duality formulae presented in Section 4.

Proposition 4.11. *Suppose that the potential functions G_n are lower and upper bounded by some positive constant, and define $g_n := \sup_{x,y} G_n(x)/G_n(y)$. The measure η_n is the unique invariant measures of the Markov transitions $\mathbb{K}_n$ and $\mathbb{K}^\flat_n$. In addition, we obtain the estimates*

$$
\beta(\mathbb{K}_n) \vee \beta\big(\mathbb{K}^\flat_n\big) \le 1 - \tau_n\left(1 - \frac{1}{N}\right)^n \quad \text{with } \tau_n = 1\Big/ \prod_{0 \le p < n} g_p. \tag{4.12}
$$

The estimate of $\beta(\mathbb{K}^\flat_n)$ is a direct consequence of the following rather crude uniform minorization condition

$$
\mathbb{K}_n(f_n)(x_n) \wedge \mathbb{K}^\flat_n(f_n)(x_n) \ge \tau_n\left(1 - \frac{1}{N}\right)^n \eta_n(f_n) \tag{4.13}
$$

for any non-negative function f_n on S_n, and any path sequence $x_n = (x'_p)_{0 \le p \le n}$. These lower bounds are easily checked by induction with respect to the time parameter. By construction, for any $\mathbf{x}_n = (x_{n-1}, x'_n) \in S_n = (S_{n-1} \times S'_n)$,

$$\mathbb{K}_n^\flat(f_n)(x_n) \ge g_{n-1}^{-1}(1 - 1/N)\mathbb{K}_{n-1}^\flat\big(\overline{Q}_n(f_n)\big)(x_{n-1}).$$

We prove similarly that

$$\mathbb{K}_n(f_n)(x_n) \ge g_{n-1}^{-1}(1 - 1/N)\mathbb{E}\big(m(\mathcal{X}_{n-1})\big(\overline{Q}_n(f_n)\big)|X_{n-1} = x_{n-1}\big)$$
$$= g_{n-1}^{-1}(1 - 1/N)\mathbb{K}_{n-1}\big(\overline{Q}_n(f_n)\big)(x_{n-1}).$$

We obtain (4.13) by iterating these estimates.

These rather crude estimates are not new. Similar estimates are obtained in [3,11,12,42] using different approaches. As shown in [3,42], the assumption that G_n is lower bounded is unecessary, although [3] shows that the assumption that G_n is upper bounded is essentially necessary for a uniform minorization condition to hold. When the function G_n can take null values, the particles are killed instantly when they hit the set with null potential values. It may also happen that all the particles hit this hard-type obstacle set at the same time. These general Feynman–Kac models can be analyzed using the techniques developed in Section 2.4.2 and in Section 7.4.1 in [16] (see also [23] for particle mean field model locally constrained models avoiding the null level sets $G_n^{-1}(\{0\})$).

Sharp estimates of the contraction properties of $\mathbb{K}_n$ and *its iterates* $\mathbb{K}_n^m$, with $m \ge 1$, are developed in Section 4.4. These quantitative estimates are based on new Taylor type expansions of the PMCMC transitions around the limiting invariant measure η_n with respect to the precision parameter $1/N$.

4.4. *Taylor type expansions around the invariant measure*

We now assume that $(\eta_n, \gamma_n, \xi_n)$ is the historical version of an auxiliary Feynman–Kac model $(\gamma'_n, \eta'_n, \xi'_n)$. Our first objective is to find a Taylor type expansion of the Markov transition $\mathbb{K}_n$ around its invariant measure η_n with respect to powers of $1/N$. We fix the time horizon n and a frozen trajectory $z_n := (z'_0, \dots, z'_n) \in S_n = (S'_0 \times \cdots \times S'_n)$, and for any $0 \le p \le n$ we set $z_p := (z'_0, \dots, z'_p) \in S_p$.

We denote by $\mathcal{X}_{z_n,n}$ the dual many body model associated with the Feynman–Kac particle model χ_n and the frozen path $X_n = z_n$. Using the exchangeability properties of the dual particles, there is no loss of generality in assuming that only the first particle $\mathcal{X}_{z_n,n}^1 = X_n$ is frozen. With this convention, for any function $f_n \in \mathcal{B}(S_n)$,

$$\mathbb{K}_n(f_n)(z_n) = \mathbb{E}\big(m(\mathcal{X}_{z_n,n})(f_n)\big) = \frac{1}{N} f_n(z_n) + \left(1 - \frac{1}{N}\right)\mathbb{E}\big(m(\mathcal{X}_{z_n,n}^-)(f_n)\big),$$

where $m(\mathcal{X}_{z_n,n}^-)$ stands for the occupation measure of the non frozen particles $m(\mathcal{X}_{z_n,n}^-) := \frac{1}{N-1}\sum_{1 < i \le N}\delta_{\mathcal{X}_{z_n,n}^i}$. This shows that whenever they exists these Taylor expansions are related to the bias and the fluctuations of the measures $m(\mathcal{X}_{z_n,n}^-)$. To analyze these properties we observe that

$$\mathbb{E}\big(m(\mathcal{X}_{z_n,n})(f_n)|\mathcal{X}_{z_{n-1},n-1}\big) = \Phi_{z_n,n}\big(m(\mathcal{X}_{z_{n-1},n-1})\big)(f_n),$$

with the one step transformations $\Phi_{\mathbf{z}_n,n}$ defined as Φ_n by replacing the Markov transitions M_n by

$$M_{z_n,n}(x_{n-1}, dx_n) = \frac{1}{N}\delta_{z_n}(dx_n) + \left(1 - \frac{1}{N}\right)M_n(x_{n-1}, dx_n).$$

In addition, the occupation measures $m(\mathcal{X}_{z_n,n}^-)$ of all the particles, except the first frozen one, are based on $(N-1)$ conditionally independent random states with common law $\Phi_n(m(\mathcal{X}_{z_{n-1},n-1}))$. Thus, the local fluctuations of $m(\mathcal{X}_{z_n,n})$ around $\Phi_{z_n,n}(m(\mathcal{X}_{z_{n-1},n-1}))$ can be expressed in terms of the local sampling random fields

$$V_n^N := \sqrt{N-1}\big[m\big(\mathcal{X}_{z_n,n}^-\big) - \Phi_n\big(m(\mathcal{X}_{z_{n-1},n-1})\big)\big]$$

with the formula

$$m(\mathcal{X}_{z_n,n})(f_n) = \Phi_{z_n,n}\big(m(\mathcal{X}_{z_{n-1},n-1})\big) + \left(1 - \frac{1}{N}\right)\frac{1}{\sqrt{N-1}}V_n^N.$$

Proposition 4.12. *Let $X_{z_n,n}$ stand for a Markov chain on S_n, with initial distribution $\eta_{z_0,0} = \frac{1}{N}\delta_{z_0} + (1 - \frac{1}{N})\eta_0$ and Markov transitions $M_{z_n,n}$ from S_{n-1} into S_n. Then,*

$$\mathbb{E}\left(m(\mathcal{X}_{z_n,n})(f_n)\prod_{0 \le p < n} m(\mathcal{X}_{z_p,p})(G_p)\right) = \mathbb{E}\left(f_n(X_{z_n,n})\prod_{0 \le p < n} G_p(X_{z_p,p})\right). \tag{4.14}$$

The proof is similar to showing that $\boldsymbol{\gamma}_n^{(N,1)}$ is an unbiased approximation of $\boldsymbol{\gamma}_n$ as in [16], and is omitted.

The right-hand side Feynman–Kac measure in (4.14) can be expressed in terms of powers of the precision parameter $1/N$. To describe these models, let ε_n be a sequence of independent $\{0, 1\}$-valued random variables with $\mathbb{P}(\varepsilon_n = 1) = 1/N$. For any $\varepsilon = (\varepsilon_p)_{0 \le p \le n} \in \{0, 1\}^{n+1}$, we set $X_{z_n,n}^{(\varepsilon)}$ be a Markov chain on S_n, with initial distribution $\eta_{z_0,0}^{(\varepsilon)}$ and Markov transitions $M_{z_n,n}^{(\varepsilon)}$ defined by

$$\eta_{z_0,0}^{(\varepsilon)} = \varepsilon_0\delta_{z_0} + (1 - \varepsilon)\eta_0,$$
$$M_{z_n,n}^{(\varepsilon)}(x_{n-1}, dx_n) = \varepsilon_n\delta_{z_n}(dx_n) + (1 - \varepsilon_n)M_n(x_{n-1}, dx_n).$$

In this notation, we can readily check that

$$\mathbb{E}\left(f_n(X_{z_n,n})\prod_{0 \le p < n} G_p(X_{z_p,p})\right)$$
$$= \left(1 - \frac{1}{N}\right)^{(n+1)}\gamma_n(f_n) + \sum_{1 \le p \le n+1}\left(\frac{1}{N}\right)^p\left(1 - \frac{1}{N}\right)^{(n+1)-p}\sum_{\varepsilon_0+\cdots+\varepsilon_n=p}\mathbb{E}\left[f_n\big(X_{z_n,n}^{(\varepsilon)}\big)\prod_{0 \le p < n} G_p\big(X_{z_p,p}^{(\varepsilon)}\big)\right].$$

These decompositions can be easily turned into Taylor type polynomial expansions in powers of $1/N$. The Taylor expansion of the normalized Feynman–Kac measures with the 0th order measure η_n follows standard arguments on quotient power series.

The next proposition is easily proved using rather standard stochastic perturbation techniques (cf. for instance [16, 20]).

Proposition 4.13. *The random fields $\sqrt{N}[m(\mathcal{X}_{z_n,n}) - \eta_n]$ and $\sqrt{N}[m(\xi_n) - \eta_n]$ converge in law as $N \uparrow \infty$ to the same Gaussian and centered random fields. The same property holds true for the random fields associated with the unnormalized particle measures. In addition, for any function $f_n \in \mathcal{B}(S_n)$ s.t. $\eta_n(f_n) = 0$, and any frozen trajectory $z_n = (z_p')_{0 \le p \le n} \in S_n = \prod_{0 \le p \le n} S_p'$ we obtain the asymptotic bias expansion*

$$\lim_{N \uparrow \infty} N\mathbb{K}_n(f_n)(z_n) = \sum_{0 \le p \le n}\eta_p\big(\overline{Q}_{p,n}(1)\big[\overline{Q}_{p,n}(f_n)(z_p) - \overline{Q}_{p,n}(f_n)\big]\big) \tag{4.15}$$

with $z_p := (z_0', \ldots, z_p') \in S_p$, for any $p \le n$.

We proceed by analyzing the propagation properties of the non-frozen particles.

Theorem 4.14. *If the condition* (G) *stated in* (2.4) *is satisfied then there exists some finite constant $c > 0$ such that for any $n \ge 0$, $m \ge 1$ and $N > cn(m\overline{g}_n)^2$,*

$$\mathbb{K}_n(z_n, dy_n) = \eta_n(dy_n) + \sum_{1 \le k \le m}\frac{1}{N^k}d^{(k)}\mathbb{K}_n(z_n, dy_n) + \frac{1}{N^{m+1}}\partial^{(m+1)}\mathbb{K}_n(z_n, dy_n). \tag{4.16}$$

In (4.16), $d^{(k)}\mathbb{K}_n$, and $\partial^{(m+1)}\mathbb{K}_n$ *stand for some sequence of signed and bounded integral operators such that*

$$d^{(k)}\mathbb{K}_n(1)(z_n) = \partial^{(m+1)}\mathbb{K}_n(1)(z_n) = 0 = \eta_n\big(d^{(k)}\mathbb{K}_n(f_n)\big) = \eta_n\big(\partial^{(m+1)}\mathbb{K}_n(f_n)\big), \tag{4.17}$$

for any function f_n on the path space S_n, and

$$\beta\big(d^{(k)}\mathbb{K}_n\big) \leq \big[cn(k\overline{g}_n)^2\big]^k \quad \text{and} \quad \beta\big(\partial^{(m+1)}\mathbb{K}_n\big) \leq \big(cn\big((m+1)\overline{g}_n\big)^2\big)^{m+1}. \tag{4.18}$$

In addition, if the Feynman–Kac model (γ_n', η_n') satisfies the regularity condition (H) *stated in* (2.6), *then the above estimate remains valid by replacing $\overline{g}_n$ by $\overline{g} := \sup_{n\geq 0}\overline{g}_n < \infty$.*

This theorem is a particular case of Theorem 5.21, which can be stated as follows. Let

$$\mathbb{P}_{z_n,n}^{(N,q)} = \mathrm{Law}\big(\mathcal{X}_{z_n,n}^2, \mathcal{X}_{z_n,n}^3, \ldots, \mathcal{X}_{z_n,n}^{q+1}\big), \tag{4.19}$$

be the distribution of the first q random non frozen particles $\mathcal{X}_{z,n}^{i+1}$ $i = 1, \ldots, q$. Using this notation, and under the regularity conditions stated in Theorem 4.14, there exists some finite constant $c > 0$ such that for any $n \geq 0, r > m \geq 1$ and $N > cn((r+q)\overline{g}_n)^2$, the Taylor expansion

$$\mathbb{P}_{z_n,n}^{(N,q)} = \eta_n^{\otimes q} + \sum_{1 \leq k \leq m} \frac{1}{N^k}d^{(k)}\mathbb{P}_{z_n,n}^{(q)} + \frac{1}{N^{m+1}}\partial^{(m+1)}\mathbb{P}_{z_n,n}^{(N,q)}, \tag{4.20}$$

exists for some signed and bounded measures $d^{(k)}\mathbb{P}_{z_n,n}^{(q)}$ with null mass $d^{(k)}\mathbb{P}_{z_n,n}^{(q)}(1) = 0$, whose values don't depend on the population size N, and such that

$$\big\|d^{(k)}\mathbb{P}_{z,n}^{(q)}\big\|_{\mathrm{tv}} \leq \big[cn\big((q+2k)\overline{g}_n\big)^2\big]^k \quad \text{and} \quad \big\|\partial^{(m+1)}\mathbb{P}_{z_n,n}^{(N,q)}\big\|_{\mathrm{tv}} \leq b(q)\big(cn\big((q+m)\overline{g}_n\big)^2\big)^{m+1},$$

with some finite constant $b(q) < \infty$ whose values only depend on the parameters q. Theorem 5.21 provides a more precise description of the derivative operators.

4.5. *Further developments and consequences*

We conclude this part of the article with a series of direct implications of the previous expansions around the fixed point Feynman–Kac measures. To illustrate our result we assume that the regularity condition (H) holds and the size N of the system is chosen such that $N > cn\overline{g}^2$ for some finite constant $c < \infty$, and the constant $\overline{g}$ defined in Theorem 4.14.

- These expansions can also be used to estimate of the behavior of the particle measures $m(\xi_{z_n,n})$ as $N \uparrow \infty$. For instance, the bias and the variance estimates are

$$\mathbb{E}\big(m(\mathcal{X}_{z_n,n})(f_n)\big) = \eta_n(f_n) + \frac{1}{N}\big(\big[f_n(z_n) - \eta_n(f)\big] + d^{(1)}\mathbb{P}_{z_n,n}^{(1)}(f)\big) + \frac{1}{N^2}r_{z_n,n}^{(N,1)}(f)$$

and

$$\mathrm{Var}\big(m(\mathcal{X}_{z_n,n})(f_n)\big) = \frac{1}{N}\big(\big[\eta_n\big(f_n^2\big) - \eta_n(f_n)^2\big] + d^{(1)}\mathbb{P}_{z_n,n}^{(2)}\big((f - \eta_n(f))^{\otimes 2}\big)\big) + \frac{1}{N^2}r_{z_n,n}^{(N,2)}(f)$$

with the two remainder terms such that

$$\sup_{i=1,2}\big|r_{z_n,n}^{(N,i)}(f)\big| \leq c\big(n\overline{g}^2\big)^2 \quad \text{for some finite } c < \infty.$$

The last estimate is related to the variance of the particle measures $m(\mathcal{X}_{z_n,n})$. Similarly, the variance of a function of the trajectory delivered by the PMCMC model is computed using the expansion of $\mathbb{E}(m(\mathcal{X}_{z_n,n})(f_n^2))$.

- Using the first order expansion (4.16), we can readily check that for any $\mu_n, \nu_n \in \mathcal{P}(S_n)$

$$N^2 \left\| (\mu_n - \nu_n)\mathbb{K}_n - \frac{1}{N}(\mu_n - \nu_n)d^{(1)}\mathbb{K}_n \right\|_{\mathrm{tv}} \leq c\left(n\overline{g}^2\right)^2,$$

for some finite constant $c < \infty$, and the first order integral operator $d^{(1)}\mathbb{K}_n$ defined in (4.15) and given by

$$d^{(1)}\mathbb{K}_n(f_n)(z_n) = \sum_{0 \leq p \leq n} \eta_p\left(\overline{Q}_{p,n}(1)\left[\overline{Q}_{p,n}(f_n)(z_p) - \overline{Q}_{p,n}(f_n)\right]\right).$$

This implies that

$$\left| \beta(\mathbb{K}_n) - \frac{1}{N}\beta\left(d^{(1)}\mathbb{K}_n\right) \right| \leq c\left(n\overline{g}^2/N\right)^2. \tag{4.21}$$

Using (4.15), we also have the crude estimate

$$\beta\left(d^{(1)}\mathbb{K}_n\right) \leq 2 \sum_{0 \leq k \leq n} \left\|\overline{Q}_{k,n}(1)\right\| \leq 2(n+1)\overline{g}.$$

To check that the right-hand side linear estimate with respect to the time horizon n is sharp, we choose unit potential functions $G_n = 1$ and a function $f_n(z_n) = \varphi(z'_0)$ that only depend on the initial state of the path $z_n = (z'_k)_{0 \leq k \leq n} \in S_n = \prod_{0 \leq k \leq n} S'_n$. Now,

$$d^{(1)}\mathbb{K}_n(f_n)(y_n) - d^{(1)}\mathbb{K}_n(f_n)(z_n) = (n+1)\left(\varphi(y'_0) - \varphi(z'_0)\right) \quad \Rightarrow \quad \beta\left(d^{(1)}\mathbb{K}_n\right) \geq (n+1).$$

The estimates (4.21) ensure that the Markov chain with transitions $\mathbb{K}_n$ converges exponentially fast to η_n with a rate that can be made arbitrarily large when the precision parameter and the size of the particle population $N \uparrow \infty$.

- Using the properties (4.17) we can readily prove the existence of the Taylor expansions of any mth iterate $\mathbb{K}_n^m = \mathbb{K}_n^{m-1}\mathbb{K}_n$ of the transition $\mathbb{K}_n$. For instance, for any $m \geq 1$, we have

$$\mathbb{K}_n^m(y_n, dz_n) = \eta_n(dz_n) + \frac{1}{N^m}\left[d^{(1)}\mathbb{K}_n\right]^m(y_n, dz_n) + \frac{1}{N^{m+1}}\partial^{(m+1)}\mathbb{K}^m(y_n, dz_n), \tag{4.22}$$

with the remainder integral operator $\partial^{(m+1)}\mathbb{K}_n^m$ such that

$$\partial^{(m+1)}\mathbb{K}_n^m(1)(y_n) = 0 \quad \text{and} \quad \beta\left(\partial^{(m+1)}\mathbb{K}_n^m\right) \leq m\left(cn\overline{g}^2\right)^{m+1}\left(1 + cn\overline{g}^2/N\right)^{m-1}.$$

This result shows that the distribution of the random state of the Markov chain with transition $\mathbb{K}_n$ after m iterations is equal to η_n, up to some remainder measure with total variation norm of order N^{-m}. In addition, arguing as above, we find that

$$N^{m+1}\left| \beta\left(\mathbb{K}_n^m\right) - \frac{1}{N^m}\beta\left(\left[d^{(1)}\mathbb{K}_n\right]^m\right) \right| \leq \beta\left(\partial^{(m+1)}\mathbb{K}_n^m\right).$$

- The decompositions (4.22) can be used to derive, without any additional work, the $\mathbb{L}_p$-norms between the distributions of the random states of the PG model and the invariant measures. For instance, for any $p \geq 1$,

$$\left\|\mathbb{K}_n^m(f_n) - \eta_n(f_n)\right\|_{\mathbb{L}_p(\eta_n)} = \frac{1}{N^m}\left\|\left[d^{(1)}\mathbb{K}_n\right]^m(f_n)\right\|_{\mathbb{L}_p(\eta_n)} + \frac{1}{N^{m+1}}\left\|\partial^{(m+1)}\mathbb{K}^m(f_n)\right\|_{\mathbb{L}_p(\eta_n)}.$$

- The proof of the Taylor expansions (4.19) is based on renormalization techniques and a differential calculus on the measures $\Upsilon_{z_n,n}^{(N,q)}$ on S_n^q defined for any $F_n \in \mathcal{B}(S_n^q)$ by

$$\Upsilon_{z_n,n}^{(N,q)}(F_n) := \mathbb{E}\left(m(\mathcal{X}_{z_n,n})^{\otimes q}(F_n) \prod_{0 \leq p < n} m(\mathcal{X}_{z_p,p})(\overline{G}_p)^q\right). \tag{4.23}$$

We will show that the $\Upsilon_{z_n,n}^{(N,q)}$ are differentiable of any order with $d^{(0)}\Upsilon_{z_n,n}^{(N,q)} = \eta_n^{\otimes q}$. Furthermore, formula (4.2) implies that

$$\int \eta_n(dz_n)\Upsilon_{z_n,n}^{(N,q-1)}(F_n) = \Upsilon_n^{(N,q)}(F_n \otimes 1), \tag{4.24}$$

for any $F_n \in \mathcal{B}(S_n^{q-1})$, with the measure $\Upsilon_n^{(N,q)}$ defined as $\Upsilon_{z_n,n}^{(N,q)}$ by replacing $(\mathcal{X}_{z_p,p})_{0 \le p \le n}$ by $(\mathcal{X}_n)_{0 \le p \le n}$. This formula can be used to compute Taylor type expansions for the occupation measures of the process $\mathcal{X}_n$, including the $(q+1)$th-moments of the unnormalized particle normalizing constants $\prod_{0 \le p < n} m(\mathcal{X}_p)(G_p)$.

In this connection, the transfer formula (4.24) also shows that the particle approximation $\prod_{0 \le p < n} m(\mathcal{X}_p)(G_p)$ of the normalizing constants associated with the particle model with a frozen trajectory is *biased even if the particle Markov chain model starts with the desired target measure*. For instance, for $q = 1$ and $F_n = 1$, equation (4.24) implies that

$$\mathbb{E}\left(\prod_{0 \le p < n} m(\mathcal{X}_p)(\overline{G}_p)\right) = 1 + \mathbb{E}\left(\left[\prod_{0 \le p < n} m(\mathcal{X}_p)(\overline{G}_p) - 1\right]^2\right) \neq 1.$$

By running a Markov chain with one of the transitions $\mathbb{K}_n$, we design an asymptotically unbiased estimate using the easily checked formula

$$\mathbb{E}\left(\left[\prod_{0 \le p < n} m(\mathcal{X}_p)(G_p)\right]^{-1}\right) = \left[\prod_{0 \le p < n} \eta_p(G_p)\right]^{-1}.$$

- We consider the Hidden Markov chain model discussed in Section 2.4.2. For any path $x_n \in S_n$, we let $P_{n,x_n}(\theta, d\overline{\theta})$ be a collection of Markov transitions in the parameter space s.t. $\lambda_{n,x_n} = \lambda_{n,x_n} P_{n,x_n}$. In this context, the Feynman–Kac model $\eta_{n,\theta}$ and the corresponding mean field particle Gibbs transitions $\mathbb{K}_{n,\theta}$ introduced in Definition 4.9 depend on the parameter θ. We set

$$\overline{\mathbb{K}}_n\big((\theta, dx_n), d(\overline{\theta}, d\overline{x}_n)\big) := P_{n,x_n}(\theta, d\overline{\theta})\mathbb{K}_{n,\overline{\theta}}(x_n, d\overline{x}_n).$$

The Taylor expansions stated in Theorem 4.14 and (4.20) and all their implications discussed above remain valid if $(\eta_n, \mathbb{K}_n)$ is replaced by $(\eta_{n,\theta}, \mathbb{K}_{n,\theta})$, for every fixed value of the parameter θ. In addition, assuming that the condition (H) stated in (2.6) is satisfied uniformly with respect to the parameter θ, the Taylor expansions stated in Theorem 4.14 and all their consequences discussed above remain valid if $(\eta_n, \mathbb{K}_n)$ is replaced by $(\overline{\eta}_n, \overline{\mathbb{K}}_n)$, with the posterior distribution $\overline{\eta}_n$ defined in (2.10).

5. Propagation of chaos expansions

This section focuses on the fine analysis of the size N dependence of PG samplers and related problems such as the asymptotic independence of $q \ll N$ subsets of the particle models investigated in the first part of the article – i.e., propagation of chaos properties.

5.1. *Combinatorial preliminaries*

We let $X = (X^i)_{2 \le i \le N}$ be a sequence of random variables on some state space S, and $z \in S$ a given fixed state. For any $q < N$ we set

$$m(X)^{\odot q} = \frac{1}{(N-1)_q} \sum_{a \in I_q^N} \delta_{(X^{a(1)},\dots,X^{a(q)})},$$

where I_q^N stands the set of all $(N-1)_q = \frac{(N-1)!}{((N-1)-q)!}$ multi-indices $a = (a(1), \dots, a(q)) \in \{2, \dots, N\}^q$ with different values, or equivalently one to one mappings from $[q] := \{1, \dots, q\}$ into $\{2, \dots, N\} = [N] - \{1\}$. The link between

these measures and tensor product measures is expressed in terms of the Markov transitions $\mathbb{A}_a^{(q)}$ indexed by the set of mappings a from $[q]$ into itself and defined for any $x = (x^1, \ldots, x^q) \in S^q$ by

$$\mathbb{A}_a^{(q)}(F)(x) = F(x^a) \quad \text{with } x^a := (x^{a(1)}, \ldots, x^{a(q)}),$$

for any function F on $\mathcal{B}(S^q)$, and any $(x^1, \ldots, x^q) \in S^q$. The connection between these measures is described in the following technical lemma taken from [21].

We emphasize that the tensor product measures discussed above are symmetry-invariant by construction. To proceed, we assume, without loss of generality, that these measures act on symmetric functions F; i.e., $F = \frac{1}{q!} \sum_{\sigma \in \mathcal{G}_q} \mathbb{A}_\sigma^{(q)}(F)$, where $\mathcal{G}_q$ stands for the symmetric group of all permutations of $[q]$.

Lemma 5.1. *For any $q < N$,*

$$m(X)^{\otimes q} = m(X)^{\odot q} \mathbb{A}^{(N,q)} \quad \text{with } \mathbb{A}^{(N,q)} = \frac{1}{(N-1)^q} \sum_{a \in [q]^{[q]}} \frac{(N-1)_{|a|}}{(q)_{|a|}} \mathbb{A}_a^{(q)},$$

where $|a|$ is the cardinality of the set $a([q])$, and $(m)_p = m!/(m-p)!$ is the number of one to one mappings from $[p]$ into $[m]$.

Definition 5.2. *For any $z \in S$, we define the random measures*

$$m_z(X) := \frac{1}{N} \delta_z + \left(1 - \frac{1}{N}\right) m(X), \qquad m_z^{(1)}(X) := \delta_z \quad and \quad m_z^{(0)}(X) := m(X).$$

For any $b \in \{0, 1\}^{[q]}$, let $\mathbb{B}_{z,b}^{(q)}$ denote the Markov transitions defined for any $x = (x^1, \ldots, x^q) \in S^q$ by

$$\mathbb{B}_{z,b}^{(q)}(F)(x) = F(x_z^b) \quad \text{with } x_z^b := (b(1)z + (1 - b(1))x^1, \ldots, b(q)z + (1 - b(q))x^q).$$

We observe that

$$m_z(X)^{\otimes q} = \sum_{b \in \{0,1\}^{[q]}} \frac{1}{N^{|b|_1}} \left(1 - \frac{1}{N}\right)^{q - |b|_1} m_z^{(b)}(X),$$

with $|b|_1 = \sum_{1 \leq p \leq q} b(p)$ and

$$m_z^{(b)}(X) = m_z^{(b(1))}(X) \otimes \cdots \otimes m_z^{(b(q))}(X).$$

Lemma 5.3. *For any $q < N$, and $b \in \{0, 1\}^{[q]}$ we have $m_z^{(b)}(X) = m^{\otimes q}(X) \mathbb{B}_{z,b}^{(N,q)}$ and*

$$m_z(X)^{\otimes q} = m^{\otimes q}(X) \mathbb{B}_z^{(N,q)} \quad \text{with } \mathbb{B}_z^{(N,q)} = \sum_{b \in \{0,1\}^{[q]}} \frac{1}{N^{|b|_1}} \left(1 - \frac{1}{N}\right)^{q - |b|_1} \mathbb{B}_{z,b}^{(q)}$$

as well as

$$m_z(X)^{\otimes q} = m(X)^{\odot q} \mathbb{C}_z^{(N,q)} \quad \text{with } \mathbb{C}_z^{(N,q)} := \mathbb{A}^{(N,q)} \mathbb{B}_z^{(N,q)}.$$

Definition 5.4. *We let p_1 and p_2 be two integers s.t. $0 \leq p_1 \leq q - 1$ and $0 \leq p_2 \leq q$.*

- *We consider the collection of sets*

$$\mathcal{I}_q := \{0, \ldots, q - 1\} \times \{0, \ldots, q\}, \qquad [r]_{q-p_1}^{[q]} := \{a \in [r]^{[q]} : |a| = q - p_1\},$$

$$\{0, 1\}_{1, p_2}^{[q]} := \{b \in \{0, 1\}^{[q]} : |b|_1 = p_2\} \quad and \quad I_q(p_1, p_2) = [q]_{q-p_1}^{[q]} \times \{0, 1\}_{1, p_2}^{[q]}.$$

- We let $\mathcal{A}_{p_1}^{(q)}$ and $\mathcal{B}_{p_2}^{(q)}$ be the uniform distributions on $[q]_{q-p_1}^{[q]}$ and $\{0,1\}_{1,p_2}^{[q]}$ respectively. We also denote by $\mathcal{C}_{p_1,p_2}^{(q)} = \mathcal{A}_{p_1}^{(q)} \otimes \mathcal{B}_{p_2}^{(q)}$ the uniform measure on $I_q(p_1,p_2)$.
- For any $c = (a,b) \in I_q(p_1,p_2)$, we let $\mathbb{C}_{z,(a,b)}^{(q)}$ be the coalescent operator defined for any $x = (x^1, \ldots, x^q) \in S^q$ by

$$\mathbb{C}_{z,(a,b)}^{(q)}(F)(x) := F\big(x_z^{(a,b)}\big),$$

with

$$x_z^{(a,b)} = \big(b(1)z + \big(1 - b(1)\big)x^{a(1)}, \ldots, b(q)z + \big(1 - b(q)\big)x^{a(q)}\big),$$

so that $\mathbb{C}_{z,(a,b)}^{(q)} = \mathbb{A}_a^{(q)}\mathbb{B}_{z,b}^{(q)}$.

Remark 5.5. When maps in $[q]^{[q]}$ are represented graphically, the parameter p_1 in $[q]_{q-p_1}^{[q]}$ represents the number of coalescences of the change of index mapping a. The p_2 is the number of $b(i)$ such that $b(i) = 0$ or $x_z^{(a,b),i} = z$; it will be referred to as the number of z-infections of the mapping b.

We recall that for the Stirling numbers of the second kind, $S(q,p)$ is the number of partitions of $[q]$ into p sets, so that

$$\#\big([r]_p^{[q]}\big) = S(q,p)(r)_p \quad \text{and} \quad r^q = \sum_{1 \leq p \leq q} S(q,p)(r)_p,$$

for any $p \leq q \leq r$. We also recall that the Stirling numbers of the first kind $s(q,p)$ provide the coefficients of the polynomial expansion of $(r)_q$

$$(r)_q = \sum_{1 \leq p \leq q} s(q,p)r^p. \tag{5.1}$$

We also use the convention that $(r)_q = 0$ and $(r)_0 = 1 = (-r)_0$ for any $q > r \geq 0$, as well as $s(q,0) = s(0,-q) = S(0,-q) = S(q,0) = 0$ except $s(0,0) = S(0,0) = 1$, for $q = 0$.

These formulae can be found in any textbook on combinatorial analysis, e.g., [13,14].

Definition 5.6. We define the sequence of probabilities $\mathcal{P}^{(N,q)} = \mathcal{P}^{[N,q,1]} \otimes \mathcal{P}^{[N,q,2]}$ on the set $\mathcal{I}_q$ as

$$\mathcal{P}^{(N,q)}(p_1,p_2) := \underbrace{\frac{1}{(N-1)^q}S(q,q-p_1)(N-1)_{q-p_1}}_{\mathcal{P}^{[N,q,1]}(p_1)} \times \underbrace{\binom{q}{p_2}\left(1 - \frac{1}{N}\right)^{q-p_2}\frac{1}{N^{p_2}}}_{\mathcal{P}^{[N,q,2]}(p_2)}. \tag{5.2}$$

We note that $\mathcal{P}^{[N,q,1]}(p_1) = \#([N-1]_{q-p_1}^{[q]})/\#[N-1]^{[q]}$ is a statistic for the number of coalescences, whereas $\mathcal{P}^{[N,q,2]}(p_2)$ is the proportion of infested mappings with p_2 infections. The following lemma is obtained by construction.

Lemma 5.7. The following formula holds for any $q < N$

$$\mathbb{C}_z^{(N,q)} = \sum_{p \in \mathcal{I}_{0,q}} \mathcal{P}^{(N,q)}(p)\widehat{\mathbb{C}}_{z,p}^{(q)} \quad \text{with} \quad \widehat{\mathbb{C}}_{z,p}^{(q)} = \sum_{c \in I_q(p)} \mathcal{C}_p^{(q)}(c)\mathbb{C}_{z,c}^{(q)}.$$

We end this section with a Taylor series expansion of the measure $\mathcal{P}^{(N,q)}$ introduced above.

Proposition 5.8. *For any $q < N$, the mapping $N \mapsto \mathcal{P}^{(N,q)}$ is differentiable of any order $m \geq 0$. The mth-order derivative is supported by*

$$\mathcal{T}_{q,n}^{(m)} := \left\{ (p_1, p_2) \in \mathcal{I}_q : 0 \leq p_1 + p_2 \leq m \right\}.$$

Equation (5.2) shows that the fraction in the variable N, $\mathcal{P}^{(N,q)}(p_1, p_2)$, can be expanded as a formal power series in $\frac{1}{N}$ (or, more precisely, as an analytic function in the neighborhood of 0) with leading term in $\frac{1}{N^{p_1+p_2}}$. The proposition follows.

Expanding the formula for $\mathcal{P}^{(N,q)}(p_1, p_2)$ using (5.1) and the Taylor expansion

$$\frac{1}{(1-x)^n} = \sum_{0 \leq k} (n+k-1)_k \frac{x^k}{k!} = \sum_{0 \leq k} \binom{n+k-1}{k} x^k,$$

with $(n-1)_0 := 1$, we get an explicit formula for the derivatives.

Proposition 5.9. *The mth-order derivative is given by the signed measure (with total null mass) supported on the set $\mathcal{T}_{q,n}^{(m)}$:*

$$d^{(m)} \mathcal{P}^{(q)} := \sum_{(p_1,p_2) \in \mathcal{T}_{q,n}^{(m)}} \tau_{q,p_1,p_2}^{(m)} \delta_{(p_1,p_2)}, \tag{5.3}$$

with

$$\tau_{q,p_1,p_2}^{(m)} = \sum_{\mathbf{k} \in \mathcal{K}_q^{(m)}(p_1,p_2)} \alpha_{q,p_1,p_2}(\mathbf{k}),$$

$$\mathcal{K}_q^{(m)}(p_1,p_2) := \left\{ (k_1, k_2, k_3) \in [0, q-p_1[\times [0, q-p_2] \times \mathbb{N} : \sum_{1 \leq i \leq 2} p_i + \sum_{1 \leq i \leq 3} k_i = m \right\},$$

$$\alpha_{q,p_1,p_2}(k_1, k_2, k_3) = S(q, q-p_1) \binom{q}{p_2} \tag{5.4}$$

$$\times s(q-p_1, q-p_1-k_1)(-1)^{k_2} \binom{q-p_2}{k_2} \binom{(p_1+k_1)+k_3-1}{k_3}.$$

Remark 5.10. *We observe that $\tau_{q,p_1,p_2}^{(0)} = 1_{(0,0)}(p_1, p_2)$. As we show below, this identity encodes the propagation of chaos properties (i.e., asymptotic independence) of PG samplers. We also mention that $\alpha_{q,p_1,p_2}(\mathbf{k}) = 0 = \tau_{q,p_1,p_2}^{(m)}$ as soon as $p_1 > q$ or $p_2 > q$.*

Remark 5.11. *The mth-order signed measure $d^{(m)} \mathcal{P}^{(q)}$ and the mapping $(p_1, p_2) \mapsto \tau_{q,p_1,p_2}^{(m)}$ in formula (5.3) only give mass to integers $(p_1, p_2) \in ([1, q] \times [0, q])$ such that $0 \leq p_1 + p_2 \leq m$. The first coordinate $0 \leq p_1 < q$ can be interpreted as the number of coalescent states, while p_2 can be interpreted as the number of z-infected states.*

By construction, the mapping $(p_1, p_2) \mapsto \tau_{q,p_1,p_2}^{(m)}$ can also be seen as a measure with null total mass supported on the set $0 \leq p_1 + p_2 \leq m$. For instance, for $m = 1, 2$, recalling that $s(q, q-1) = -q(q-1)/2 = -S(q, q-1)$, $s(q, q-2) = \frac{q!}{3!(q-3)!} \frac{3q-1}{4}$, and $S(q, q-2) = \frac{q!}{3!(q-3)!}(3q-5)/4$, we have

$$\tau_{q,2,0}^{(2)} = \frac{q!}{3!(q-3)!} \frac{3q-5}{4}, \qquad \tau_{q,0,2}^{(2)} = \frac{q(q-1)}{2},$$

$$\tau_{q,0,0}^{(2)} = \frac{q^2(q-1)}{2} + \frac{q!}{3!(q-3)!} \frac{3q-1}{4}, \qquad \tau_{q,1,0}^{(2)} = -\left(\frac{q(q-1)}{2} \right)^2, \tag{5.5}$$

$$\tau_{q,0,1}^{(2)} = -\frac{q^2(q-1)}{2} - q(q-1), \qquad \tau_{q,1,1}^{(2)} = q\frac{q(q-1)}{2},$$

$$\tau_{q,1,0}^{(1)} = \frac{q(q-1)}{2}, \qquad \tau_{q,0,1}^{(1)} = q, \qquad \tau_{q,0,0}^{(1)} = -\big(\tau_{q,1,0}^{(1)} + \tau_{q,0,1}^{(1)}\big).$$

Definition 5.12. *We denote by* $\mathbf{p}_n := (p_0, \ldots, p_n)$ *a given multi-index in* $\mathcal{I}_{n,q} := (\mathcal{I}_q)^{n+1}$, *with* $p_k = (p_k^1, p_k^2) \in \mathcal{I}_q$ *for any* $0 \le k \le n$. *We also denote by* $\mathbf{c}_n = (c_0, \ldots, c_n)$ *a sequence of mappings in the set*

$$\mathcal{J}_{q,n} = \bigcup_{\mathbf{p}_n \in \mathcal{I}_{n,q}} \mathbf{I}_q(\mathbf{p}_n) \quad \text{with } \mathbf{I}_q(\mathbf{p}_n) := \prod_{0 \le k \le n} I_q(p_k).$$

For any $\mathbf{m}_n = (m_0, \ldots, m_n) \in \mathbb{N}^{n+1}$, *we set* $|\mathbf{m}_n| = \sum_{0 \le k \le n} m_k$, *and we use the multi-index notation*

$$\boldsymbol{\tau}_{q,\mathbf{p}_n}^{(\mathbf{m}_n)} = \prod_{0 \le k \le n} \tau_{q,p_k^1,p_k^2}^{(m_k)}, \qquad \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(m)} := \sum_{|\mathbf{m}_n|=m} \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(\mathbf{m}_n)}, \qquad \mathcal{T}_{q,n}^{(m)} := \bigsqcup_{|\mathbf{m}_n|=m} \prod_{0 \le k \le n} \mathcal{T}_{q,n}^{(m_k)}$$

and

$$\mathcal{C}_{\mathbf{p}_n}^{(q)}(\mathbf{c}_n) := \prod_{0 \le k \le n} \mathcal{C}_{p_k}^{(q)}(c_k), \qquad \mathcal{P}_n^{(N,q)}(\mathbf{p}_n) := \prod_{0 \le k \le n} \mathcal{P}^{(N,q)}(p_k).$$

In this notation, and recalling that $p_n^1 + p_n^2 > m_n \Rightarrow \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(\mathbf{m}_n)} = 0$, we can prove the following extension of Lemma 5.9:

Proposition 5.13. *For any* $q < N$ *and* $n \ge 0$, *the mapping* $N \mapsto \mathcal{P}_n^{(N,q)}$ *is differentiable of any order. Furthermore, the* m*th-order derivative is the signed measure with null mass*

$$d^{(m)}\mathcal{P}_n^{(q)} = \sum_{\mathbf{p}_n \in \mathcal{T}_{q,n}^{(m)}} \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(m)} \delta_{\mathbf{p}_n}.$$

Finally,

$$\sum_{\mathbf{p}_n \in \mathcal{T}_{q,n}^{(m)}} \big|\boldsymbol{\tau}_{q,\mathbf{p}_n}^{(m)}\big| \le \frac{(m+n)!}{m!\,n!}(cq)^{2m}, \tag{5.6}$$

for some finite constant $c < \infty$.

Proof. By Theorem 2 in [4], for any $p \le q$ we have the rather crude estimates

$$S(q, q-p) \le c\frac{q^{2p}}{2^p\, p!} \le cq^{2p} \quad \text{and} \quad \big|s(q, q-p)\big| \le c\binom{q}{q-p}\left(\frac{q-p}{2}\right)^p \le cq^{2p},$$

for some finite constant $c < \infty$. We also note that

$$\binom{q}{p_2} \le q^{p_2}\binom{q-p_2}{k_2} \le q^{k_2} \quad \text{and} \quad \binom{(p_1+k_1)+k_3-1}{k_3} \le (2eq)^{k_3}. \tag{5.7}$$

To obtain the right-hand side estimate, we use Stirling's approximation to check that

$$\binom{(p_1+k_1)+k_3-1}{k_3} \le (q+k_3)^{k_3}/k_3! \le \big(e(q+k_3)/k_3\big)^{k_3} \le \big(e(q+1)\big)^{k_3} \le (2eq)^{k_3}.$$

Combining (5.7) and (5.5) with the inequality

$$\big|\alpha_{q,p_1,p_2}(\mathbf{k})\big| \le (cq)^{2p_1+p_2+2k_1+k_2+k_3} \le (cq)^{2m},$$

for any $\mathbf{k} \in \mathcal{K}_q^{(m)}(p_1, p_2)$ and some finite constant $c < \infty$, we conclude that

$$\left| \tau_{q,p_1,p_2}^{(m)} \right| \leq \left(m - (p_1 + p_2) \right)^2 \times (cq)^{2m} \leq m^2 (cq)^{2m} \leq \left(c'q \right)^{2m},$$

for some finite constant $c' < \infty$. Similarly,

$$\sum_{0 \leq p_1 + p_2 \leq m} \left| \tau_{q,p_1,p_2}^{(m)} \right| \leq m^2 \left(c'q \right)^{2m} \leq (cq)^{2m},$$

for some finite constant $c < \infty$. This yields the rather crude estimates

$$\forall m_0 + \cdots + m_n = m \qquad \sum_{0 \leq p_0^1 + p_0^2 \leq m} \cdots \sum_{0 \leq p_n^1 + p_n^2 \leq m} \left| \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(m)} \right| \leq (cq)^{2m}.$$

The estimate (5.6) holds because the cardinality of the set $\{(m_0, \ldots, m_n) : m_0 + \cdots + m_n = m\}$ coincides with the number $\frac{(m+n)!}{m!n!}$ of finite multisets of size m whose elements are drawn from a set of $(n+1)$ elements. $\square$

Definition 5.14. *To proceed further, let* $\mathbf{c} = (c_0, \ldots, c_n)$, $c_i = (a_i, b_i)$ *be a sequence of mappings in the set* $\mathcal{J}_{q,n}$*. We say that*

- *the pth trajectory,* $1 \leq p \leq q$ *of* $\mathbf{c}$ *is free if* $\forall i \leq n, \forall m \neq p$,

  $$a_i \circ \cdots \circ a_n(p) \neq a_i \circ \cdots \circ a_n(m) \quad \text{and} \quad b_i \left(a_{i+1} \circ \cdots \circ a_n(p) \right) \neq 1,$$

- *the pth trajectory is coalescent if* $\exists i \leq n, \exists m \neq p, a_i \circ \cdots \circ a_n(p) = a_i \circ \cdots \circ a_n(m)$,
- *the pth trajectory is infected if* $\exists i \leq n, b_i(a_{i+1} \circ \cdots \circ a_n(p)) = 1$.

5.2. *Unnormalized tensor product measures*

We now apply these combinatorial results to PG samplers. Our first result is concerned with tensor product measures. Given a frozen trajectory $z := (z_n)_{n \geq 0} \in \prod_{n \geq 0} S_n$, we denote by $\mathcal{X}_{z,n}$ the dual many body model associated with the Feynman–Kac particle model $\mathcal{X}_n$ and the frozen path $X_n = z_n$.

We also set

$$\eta_{z,n}^N := m(\mathcal{X}_{z,n}) = m_{z_n}\left(\mathcal{X}_{z,n}^- \right), \qquad \gamma_{z,n}^N(1) := \prod_{0 \leq p < n} \eta_{z,p}^N(G_p), \qquad \gamma_{z,n}^N := \gamma_{z,n}^N(1) \cdot \eta_{z,n}^N,$$

and finally, for any function F on S_n^q,

$$\Upsilon_{z,n}^{(N,q)}(F) := \mathbb{E}\left(\left(\gamma_{z,n}^N \right)^{\otimes q}(F) \right) / \gamma_n(1)^q.$$

Definition 5.15. *We consider the tensor product measures*

$$\Delta_{z,\mathbf{p}_n}^{(q)} = \left(\eta_0^{\otimes q} \widehat{\mathbb{C}}_{z_0,p_0}^{(q)} \right) \left(\overline{Q}_1^{\otimes q} \widehat{\mathbb{C}}_{z_1,p_1}^{(q)} \right) \cdots \left(\overline{Q}_n^{\otimes q} \widehat{\mathbb{C}}_{z_n,p_n}^{(q)} \right) = \sum_{\mathbf{c}_n \in I_q(\mathbf{p}_n)} \mathcal{C}_{\mathbf{p}_n}^{(q)}(\mathbf{c}_n) \Delta_{z,\mathbf{c}_n}^{(q)} \tag{5.8}$$

with the conditional expectation operators

$$\Delta_{z,\mathbf{c}_n}^{(q)} := \left(\eta_0^{\otimes q} \mathbb{C}_{z_0,c_0}^{(q)} \right) \left(\overline{Q}_1^{\otimes q} \mathbb{C}_{z_1,c_1}^{(q)} \right) \cdots \left(\overline{Q}_n^{\otimes q} \mathbb{C}_{z_n,c_n}^{(q)} \right).$$

Theorem 5.16. *For any* $q < N, n \geq 0$,

$$\Upsilon_{z,n}^{(N,q)} = \sum_{\mathbf{p}_n \in \mathcal{I}_{n,q}} \sum_{\mathbf{c}_n \in I_q(\mathbf{p}_n)} \left[\mathcal{P}_n^{(N,q)}(\mathbf{p}_n) \mathcal{C}_{\mathbf{p}_n}^{(q)}(\mathbf{c}_n) \right] \Delta_{z,\mathbf{c}_n}^{(q)}.$$

Proof. Define $\eta_{z,n}^N := m_{z_n}(\mathcal{X}_{z,n}^-)$, by Lemma 5.1 we have

$$m_{z_n}(\mathcal{X}_{z,n}^-)^{\otimes q} = m(\mathcal{X}_{z,n}^-)^{\odot q}\mathbb{C}_{z_n}^{(N,q)}.$$

Furthermore, for any function F on S_n^q,

$$\mathbb{E}\big(m(\mathcal{X}_{z,n+1}^-)^{\odot q}(F)|\mathcal{F}_n\big) = \big(\eta_{z,n}^N\big)^{\otimes q}\big(Q_{n+1}^{\otimes q}(F)\big)/\eta_{z,n}^N(G_n)^q,$$

implying that

$$\mathbb{E}\big(\big(\gamma_{z,n+1}^N\big)^{\otimes q}(F)|\mathcal{F}_n\big) = \gamma_{z,n}^N(1)^q \times \big(\eta_{z,n}^N\big)^{\otimes q}\big(Q_{n+1}^{\otimes q}\mathbb{C}_{z_{n+1}}^{(N,q)}(F)\big)$$
$$= \big(\gamma_{z,n}^N\big)^{\otimes q}\big(Q_{n+1}^{\otimes q}\mathbb{C}_{z_{n+1}}^{(N,q)}(F)\big),$$

from which we conclude that

$$\Upsilon_{z,n}^{(N,q)}(F) = \big(\eta_0^{\otimes q}\mathbb{C}_{z_0}^{(N,q)}\big)\big(\overline{Q}_1^{\otimes q}\mathbb{C}_{z_1}^{(N,q)}\big)\dots\big(\overline{Q}_n^{\otimes q}\mathbb{C}_{z_n}^{(N,q)}\big)(F).$$

The result follows by expanding the $\mathbb{C}_{z_i}^{(N,q)}$ in terms of the $\mathbb{C}_{z_i,c_i}^{(q)}$. $\square$

The next corollary is a direct consequence of the proof of Theorem 5.16. It provides a more probabilistic description of the measure $\Upsilon_n^{(N,q)}$ in terms of expectation operators.

Corollary 5.17. *For any $q < N$ and $n \geq 0$, $\Upsilon_{z,n}^{(N,q)}$ is differentiable of any order. In addition, for any $n \geq 0$, its derivatives are given by the recursion*

$$d^{(m)}\Upsilon_{z,n}^{(q)}(F) = \sum_{r_1+r_2=m}\sum_{p\in\mathcal{I}_q} d^{(r_1)}\mathcal{P}^{(q)}(p)d^{(r_2)}\Upsilon_{z,n-1}^{(q)}\big(\overline{Q}_n^{\otimes q}\widehat{\mathbb{C}}_{z_n,p}^{(q)}(F)\big),$$

using the convention that $\Upsilon_{z,-1}^{(q)}\overline{Q}_0^{\otimes q} = \eta_0^{\otimes q}$ and $d^{(r_2)}\Upsilon_{z,-1}^{(q)}\overline{Q}_0^{\otimes q} = 0$ for $r_2 > 0$. In particular, we obtain the expansions

$$d^{(m)}\Upsilon_{z,n}^{(q)} = \sum_{\mathbf{p}_n\in\mathcal{T}_{q,n}^{(m)}} \boldsymbol{\tau}_{q,\mathbf{p}_n}^{(m)} \times \Delta_{z,\mathbf{p}_n}^{(q)}. \tag{5.9}$$

To proceed, we study the action of the operators $\Delta_{z,\mathbf{c}_n}^{(q)}$. We already know that they contribute to $d^{(m)}\Upsilon_{z,n}^{(q)}$ only if the total number of coalescences and infections of $\mathbf{c}_n$, written $\text{Tot}(\mathbf{c}_n)$, is less than m.

Lemma 5.18. *For any $\mathbf{p}_n \in \mathcal{T}_{q,n}^{(m)}$, $m \geq 1$, and $n' \geq n$,*

$$\big\|\Delta_{z,\mathbf{p}_n}^{(q)}\overline{Q}_{n,n'}^{\otimes q}\big\|_{\text{tv}} \leq \overline{g}_{n'}^{2m} \quad and \quad \big\|d^{(m)}\Upsilon_{z,n}^{(q)}\overline{Q}_{n,n'}^{\otimes q}\big\|_{\text{tv}} \leq \big(cn(q\overline{g}_{n'})^2\big)^m, \tag{5.10}$$

for some finite constant $c < \infty$. In addition, let f be a η_n-centered function on S_n (i.e., $\eta_n(f) = 0$). Then, for any sequence of mappings $\mathbf{c}_n$,

$$\text{Tot}(\mathbf{c}_n) < \frac{q}{2} \quad \Rightarrow \quad \Delta_{z,\mathbf{c}_n}^{(q)}(f^{\otimes q}) = 0.$$

In particular, $d^{(m)}\Upsilon_{z,n}^{(q)}(f^{\otimes q}) = 0$ whenever $m < \frac{q}{2}$.

Proof. The first assertion follows because for any $\mathbf{c}_n$ with p_k^1-coalescences and p_k^2 infections at levels $0 \leq k \leq n$, we have the rather crude estimates

$$\big\|\Delta_{z,\mathbf{c}_n}^{(q)}\big(\overline{Q}_{n,n'}^{\otimes q}(F)\big)\big\| \leq \prod_{0\leq k\leq n} \overline{g}_{n'}^{2p_k^1+p_k^2} \leq \overline{g}_{n'}^{2|\mathbf{p}_n|},$$

with

$$|\mathbf{p}_n| := \sum_{0 \le k \le n} \left(p_k^1 + p_k^2 \right) \le \sum_{0 \le k \le n} m_k = m$$

for any function F on $S_{n'}^q$ s.t. $\|F\| \le 1$. The proof of (5.10) is completed as a direct consequence of (5.6), (5.8) and

$$\frac{(m+n)!}{m!n!n^m} \le \frac{(n+m)^m}{n^m m!} \le \frac{m^m}{m!} \left(\frac{1}{m} + \frac{1}{n} \right) \le e^{2m}.$$

To prove the second assertion, suppose that $\mathrm{Tot}(\mathbf{c}_n) < \frac{q}{2}$. It follows immediately that one trajectory is free in the sense of Definition 5.14. Because of the symmetry of the problem (which, as usual, is invariant to a permutation of the particles), we may assume without restriction that the particles of this free trajectory all have the same index q ($a_i(q) = q \ \forall i \le n$). Denote by $\hat{\mathbf{c}}_n$ the sequence of mappings obtained by restricting each a_i to a map from $[q-1]$ to itself (this process is well-defined because of the freeness assumption) and by similarly restricting b_i to $[q-1]$. It then follows from the definition of $\Delta_{z,\mathbf{c}_n}^{(q)}(f^{\otimes n})$ that

$$\Delta_{z,\mathbf{c}_n}^{(q)}\left(f^{\otimes q} \right) = \Delta_{z,\hat{\mathbf{c}}_n}^{(q-1)}\left(f^{\otimes q-1} \right) \cdot \eta_n(f) = 0. \qquad \square$$

Corollary 5.19. *For an arbitrary $q \le N$,*

$$\mathbb{E}\left[\left(\gamma_{z,n}^N\left(G_n - \eta_n(G_n) \right) \right)^q \right] = \mathbb{E}\left[\left(\gamma_{z,n}^N \right)^{\otimes q}\left(\left(G_n - \eta_n(G_n) \right)^{\otimes q} \right) \right] = O\left(N^{-q/2} \right).$$

Corollary 5.20. *For an arbitrary $q \le N$,*

$$\mathbb{E}\left[\left(\gamma_{z,n}^N(G_n) - \gamma_n(G_n) \right)^q \right] = O\left(N^{-q/2} \right).$$

Indeed,

$$\begin{aligned} \gamma_{z,n}^N(G_n) - \gamma_n(G_n) &= \prod_{0 \le p \le n} \eta_{z,p}^N(G_p) - \prod_{0 \le p \le n} \eta_p(G_p) \\ &= \gamma_{z,n}^N\left(G_n - \eta_n(G_n) \right) + \left[\gamma_{z,n-1}^N(G_{n-1}) - \gamma_{n-1}(G_{n-1}) \right] \eta_n(G_n) \\ &= \sum_{i=0}^n \left[\gamma_{z,i}^N(G_i - \eta_i(G_i)) \right] \prod_{j=i+1}^n \eta_i(G_i). \end{aligned}$$

The proof follows from the previous corollary and the Minkowski inequality.

5.3. *Normalized tensor product measures*

This section shows that the distribution $\mathbb{P}_{z,n+1}^{(N,q)}$ of the first q random non frozen particles (see Definition 4.19) has derivatives of all orders.

We recall the instrumental identity: for any $u \ne 1, q \ge 0$ and $m \ge 1$

$$\frac{1}{(1-u)^{q+1}} = \sum_{0 \le k \le m} (q+k)_k \frac{u^k}{k!} + u^m \sum_{1 \le k \le q+1} \binom{(q+1)+m}{k+m} \left(\frac{u}{1-u} \right)^k. \qquad (5.11)$$

See [21, Lemma 4.11 on page 820] for a detailed proof. Using the identity

$$\binom{n+1}{k} = \sum_{k \le l \le n} \binom{n}{l} \quad \text{following, e.g., from } 1 - (1-x)^{n+1} = x \sum_{0 \le k \le n} (1-x)^k,$$

we obtain

$$\frac{1}{x^q} = \frac{(q+r)!}{(q-1)!} \sum_{0 \le l \le r} \frac{1}{(q+l)} \frac{(-1)^l}{l!(r-l)!} x^l + \sum_{1 \le k \le q} \binom{q+r}{k+r} \frac{(1-x)^{r+k}}{x^k}. \tag{5.12}$$

Theorem 5.21. *There exists some finite constant $c > 0$ such that for any $n \ge 0$, $r > m \ge 1$ and $N > cn((r+q)\overline{g}_{n+1})^2$, such that*

$$\left\| \mathbb{P}_{z,n+1}^{(N,q)} - \eta_{n+1}^{\otimes q} - \sum_{1 \le k \le m} \frac{1}{N^k} d^{(k)} \mathbb{P}_{z,n+1}^{(q)} \right\|_{\mathrm{tv}} \le b(q) \left(\frac{cn((q+r)\overline{g}_{n+1})^2}{N} \right)^{m+1},$$

with some finite constant $b(q) < \infty$ whose values only depend on the parameters q, and the kth order derivatives given for any function F on S_n^q by

$$d^{(k)} \mathbb{P}_{z,n+1}^{(q)}(F) = \frac{(q+2k)!}{(q-1)!} \sum_{0 \le l \le 2k} \frac{(-1)^l}{(q+l)} \frac{1}{l!(2k-l)!} d^{(k)} \Upsilon_{z,n}^{(l+q)} \left[\overline{Q}_{n,n+1}^{\otimes(l+q)} (1^{\otimes l} \otimes F) \right]. \tag{5.13}$$

In addition,

$$\left\| d^{(k)} \mathbb{P}_{z,n+1}^{(q)} \right\|_{\mathrm{tv}} \le \left[cn \left((q+2k)\overline{g}_{n+1} \right)^2 \right]^k, \tag{5.14}$$

for some finite constant $c < \infty$.

Proof. The proof of (5.14) is a direct consequence of (5.10). Let $\overline{\gamma}_{z,n}^N(f) = \gamma_{z,n}^N(f)/\gamma_n(1)$. In this notation,

$$\gamma_{z,n}^N(\overline{G}_n) - \gamma_n(\overline{G}_n) = \gamma_n(1)\left(\overline{\gamma}_{z,n}^N(\overline{G}_n) - 1 \right).$$

If q is an even integer, then by Corollary 5.20,

$$\mathbb{E}\left(\left(\overline{\gamma}_{z,n}^N(\overline{G}_n) - 1 \right)^q \right) = \sum_{k \ge q/2} \frac{1}{N^k} \sum_{0 \le p \le q} \binom{q}{p} (-1)^{q-p} d^{(k)} \Upsilon_{z,n}^{(p)}(\overline{G}_n^{\otimes p})$$

$$\le 2^q \sum_{k \ge q/2} \left(\frac{cn(q\overline{g}_{n+1})^2}{N} \right)^k,$$

for some finite constant $c < \infty$ and for any $N \ge cn(q\overline{g}_{n+1}^2)^2$. The right-hand side estimate is readily checked by recalling that $\overline{Q}_{n,n+1}^{\otimes p}(1) = \overline{G}_n^{\otimes p}$ and applying (5.10). Thus, there exists some finite universal constant $c < \infty$ such that

$$\mathbb{E}\left(\left(\overline{\gamma}_{z,n}^N(\overline{G}_n) - 1 \right)^q \right)^{1/q} \le \left(cn(q\overline{g}_{n+1})^2 / N \right)^{1/2}$$

and

$$\mathbb{E}\left(\left(\overline{\gamma}_{z,n}^N(\overline{G}_n) \right)^q \right)^{1/q} \le 1 + \left(cn(q\overline{g}_{n+1})^2 / N \right)^{1/2} \le 2,$$

as soon as $N \ge cn(q\overline{g}_{n+1})^2$.

It follows from the proof of Theorem 5.16 that

$$\mathbb{E}\left(m\left(\mathcal{X}_{z,n+1}^- \right)^{\odot q}(F) \right) = \mathbb{E}\left[\overline{\gamma}_{z,n}^N(\overline{G}_n)^{-q} \times \left(\overline{\gamma}_{z,n}^N \right)^{\otimes q} \left(\overline{Q}_{n,n+1}^{\otimes q}(F) \right) \right].$$

Furthermore, for any $1 \le k \le q$, $r \ge 1$ and $\|F\| \le 1$,

$$\left| \mathbb{E}\big[\overline{\gamma}_{z,n}^N(\overline{G}_n)^{-k} \times \big(1 - \overline{\gamma}_{z,n}^N(\overline{G}_n)\big)^{r+k} \big(\overline{\gamma}_{z,n}^N\big)^{\otimes q}\big(\overline{Q}_{n,n+1}^{\otimes q}(F)\big) \big] \right|$$

$$\le \mathbb{E}\big[\overline{\gamma}_{z,n}^N(\overline{G}_n)^{q-k} \times \big|1 - \overline{\gamma}_{z,n}^N(\overline{G}_n)\big|^{r+k} \big]$$

$$\le 2^{q-k}\big(\mathbb{E}\big[\big(\overline{\gamma}_{z,n}^N(\overline{G}_n) - 1\big)^{2(r+k)}\big]\big)^{1/2} \le 2^q \big(cn\big((r+k)\overline{g}_{n+1}\big)^2/N\big)^{(r+k)/2},$$

as soon as $N \ge cn((r+q)\overline{g}_{n+1})^2$. Recall that

$$\overline{G}_n^{\otimes l} \otimes \big(\overline{Q}_{n,n+1}^{\otimes q}(F)\big) = \overline{Q}_{n,n+1}^{\otimes(q+l)}\big(1^{\otimes l} \otimes F\big)$$

and combine (5.12) with Corollary 5.20, to obtain

$$\left| \mathbb{P}_{z,n+1}^{(N,q)}(F) - \frac{(q+r)!}{(q-1)!} \sum_{0 \le l \le r} \frac{1}{(q+l)} \frac{(-1)^l}{l!(r-l)!} \Upsilon_{z,n}^{(N,l+q)}\big[\overline{Q}_{n,n+1}^{\otimes(l+q)}\big(1^{\otimes l} \otimes F\big)\big] \right|$$

$$\le a(q)\big(cn\big((r+q)\overline{g}_{n+1}\big)^2/N\big)^{(r+1)/2},$$

for any $r \ge 0$ and some finite constant $a(q) < \infty$. We obtain the kth differential operator formulae (5.13) by choosing $r = 2k$. After some elementary manipulations, this yields the estimate

$$\left\| \mathbb{P}_{z,n+1}^{(N,q)} - \eta_{n+1}^{\otimes q} - \sum_{1 \le k \le r} \frac{1}{N^k} d^{(k)}\mathbb{P}_{z,n+1}^{(q)} \right\|_{\mathrm{tv}} \le b(q)\big(cn\big((q+r)\overline{g}_{n+1}\big)^2/N\big)^{r+1/2},$$

with some finite constant $b(q) < \infty$ and the derivatives operators $d^{(k)}\mathbb{P}_{z,n+1}^{(q)}$ given in (5.13). It is now straightforward to complete the proof. □

It is instructive to derive explicit expressions for the derivatives – this is one of the topics addressed below. To facilitate these developments we make explicit the first order derivative in a simple case. For $k = q = 1$, and any function f on S_n, with $\eta_n(f) = 0$, using the first order expansions that will be stated in Corollary 5.24, it is readily checked that

$$d^{(1)}\mathbb{P}_{z,n+1}^{(1)}(f) = \sum_{0 \le k \le n} \overline{Q}_{k,n+1}(f)(z_k) - \sum_{0 \le k \le n} \eta_k\big(\overline{Q}_{k,n+1}(1)\overline{Q}_{k,n+1}(f)\big).$$

5.4. *Infected forest expansions*

We know that $\mathbb{P}_{z,n}^{(q,N)}$ has derivatives of all orders and can be expanded in terms of the derivatives of $\Upsilon_{z,n}^{(N)}$. In turn, these derivatives can be expanded in terms of the elementary integral operators $\Delta_{z,n,\mathbf{c}}^{(q)}$. However, because of the symmetries of Feynman–Kac models, many of these operators coincide and this expansion is inefficient, both computationally and theoretically. This section clarifies these issues and eliminates redundancies in the combinatorial expansions of the derivatives.

The results in this section build largely on [21]. We therefore skip the details of the arguments as they follow closely the ones in [21] and refer the reader to that article for further details on the definitions, proofs, reasoning, etc. on trees, forests and jungles.

5.4.1. *Forests and jungles*
We start by recalling some classical terminology on trees and forests introduced in [21].

A tree is a (isomorphism class of) finite non-empty oriented connected graph $\mathbf{t}$ without loops such that any vertex of $\mathbf{t}$ has at most one outgoing edge. Paths are oriented from the vertices to the root. The height of a tree is the maximum lenght of a path. Similarly, the level of a vertex in a tree is the length of the path that connects it to the root. This definition extends in a straightforward way to the objects introduced below (forests and jungles).
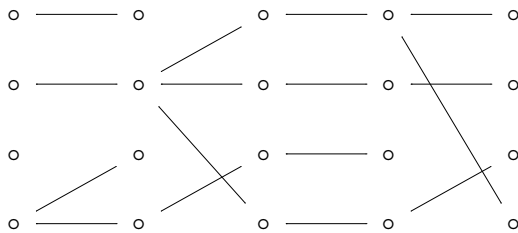
Fig. 4. A forest representing a sequence of maps in $([4]^{[4]})^4$.

A forest $\mathbf{f}$ is a multi-set 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 $T_0 = \varnothing$ as a unit. Since the algebraic notation is the most convenient, we write $\mathbf{f} = \mathbf{t}_1^{m_1}, \ldots, \mathbf{t}_k^{m_k}$, for the forest with the tree $\mathbf{t}_i$ appearing with multiplicity m_i, $i \leq k$. When $\mathbf{t}_i \neq \mathbf{t}_j$ for $i \neq j$, we say that $\mathbf{f}$ is written in normal form.

The set of forests with height $(n + 1)$, and with q vertices at each level set is written as $Forest_{q,n}$.

The sequence $\mathbf{a} = (a_0, \ldots, a_n) \in \mathcal{A}_{q,n} := ([q]^{[q]})^{n+1}$ is naturally associated with a forest $F(\mathbf{a})$: the one with one vertex for each element of $[q]^{n+1}$, and an edge for each pair $(i, a_k(i))$, $i \in [q]$. 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 such jungles is written $Jungle_{q,n}$. Figure 4 is a graphical representation of a jungle (for consistency with the probabilistic interpretation of heights and levels as time-indices, we represent trees, forest and jungles *horizontally* and from left to right – roots are on the left!).

The group $\mathcal{G}_{q,n} := \mathcal{G}_q^{n+2}$ also acts naturally on sequences of maps $\mathbf{a} \in \mathcal{A}_{q,n}$, and on jungles $J(\mathbf{a}) \in Jungle_{q,n}$ by permuting the vertices at each level. More precisely, for all $\mathbf{a} \in \mathcal{A}_{n,q}$ and all $\boldsymbol{\sigma} = (\sigma_0, \ldots, \sigma_{n+1}) \in \mathcal{G}_{q,n}$, we use the two formulae

$$\boldsymbol{\sigma}(\mathbf{a}) := \left(\sigma_0 a_0 \sigma_1^{-1}, \sigma_1 a_1 \sigma_2^{-1}, \ldots, \sigma_n a_n \sigma_{n+1}^{-1}\right) \quad \text{and} \quad \boldsymbol{\sigma} J(\mathbf{a}) := J\big(\boldsymbol{\sigma}(\mathbf{a})\big). \tag{5.15}$$

We note that if two sequences $\mathbf{a}$ and $\mathbf{a}' \in \mathcal{A}_{q,n}$ differ only by the order of the vertices in $J(\mathbf{a})$ and $J(\mathbf{a}')$ (i.e. by the action of an element of $\mathcal{G}_{q,n}$) then the associated forests are identical: $F(\mathbf{a}) = F(\mathbf{a}')$. The converse is also true: if $F(\mathbf{a}) = F(\mathbf{a}')$, then $J(\mathbf{a})$ and $J(\mathbf{a}')$ differ only by the ordering of the vertices, since they have the same underlying non planar graph. Here, $\mathbf{a}$ and $\mathbf{a}'$ belong to the same orbit

$$[\mathbf{a}] := \left\{\boldsymbol{\sigma}(\mathbf{a}) : \boldsymbol{\sigma} \in \mathcal{G}_{q,n}\right\},$$

under the action of $\mathcal{G}_{q,n}$. In particular, the set of equivalence classes of jungles in $Jungle_{q,n}$ under the action of the permutation groups $\mathcal{G}_{q,n}$ is a bijection with both the set of $\mathcal{G}_{q,n}$-orbits of maps in $\mathcal{A}_{q,n}$ and the set of forests $Forest_{q,n}$.

Writing

$$\text{Stab}(\mathbf{a}) := \left\{\boldsymbol{\tau} \in \mathcal{G}_{q,n} : \boldsymbol{\tau}(\mathbf{a}) = \mathbf{a}\right\}$$

for the stabilizer subgroup of $\mathbf{a}$, the class formula yields

$$\#[\mathbf{a}] = \#\mathcal{G}_{q,n}/\#\text{Stab}(\mathbf{a}) = (q!)^{n+2}/\#\text{Stab}(\mathbf{a}).$$

To compute the cardinality of the set $\text{Stab}(\mathbf{a})$ in terms of forests and trees, we denote by $\text{Cut}(\mathbf{t})$ the forest deduced from cutting the root of the tree $\mathbf{t}$; that is, removing its root vertex, and all its incoming edges. Conversely, we denote by $\text{Cut}^{-1}(\mathbf{f})$ the tree deduced from the forest $\mathbf{f}$ by adding a common root to its rooted tree. The symmetry multiset $\mathbf{S}(\mathbf{t})$ of a tree $\mathbf{t} = \text{Cut}^{-1}(\mathbf{t}_1^{m_1}, \ldots, \mathbf{t}_k^{m_k})$ associated with a forest written in normal form, is defined by $\mathbf{S}(\mathbf{t}) := (m_1, \ldots, m_k)$. The symmetry multiset of a forest is

$$\mathbf{S}\big(\mathbf{t}_1^{m_1} \ldots \mathbf{t}_k^{m_k}\big) := \big(\underbrace{\mathbf{S}(\mathbf{t}_1), \ldots, \mathbf{S}(\mathbf{t}_1)}_{m_1\text{-terms}}, \ldots, \underbrace{\mathbf{S}(\mathbf{t}_k), \ldots, \mathbf{S}(\mathbf{t}_k)}_{m_k\text{-terms}}\big).$$

We also extend Cut(**f**) to forests $\mathbf{f} = \mathbf{t}_1^{m_1}, \ldots, \mathbf{t}_k^{m_k}$ by setting

$$\text{Cut}(\mathbf{f}) = \text{Cut}(\mathbf{t}_1)^{m_1} \cdots \text{Cut}(\mathbf{t}_k)^{m_k}, \tag{5.16}$$

where $\text{Cut}(\mathbf{t}_i)^{m_i}$ stands for the forest deduced from $\text{Cut}(\mathbf{t}_i)$ repeated m_i times. Combining the class formula with a recursion with respect to the height parameter, we obtain

$$\#([\mathbf{a}]) = (q!)^{n+2}/\#(\text{Stab}(\mathbf{a})) \quad \text{with } \#(\text{Stab}(\mathbf{a})) = \prod_{i=-1}^{n} \mathbf{S}(\text{Cut}^i(F(\mathbf{a})))!, \tag{5.17}$$

where we have used the multi-index factorial notation $(n_1, \ldots, n_k) = n_1! \cdots n_k!$, for any $n_k \in \mathbb{N}$, with $k \geq 0$. A detailed proof of this closed formula is provided in [21].

We leave it to the reader check that, for example, for **a** as in the above graphical representation, $\#(\text{Stab}(\mathbf{a})) = 1 \cdot 1 \cdot 2! \cdot 2! = 4$ and $\#([\mathbf{a}]) = (4!)^4 \cdot 3!$.

5.4.2. *Infected forests*

Recall that the study of PG samplers requires the introduction of sequences of mappings $\mathbf{c} = (\mathbf{a}, \mathbf{b}) \in \mathcal{J}_{q,n}$, where the maps b_k can be thought of as "infections" (using the terminology previously introduced). The infection of a jungle $J(\mathbf{a})$ (or of the associated sequence of maps **a**) is defined accordingly by a sequence of functions $\mathbf{b} = (b_0, \ldots, b_n) \in (\{0, 1\}^{[q]})^{n+1}$.

Figure 5 provides an example of an infected planar forest.

By construction, there are $\prod_{0 \leq k \leq n} \binom{q}{i_k}$ ways of infecting a given forest with $0 \leq i_k \leq q$ infections at each level $0 \leq k \leq n$. Some of them are clearly equivalent. More precisely, we consider the following equivalence relation on infected jungles

$$(\mathbf{a}, \mathbf{b}) \sim (\mathbf{a}', \mathbf{b}') \Longleftrightarrow \exists \sigma \in \mathcal{G}_{q,n} : \sigma(\mathbf{a}, \mathbf{b}) = (\mathbf{a}', \mathbf{b}').$$

The equivalence classes are denoted by

$$[\mathbf{a}, \mathbf{b}] := \{\sigma(\mathbf{a}, \mathbf{b}) : \sigma \in \mathcal{G}_{q,n}\} = \{(\sigma(\mathbf{a}), \mathbf{b}\bar{\sigma}^{-1}) : \sigma \in \mathcal{G}_{q,n}\}$$

with

$$\bar{\sigma} := (\sigma_1, \ldots, \sigma_{n+1}) \quad \text{and} \quad \bar{\sigma}^{-1} = (\sigma_1^{-1}, \ldots, \sigma_{n+1}^{-1}).$$
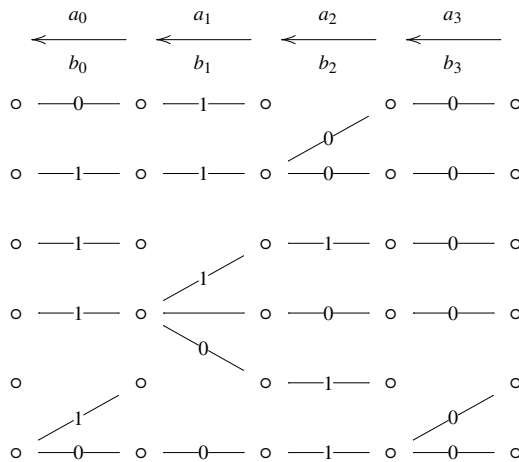

Fig. 5. Example of an infected planar forest of height 4 with 5 trees and 6 leaves, and the corresponding sequence of infection mappings. The infection is represented by the label 1, and the non-infection by the label 0 on the edges of the jungle.

The definitions of forests and jungles discussed in the previous section extend in a straightforward way to the infected case (edges being colored by 0 or 1). It is natural to associate with a sequence $(\mathbf{a}, \mathbf{b})$ an infected forest $F(\mathbf{a}, \mathbf{b})$: the one with one vertex for each element of $[q]^{n+1}$, and an infected edge for each triplet $(i, b_k(i), a_k(i)), i \in [q]$. The set of infected forests is a bijection with the set of $\mathcal{G}_{q,n}$-orbits of maps in $\mathcal{J}_{q,n}$.

The class formula yields once again a way to compute the cardinals of the classes $[\mathbf{a}, \mathbf{b}]$ from the action of the symmetry group $\mathcal{G}_{q,n}$.

Lemma 5.22. *The number of infected jungles in* $[\mathbf{a}, \mathbf{b}]$ *is*

$$\#[\mathbf{a}, \mathbf{b}] = (q!)^{n+2} / \mathrm{Stab}_{\mathbf{a}}(\mathbf{b}) = \#[\mathbf{a}] \times \frac{\#(\mathrm{Stab}(\mathbf{a}))}{\#(\mathrm{Stab}_{\mathbf{a}}(\mathbf{b}))},$$

with

$$\mathrm{Stab}_{\mathbf{a}}(\mathbf{b}) := \left\{ \tau \in \mathrm{Stab}(\mathbf{a}) : \mathbf{b}\overline{\tau} = \mathbf{b} \right\}.$$

Similarly to the non-infected case, $\#(\mathrm{Stab}_{\mathbf{a}}(\mathbf{b}))$ can be computed inductively, following essentially the same principles. We describe briefly how to do so.

Let $\mathbf{t}_1, \ldots, \mathbf{t}_n$ and $\mathbf{t}'_1, \ldots, \mathbf{t}'_m$ be two families of distinct infected trees and $l_i, i = 1, \ldots, n, p_j, j = 1, \ldots, m$ two sequences of positive integers. We write $\mathbf{t}_1^{l_1}, \ldots, \mathbf{t}_n^{l_n} \circledast \mathbf{t}_1'^{p_1}, \ldots, \mathbf{t}_n'^{p_m}$ for the infected tree obtained by joining, for $i = 1, \ldots, n, l_i$ copies of $\mathbf{t}_i$ to a common root with infection index 0 and for $i = 1, \ldots, m, p_i$ copies of $\mathbf{t}'_i$ to the same common root with infection index 1. Any infected tree $\mathbf{t}$ can be written uniquely in this way: we write $\mathbf{S}'(\mathbf{t}) = (l_1, \ldots, l_n, p_1, \ldots, p_m)$ for the corresponding multiset and call it the symmetry multiset of $\mathbf{t}$.

Cuts of infected trees and infected forests are infected forests that are defined as in the non-infected case by removing the root and erasing all infected edges connected to the root. A (right only) inverse operation Cut^{-1} acting on an infected forest $\mathbf{t}_1^{k_1}, \ldots, \mathbf{t}_n^{k_n}$ is defined by linking all the infected trees to a common root with non infected edges.

Mimicking the inductive arguments for counting jungles using cardinals of stabilizers in [21], we obtain

$$\mathrm{Stab}_{\mathbf{a}}(\mathbf{b}) = \prod_{i=-1}^{n} \mathbf{S}'\left(\mathrm{Cut}^i\left([\mathbf{a}, \mathbf{b}]\right)\right)!. \tag{5.18}$$

5.4.3. *Expectation operators on infected forests*

Recall that $\mathcal{J}_{q,n}$ is the set of $(n+1)$ mappings $\mathbf{c} = (\mathbf{a}, \mathbf{b}) = (c_0, \ldots, c_n)$ with $c_k = (a_k, b_k) \in I_q(p_k^1, p_k^2)$, for any $0 \le k \le n$.

For any symmetric function F on S_n^q, and any $\mathbf{c} = (\mathbf{a}, \mathbf{b})$ and $\mathbf{c}' := (\mathbf{a}', \mathbf{b}')$ we have

$$\mathbf{c} \sim \mathbf{c}' \quad \Longrightarrow \quad \Delta_{z,\mathbf{c}}^{(q)}(F) = \Delta_{z,\mathbf{c}'}^{(q)}(F).$$

We check this claim by using the result that for any $a_1, a_2 \in [q]^{[q]}$, and any $b \in \{0, 1\}^{[q]}$, and $\sigma \in \mathcal{G}_q$,

$$\mathbb{A}_{a_1}\mathbb{A}_{a_2} = \mathbb{A}_{a_1 a_2} \quad \text{and} \quad \mathbb{B}_{z,b} = \mathbb{A}_{\sigma}\mathbb{B}_{z,b\sigma}\mathbb{A}_{\sigma^{-1}}.$$

Thus, for any $\mathbf{f} \in \mathcal{F}_{q,n}$, we can define unambiguously $\Delta_{z,\mathbf{f}}^{(q)} = \Delta_{z,\mathbf{c}}^{(q)}$ for any choice $\mathbf{c}$ of a representative of $\mathbf{f}$ in $\mathcal{J}_{q,n}$.

We also denote by $\mathcal{F}_q(\mathbf{p}_n)$ the set of forests with p_k^1-coalescences and p_k^2 infections at each level $0 \le k \le n$. By construction, these forests are associated with the mappings $\mathbf{c}_n \in \mathbf{I}_q(\mathbf{p}_n)$. In this notation, the operators (5.8) can be rewritten in terms of the expectations operators on the set of infected forests

$$\Delta_{z,\mathbf{p}_n}^{(q)} = \sum_{\mathbf{c}_n \in \mathbf{I}_q(\mathbf{p}_n)} \mathcal{C}_{(\mathbf{p}_n)}^{(q)}(\mathbf{c}_n) \Delta_{z,\mathbf{c}_n}^{(q)} = \sum_{\mathbf{f} \in \mathcal{F}_q(\mathbf{p}_n)} \lambda_{q,\mathbf{p}_n}(\mathbf{f}) \Delta_{z,\mathbf{f}}^{(q)}, \tag{5.19}$$

with the probability measure $\lambda_{q,\mathbf{p}_n}$ given by

$$\lambda_{q,\mathbf{p}_n}(\mathbf{f}) = \#(\mathbf{f}) / \#\left(\mathbf{I}_q(\mathbf{p}_n)\right),$$

where we used the shortcut notation $\#(\mathbf{f}) := \#[\mathbf{c}]$ for an arbitrary representative of $\mathbf{f}$ in $\mathcal{J}_{q,n}$. We summarize the above discussion with the following theorem.

Theorem 5.23. *For any $m \geq 0$,*

$$d^{(m)}\Upsilon_{z,n}^{(q)} = \sum_{\mathbf{p}_n \in \mathcal{T}_{q,n}^{(m)}} \tau_{q,\mathbf{p}_n}^{(m)}\left(\sum_{\mathbf{f} \in \mathcal{F}_q(\mathbf{p}_n)} \lambda_{q,\mathbf{p}_n}(\mathbf{f}) \Delta_{z,\mathbf{f}}^{(q)} \right).$$

5.4.4. *Infected forests*

The first order derivative is expressed in terms of two classes of infected forests. The explicit description of the second derivative depends on 20 different types of infected forests. We investigate them in this section.

Fixing $3 < q < N$ and the time horizon n, there exists a single forest $\mathbf{f}_0$ with trivial trees with no infection. There is also a single non infected forest $\mathbf{f}_{1,0}^k$ with only one coalescence at level k. We also have a single forest $\mathbf{f}_{0,1}^k$ with trivial trees and an infection at level k. Figure 6 gives a synthetic description of these forests.

The corresponding measures are $\Delta_{z,\mathbf{f}_0}^{(q)} = \eta_n^{\otimes q}$, and the pair of measures

$$\Delta_{z,\mathbf{f}_{1,0}^k}^{(q)} = \eta_n^{\otimes(q-2)} \otimes \left[\int \eta_k(dw)(\delta_w \overline{Q}_{k,n})^{\otimes 2} \right] \quad \text{and} \quad \Delta_{z,\mathbf{f}_{0,1}^k}^{(q)} = \eta_n^{\otimes(q-1)} \otimes \delta_{z_k} \overline{Q}_{k,n}. \tag{5.20}$$

It is also immediate to check using (5.18) that

$$\#(\mathbf{f}_0) = q!^{n+1} \#\big(\mathbf{f}_{1,0}^k\big) = q!^{n+2}/\big((q-2)!2!\big) \quad \text{and} \quad \#\big(\mathbf{f}_{0,1}^k\big) = q!^{n+1}q.$$

There are two non-infected forests $\mathbf{f}_{2,0}^{k,1}$ and $\mathbf{f}_{2,0}^{k,2}$ with two coalescences at level k. The first one has a non-trivial tree with three leaves, the second one has two trees with two leaves (see Figure 7).

The corresponding measures are

$$\Delta_{z,\mathbf{f}_{2,0}^{k,1}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_k(dw)\big(\delta_w \overline{Q}_{k,n}^{\otimes 3}\big) \right],$$

$$\Delta_{z,\mathbf{f}_{2,0}^{k,2}}^{(q)}(F) = \eta_n^{\otimes(q-4)} \otimes \left\{ \int \eta_k(dw_1)\eta_k(dw_2)\big[(\delta_{w_1}\overline{Q}_{k,n})^{\otimes 2} \otimes (\delta_{w_2}\overline{Q}_{k,n})^{\otimes 2}\big] \right\} \tag{5.21}$$
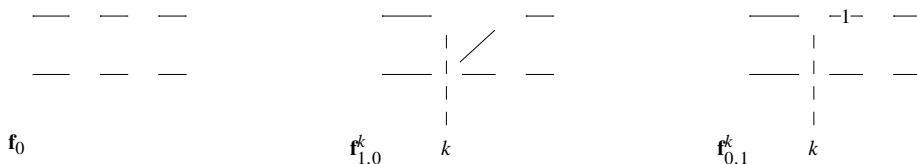

Fig. 6. Symbolic representation of a forest with trivial trees and no infection, resp. only one coalescence at level k, resp. trivial trees and an infection at level k.
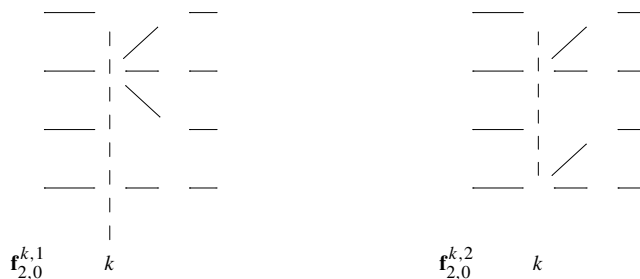

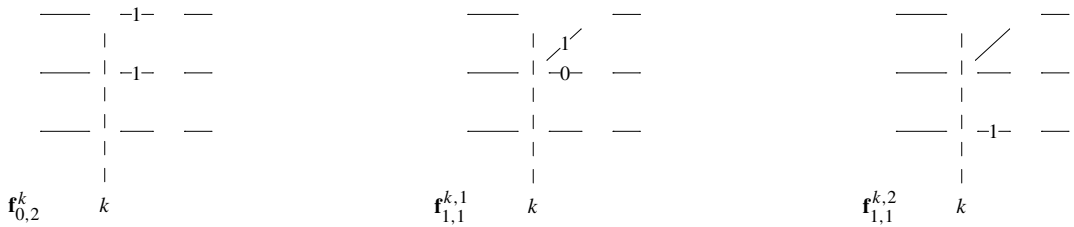Fig. 7. Symbolic representation of non-infected forests with two coalescences at level k.

Fig. 8. Symbolic representation of forests with two infections at level k, resp. one infection and a coalescence at level k.
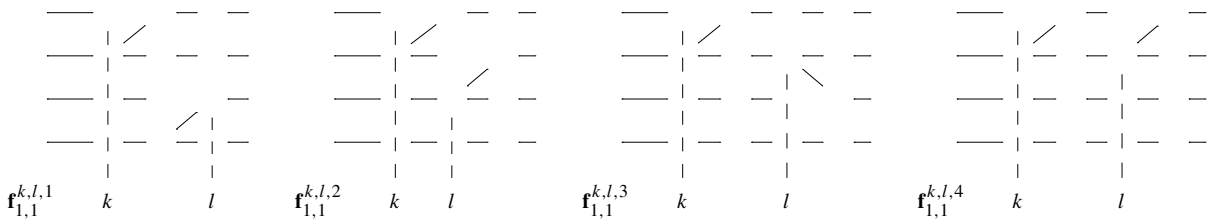

Fig. 9. Symbolic representation of $\mathbf{f}_{1,1}^{k,l,1}, \mathbf{f}_{1,1}^{k,l,2}, \mathbf{f}_{1,1}^{k,l,3}, \mathbf{f}_{1,1}^{k,l,4}$.

with $\#(\mathbf{f}_{2,0}^{k,1}) = (q!)^{n+2}/((q-3)!3!)$ and $\#(\mathbf{f}_{2,0}^{k,2}) = (q!)^{n+2}/((q-4)!2^3)$.

There is also a single non coalescent forest $\mathbf{f}_{0,2}^{k}$ with two trivial infected trees at level k. There are two forests $\mathbf{f}_{1,1}^{k,i}$, $i = 1, 2$, with one infection and one coalescence at level k. The first one has a single coalescent tree with only one infected leaf. The last one has a non infected coalescent tree and a single infected trivial tree (see Figure 8).

The corresponding measures are

$$
\Delta_{z,\mathbf{f}_{0,2}^{k,1}}^{(q)} = \eta_n^{\otimes(q-2)} \otimes (\delta_{z_k}\overline{Q}_{k,n})^{\otimes 2}, \qquad \Delta_{z,\mathbf{f}_{1,1}^{k,1}}^{(q)} = \Delta_{z,n,\mathbf{f}_{0,1}^{k}}^{(q)},
$$

$$
\Delta_{z,\mathbf{f}_{1,1}^{k,2}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_k(dw)(\delta_w\overline{Q}_{k,n})^{\otimes 2} \right] \otimes (\delta_{z_k}\overline{Q}_{k,n}).
$$

(5.22)

One checks that $\#(\mathbf{f}_{0,2}^{k}) = q!^{n+1}q(q-1)/2$, $\#(\mathbf{f}_{1,1}^{k,1}) = q!^{n+1}q(q-1)$ and $\#(\mathbf{f}_{1,1}^{k,2}) = \frac{(q!)^{n+2}}{2(q-3)!}$.

We also have the traditional four non infected forests $\mathbf{f}_{1,1}^{k,l,i}$, $i = 1, 2, 3, 4$ with two coalescences at level k and l [21]. The first one has two coalescent trees with all the leaves at level n. The second one also has two coalescent trees but one has two leaves at level n, the other has a leaf at level l and another at level n. The third one has a single coalescent tree with three leaves at level n, and a coalescent branch at level l. The last one has a single coalescent tree with two leaves at level n and a coalescent branch at level l (see Figure 9).

In this case, we readily check that

$$
\#(\mathbf{f}_{1,1}^{k,l,1}) = \frac{q!^{n+2}}{4(q-4)!}, \qquad \#(\mathbf{f}_{1,1}^{k,l,2}) = \frac{q!^{n+2}}{(q-3)!2!}, \qquad \#(\mathbf{f}_{1,1}^{k,l,3}) = \frac{q!^{n+2}}{(q-3)!2!}, \qquad \#(\mathbf{f}_{1,1}^{k,l,4}) = \frac{q!^{n+2}}{(q-2)!2!}
$$

and the corresponding measures are

$$
\Delta_{z,\mathbf{f}_{1,1}^{k,l,1}}^{(q)} = \eta_n^{\otimes(q-4)} \otimes \left[\int \eta_k(du)(\delta_u\overline{Q}_{k,n})^{\otimes 2} \right] \otimes \left[\int \eta_l(dv)(\delta_v\overline{Q}_{l,n})^{\otimes 2} \right],
$$

$$
\Delta_{z,\mathbf{f}_{1,1}^{k,l,2}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_k(du)\overline{Q}_{k,l}(1)(u)\delta_u\overline{Q}_{k,n} \right] \otimes \left[\int \eta_l(dv)(\delta_v\overline{Q}_{l,n})^{\otimes 2} \right],
$$

(5.23)

Fig. 10. Symbolic representation of $\mathbf{f}_{0,1,1}^{k,l,1}, \mathbf{f}_{0,1,1}^{k,l,2}$.
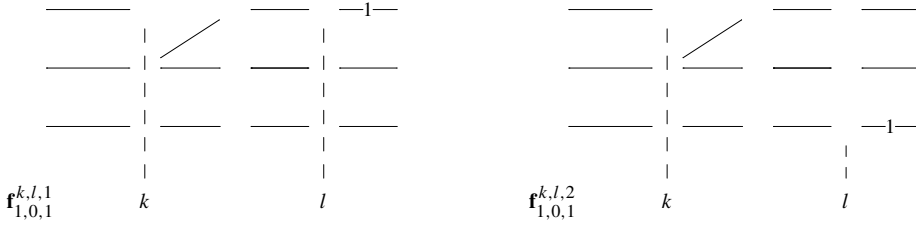
Fig. 11. Symbolic representation of $\mathbf{f}_{1,0,1}^{k,l,1}, \mathbf{f}_{1,0,1}^{k,l,2}$.

$$\Delta_{z,\mathbf{f}_{1,1}^{k,l,3}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_k(du) \left(\left\{ \int \overline{Q}_{k,l}(u,dv)(\delta_v \overline{Q}_{l,n})^{\otimes 2} \right\} \otimes \delta_u \overline{Q}_{k,n} \right) \right],$$

$$\Delta_{z,\mathbf{f}_{1,1}^{k,l,4}}^{(q)} = \eta_n^{\otimes(q-2)} \otimes \left[\int \eta_k(du) \overline{Q}_{k,l}(1)(u) \overline{Q}_{k,l}(u,dv)(\delta_v \overline{Q}_{l,n})^{\otimes 2} \right].$$

We also have two non coalescent forests $\mathbf{f}_{0,1,1}^{k,l,i}$, $i = 1, 2$, with two infections at level k and l. The first one has two infected trivial trees. The second one has a trivial tree with two infections (see Figure 10).

In this case, we have $\#(\mathbf{f}_{0,1,1}^{k,l,1}) = q!^{n+1} q(q-1)$ and $\#(\mathbf{f}_{0,1,1}^{k,l,2}) = q!^{n+1} q$, and

$$\Delta_{z,\mathbf{f}_{0,1,1}^{k,l,1}}^{(q)} = \eta_n^{\otimes(q-2)} \otimes \delta_{z_k} \overline{Q}_{k,n} \otimes \delta_{z_l} \overline{Q}_{l,n} \quad \text{and}$$

$$\Delta_{z,\mathbf{f}_{0,1,1}^{k,l,2}}^{(q)} = \overline{Q}_{k,l}(1)(z_k) \left[\eta_n^{\otimes(q-1)} \otimes \delta_{z_l} \overline{Q}_{l,n} \right]. \tag{5.24}$$

We also have two forests $\mathbf{f}_{1,0,1}^{k,l,i}$, $i = 1, 2$, with a coalescence at level k and an infection at level $l > k$. The first one has a coalescent tree with an infection. The second one has a non infected coalescent tree and an infected trivial tree (see Figure 11). In this case we have $\#(\mathbf{f}_{1,0,1}^{k,l,1}) = q!^{n+2}/(q-2)!$, and $\#(\mathbf{f}_{1,0,1}^{k,l,2}) = q!^{n+2}/(2(q-3)!)$. The corresponding measures are

$$\Delta_{z,\mathbf{f}_{1,0,1}^{k,l,1}}^{(q)} = \eta_n^{\otimes(q-2)} \otimes \left[\int \eta_k(du) \overline{Q}_{k,l}(1)(u) \delta_u \overline{Q}_{k,n} \right] \otimes \delta_{z_l} \overline{Q}_{l,n},$$

$$\Delta_{z,\mathbf{f}_{1,0,1}^{k,l,2}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_k(du)(\delta_u \overline{Q}_{k,n})^{\otimes 2} \right] \otimes \delta_{z_l} \overline{Q}_{l,n}. \tag{5.25}$$

Finally, there are three forests $\mathbf{f}_{0,1,0,1}^{k,l,i}$, $i = 1, 2, 3$, with an infection at k and a coalescence at level $l > k$. The first one has a infected tree with a leaf at level n and a non-infected coalescent tree. The second one has a infected tree with a leaf at level l and a non-infected coalescent tree. And finally, the last one has an infected coalescent tree (see Figure 12).

In this case we have $\#(\mathbf{f}_{0,1,0,1}^{k,l,1}) = q!^{n+2}/(2(q-3)!)$ and for any $i \in \{2, 3\}$ $\#(\mathbf{f}_{1,0,1}^{k,l,i}) = q!^{n+2}/(2(q-2)!)$. In addition, the corresponding measures are

$$\Delta_{z,\mathbf{f}_{0,1,0,1}^{k,l,1}}^{(q)} = \eta_n^{\otimes(q-3)} \otimes \left[\int \eta_l(du)(\delta_u \overline{Q}_{l,n})^{\otimes 2} \right] \otimes \delta_{z_k} \overline{Q}_{k,n},$$
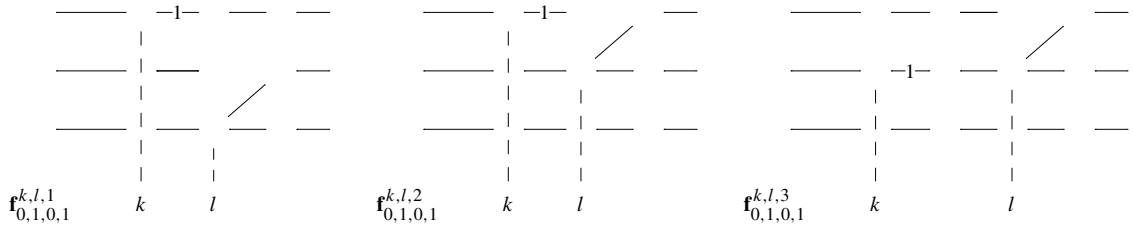
Fig. 12. Symbolic representation of $\mathbf{f}_{0,1,0,1}^{k,l,1}, \mathbf{f}_{0,1,0,1}^{k,l,2}, \mathbf{f}_{0,1,0,1}^{k,l,3}$.

$$\Delta_{z,\mathbf{f}_{0,1,0,1}^{k,l,2}}^{(q)} = \overline{Q}_{k,l}(1)(z_k)\left[\eta_n^{\otimes(q-2)} \otimes \left\{\int \eta_l(du)(\delta_u \overline{Q}_{l,n})^{\otimes 2}\right\}\right],$$

$$\Delta_{z,\mathbf{f}_{0,1,0,1}^{k,l,3}}^{(q)} = \eta_n^{\otimes(q-2)} \otimes \left[\int \overline{Q}_{k,l}(z_k,du)(\delta_u \overline{Q}_{l,n})^{\otimes 2}\right].$$

(5.26)

For any multi-index κ, and any integer i we set

$$\overline{\Delta}_{z,\mathbf{f}_\kappa^{\cdot,\cdot,i}}^{(q)} := \sum_{0 \le k < l \le n} \overline{\Delta}_{z,\mathbf{f}_\kappa^{k,l,i}}^{(q)} \quad \text{with } \overline{\Delta}_{z,n,\mathbf{f}_\kappa^{k,l,i}}^{(q)} := \Delta_{z,n,\mathbf{f}_\kappa^{k,l,i}}^{(q)} - \eta_n^{\otimes q}.$$

5.4.5. *First and second derivatives*

To describe with some precision the first two order derivatives of the mapping $N \mapsto \Upsilon_{z,n}^{(q)}$ we need to compute the expectation operators on random infected forests defined in (5.19). The ones associated with forests with at most one infection or one coalescence at some level only depend one class of forests. Thus, using (5.20) their description is immediate. Using (5.21), the centered operator associated with non infected forests with a couple of coalescence at some level is

$$\overline{\Delta}_{z,\mathbf{f}_{2,0}^{\cdot,\star}}^{(q)} := \frac{1}{1+(3/4)(q-3)} \overline{\Delta}_{z,\mathbf{f}_{2,0}^{\cdot,1}}^{(q)} + \left(1 - \frac{1}{1+(3/4)(q-3)}\right)\overline{\Delta}_{z,\mathbf{f}_{2,0}^{\cdot,2}}^{(q)}.$$

Similarly, by (5.22), the one associated with forests with a single coalescence and a single infection at some level is

$$\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\star}}^{(q)} := \frac{2}{q}\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,1}}^{(q)} + \left(1 - \frac{2}{q}\right)\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,2}}^{(q)}.$$

In view of (5.23), the centered expectation operator associated with forests with a single coalescence at two different levels is

$$\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\cdot,\star}}^{(q)} := \frac{(q-2)(q-3)}{(q-2)(q-3)+4(q-2)+2}\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\cdot,1}}^{(q)}$$

$$+ \frac{2(q-2)}{(q-2)(q-3)+4(q-2)+2}\left[\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\cdot,2}}^{(q)} + \overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\cdot,3}}^{(q)}\right]$$

$$+ \frac{2}{(q-2)(q-3)+4(q-2)+2}\overline{\Delta}_{z,\mathbf{f}_{1,1}^{\cdot,\cdot,4}}^{(q)}.$$

Using (5.24), the one associated with non coalescent forests with a single infection at two different levels is

$$\overline{\Delta}_{z,\mathbf{f}_{0,1,1}^{\cdot,\cdot,\star}}^{(q)} := \left(1 - \frac{1}{q}\right)\overline{\Delta}_{z,\mathbf{f}_{0,1,1}^{\cdot,\cdot,1}}^{(q)} + \frac{1}{q}\overline{\Delta}_{z,\mathbf{f}_{0,1,1}^{\cdot,\cdot,2}}^{(q)}.$$

Finally, using (5.25) and (5.26), the operators associated with a single coalescence and a single infection at two different levels are given by

$$\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{1,0,1}} := \frac{2}{q}\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,1}_{1,0,1}} + \left(1 - \frac{2}{q}\right)\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,2}_{1,0,1}}$$

and

$$\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{0,1,0,1}} := \left(1 - \frac{2}{q}\right)\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,1}_{0,1,0,1}} + \frac{1}{q}\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,2}_{0,1,0,1}} + \frac{1}{q}\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,3}_{0,1,0,1}}.$$

Expanding the formulae stated in Theorem 5.23, extending the combinatorial methods developed in [21] for computing the cardinals #($\mathbf{f}$) we prove the following expansions.

Corollary 5.24. *The first three derivatives of* $\Upsilon^{(N,q)}_{z,n}$ *are*

$$d^{(0)}\Upsilon^{(q)}_{z,n} = \eta^{\otimes q}_n,$$

$$d^{(1)}\Upsilon^{(q)}_{z,n} = \tau^{(1)}_{q,1,0}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{1,0}} + \tau^{(1)}_{q,0,1}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{0,1}},$$

$$d^{(2)}\Upsilon^{(q)}_{z,n}$$
$$= \tau^{(2)}_{q,1,0}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{1,0}} + \tau^{(2)}_{q,0,1}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{0,1}} + \tau^{(2)}_{q,1,1}\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\star}_{1,1}} + \tau^{(2)}_{q,2,0}\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\star}_{2,0}} + \tau^{(2)}_{q,0,2}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{0,2}}$$
$$+ \left(\tau^{(1)}_{q,1,0}\right)^2\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{1,1}} + \left(\tau^{(1)}_{q,0,1}\right)^2\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{0,1,1}} + \tau^{(1)}_{q,1,0}\tau^{(1)}_{q,0,1}\left\{\overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{1,0,1}} + \overline{\Delta}^{(q)}_{z,\mathbf{f}^{\cdot,\cdot,\star}_{0,1,0,1}}\right\}$$
$$+ n\tau^{(1)}_{q,0,0}\left[\tau^{(1)}_{q,1,0}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{1,0}} + \tau^{(1)}_{q,0,1}\overline{\Delta}^{(q)}_{z,\mathbf{f}_{0,1}}\right]$$

with the parameters $\tau^{(m)}_{q,p_1,p_2}$ *given in* (5.5).

When $q = 1$, all the terms are null except $\tau^{(1)}_{1,0,1} = 1 = -\tau^{(1)}_{1,0,0}$, and

$$d^{(1)}\Upsilon^{(1)}_{z,n} = \sum_{0 \leq k \leq n}\left[\Delta^{(1)}_{z,\mathbf{f}^k_{0,1}} - \eta_n\right] = \sum_{0 \leq k \leq n}\delta_{z_k}(\overline{Q}_{k,n} - \eta_n),$$

$$d^{(2)}\Upsilon^{(1)}_{z,n} = \overline{\Delta}^{(1)}_{z,\mathbf{f}^{\cdot,\cdot,2}_{0,1,1}} - n\overline{\Delta}^{(1)}_{z,\mathbf{f}_{0,1}}$$
$$= \sum_{0 \leq k < l \leq n}\left[\overline{Q}_{k,l}(1)(z_k)\delta_{z_l}\overline{Q}_{l,n} - \eta_n\right] - n\sum_{0 \leq k \leq n}\left[\delta_{z_k}\overline{Q}_{k,n} - \eta_n\right].$$

6. Some extensions and open questions

6.1. *Island type methodologies*

Particle MCMC methods are computationally intensive sampling techniques. As discussed in [33,55], parallel and distributed computations provide an appealing solution of tackling these issues. The central idea of Island models is to run N_2 particle models in parallel with N_1 individuals, instead of running a single particle model with $N_1 N_2$ particles. These N_2 batches are termed islands in reference to dynamic population models. Within each island the N_1 individuals evolve as a standard genetic type particle model. In this interpretation, island particle models can be thought as a parallel implementation of particle models. In the further development of this section, we show that these methodologies can also be used in a natural way to design island type particle MCMC samplers.

To design these models, we consider a collection of bounded and non-negative potential functions $\mathfrak{G}_n$ on some measurable state spaces $\mathfrak{E}_n$, with $n \in \mathbb{N}$. We let $\mathfrak{X}_n$ be a Markov chain on $\mathfrak{E}_n$ with initial distribution $\mu_0 \in \mathcal{P}(\mathfrak{E}_0)$ and

some Markov transitions $\mathfrak{M}_n$ from $\mathfrak{E}_{n-1}$ into $\mathfrak{E}_n$. The Feynman–Kac measures (μ_n, ν_n) associated with the parameters $(\mathfrak{G}_n, \mathfrak{M}_n)$ are defined for any $\mathfrak{f}_n \in \mathcal{B}(\mathfrak{E}_n)$ by the formulae

$$\mu_n(\mathfrak{f}_n) := \nu_n(\mathfrak{f}_n)/\nu_n(1) \quad \text{with } \nu_n(\mathfrak{f}_n) := \mathbb{E}\bigg(\mathfrak{f}_n(\mathfrak{X}_n) \prod_{0 \leq p < n} \mathfrak{G}_p(\mathfrak{X}_p)\bigg). \tag{6.1}$$

The many body N'-particle approximation

$$X'_n = \big\{X'^i_n\big\}_{1 \leq i \leq N'} \in S'_n := \mathfrak{E}_n^{[N']}$$

of these Feynman–Kac models is defined as in (2.5) by considering the evolution semigroup of the Feynman–Kac model μ_n.

We let M'_n be Markov transitions of X'_n and we consider the potential functions G'_n on S'_n defined by

$$G'_n(X'_n) = m(X'_n)(\mathfrak{G}_n) = \frac{1}{N'} \sum_{1 \leq i \leq N'} \mathfrak{G}_n(X'^i_n). \tag{6.2}$$

We let (η'_n, γ'_n) be the Feynman–Kac measures associated with the parameters (G'_n, M'_n). In this framework, the unbiasedness properties of the unnormalized Feynman–Kac particle measures take the form

$$f'_n(X'_n) = m(X'_n)(\mathfrak{f}_n)$$

$$\implies \quad \nu_n(\mathfrak{f}_n) = \mathbb{E}\bigg(\mathfrak{f}_n(\mathfrak{X}_n) \prod_{0 \leq p < n} \mathfrak{G}_p(\mathfrak{X}_p)\bigg) = \mathbb{E}\bigg(f'_n(X'_n) \prod_{0 \leq p < n} G'_p(X'_p)\bigg) = \gamma'_n(f'_n). \tag{6.3}$$

The path space version (η_n, γ_n) of these measures are defined by the Feynman–Kac measures associated with the historical process X_n and the potential function G_n given by

$$X_n = \big(X'_0, \ldots, X'_n\big) \in S_n = \big(S'_0 \times \cdots \times S'_n\big) \quad \text{and} \quad G_n(X_n) = G'_n(X'_n).$$

The mean field N-particle approximations $\xi'_n = (\xi'^i_n)_{1 \leq i \leq N}$ of the measures (η'_n, γ'_n) can be interpreted as a genetic type model with island type particles

$$\forall 1 \leq i \leq N \quad \xi'^i_n = \big\{\xi'^{i,j}_n\big\}_{1 \leq j \leq N'} \in S'_n := \mathfrak{E}_n^{[N']},$$

with mutation transitions M'_n and the selection potential functions G'_n given in (6.2).

By construction, the N-particle approximation ξ_n of the path space measures (η_n, γ_n) is a genealogical tree based model in the space of islands. Each particle

$$\xi^i_n = \big(\xi^i_{0,n}, \ldots, \xi^i_{n,n}\big) \in S_n = \big(\mathfrak{E}_0^{[N']} \times \cdots \times \mathfrak{E}_n^{[N']}\big)$$

represents the line of island ancestors $\xi^i_{p,n} \in \mathfrak{E}_p^{[N']}$ of the ith island $\xi^i_{n,n} = \xi'^i_n \in \mathfrak{E}_n^{[N']}$ at time n, at every level $0 \leq p \leq n$, with $1 \leq i \leq N$. In other words, $(\eta_n, \gamma_n, \xi_n)$ is the historical version of the Feynman–Kac model $(\gamma'_n, \eta'_n, \xi'_n)$. In this case, the dual mean field particle model $\mathcal{X}_n$ evolves on the state spaces $\mathcal{S}_n = S_n^{[N]}$, with a frozen trajectory of islands X_n.

This model can be interpreted as the evolution of N interacting islands

$$\forall 1 \leq i \leq N \quad \mathcal{X}^i_n = \big\{\mathcal{X}^{i,j}_n\big\}_{1 \leq j \leq N'} \in \mathfrak{E}_n^{[N']},$$

with N' individuals in each island. The PG models discussed in Section 4.3 can be used without further work to design island type particle Markov chain models with the target measure η_n. Using (6.3), we see that the S'_n-marginal of η_n can be used to compute any Feynman measures of the form (6.1). Similar constructions can be developed to design a backward-sampling based particle MCMC model.

These nested constructions can be iterated at any level. For a more thorough discussion on these island type particle methodologies, see [19,20], and the recent article [55]. An important open question is analysis of the convergence properties of the islands type particle models presented above in terms of the number of individual and the number of islands.

6.2. *Concluding remarks*

The duality relations and differential calculus developed in this article open up research problems in the field of Feynman–Kac particle models and PMCMC methods. One important problem is extending the results developed in the article to continuous time Feynman–Kac models observed discretely, where the model is also discretized to facilitate estimation. In this situation it is important to analyze the effect of the time discretization of the models on the corresponding numerical error, when we add such an extra level of approximation.

Our stability analysis is restricted to particle Gibbs methods on the space of ancestral lines. However, an alternative particle approach is to use a Metropolis–Hastings scheme to sample a target Boltzmann–Gibbs measure associated with a many-body Feyman–Kac measure. The resulting algorithm is sometimes called a particle independent Metropolis–Hastings algorithm [2]. An alternative approach is to use an interacting MCMC methodology to sample sequentially these many-body measures on path spaces (see for instance Section 2.2.1 and 2.2.5 in [25]). For parameter inference in Hidden Markov chain problems this strategy amounts to estimating at each time step the likelihood unbiasedly and then using a Metropolis–Hastings scheme to generate the parameters. In future work, we plan to analyze the convergence to equilibrium of these particle Metropolis–Hastings samplers and compare their stability properties to those of particle Gibbs sampling with a target many body Feynman–Kac measure.

Acknowledgement

References

[1] C. Andrieu, A. Doucet and R. Holenstein. Particle Markov chain Monte Carlo for efficient numerical simulation. In *Monte Carlo and Quasi-Monte Carlo Methods 2008* 45–60. P. L'Ecuyer and A. B. Owen (Eds). Springer, Berlin, 2009. MR2743887

[2] C. Andrieu, A. Doucet and R. Holenstein. Particle Markov chain Monte Carlo methods (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **72** (3) (2010) 1–269. MR2758115

[3] C. Andrieu, A. Lee and M. Vihola. Uniform ergodicity of the iterated conditional SMC and geometric ergodicity of particle Gibbs samplers. Preprint, 2013. Available at arXiv:1312.6432.

[4] R. Arratia and S. DeSalvo. Completely effective error bounds for Stirling numbers of the first and second kinds via Poisson approximation. Preprint, 2014. Available at arXiv:1404.3007.

[5] S. Barthelmé and N. Chopin. Expectation propagation for likelihood-free inference. *J. Amer. Statist. Assoc.* **109** (2014) 315–333. MR3180566

[6] R. Assaraf and M. Caffarel. A pedagogical introduction to quantum Monte Carlo. In *Mathematical Models and Methods for Ab Initio Quantum Chemistry* 45–73. M. Defranceschi and C. Le Bris (Eds). *Lecture Notes in Chemistry* **74**. Springer, Berlin, 2000. MR1857459

[7] E. Cancès, B. Jourdain and T. Lelièvre. Quantum Monte Carlo simulations of fermions. A mathematical analysis of the fixed-node approximation. *Math. Models Methods Appl. Sci.* **16** (9) (2006) 1403–1440. MR2254292

[8] O. Cappé, E. Moulines and T. Rydèn. *Inference in Hidden Markov Models*. Springer, New York, 2005. MR2159833

[9] R. Carmona, P. Del Moral, P. Hu and N. Oudjane. An introduction to particle methods with financial applications. In *Numerical Methods in Finance* 3–49. *Springer Proceedings in Mathematics* **12**. Springer, New York, 2012. MR3287602

[10] F. Cerou, P. Del Moral and A. Guyader. A non-asymptotic variance theorem for unnormalized Feynman–Kac particle models. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (3) (2011) 629–649. MR2841068

[11] H. P. Chan, P. Del Moral and A. Jasra. A sharp first order analysis of Feynman–Kac particle models. Preprint, 2014. Available at arXiv:1411.3800.

[12] N. Chopin and S. S. Singh. On particle Gibbs sampling. *Bernoulli* **21** (2015) 1855–1883.

[13] C. Berge. *Principles of Combinatorics*. Academic Press, New York, 1971. MR0270922

[14] L. Comtet. *Analyse Combinatoire. Tomes I, II. Collection SUP. Le Mathématicien 4* **5**. Presses Universitaires de France, Paris, 1970. MR0262087

[15] D. Creal. A survey of sequential Monte Carlo methods for economics and finance. *Econometric Rev.* **31** (3) (2012) 245–296.

[16] P. Del Moral. *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications. Probability and Its Applications*. Springer, New York, 2004. MR2044973

[17] P. Del Moral, M. Ledoux and L. Miclo. On contraction properties of Markov kernels. *Probab. Theory Related Fields* **126** (2003) 395–420. MR1992499

[18] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités XXXIV* 1–145. *Lecture Notes in Math.* **1729**. Springer, Berlin, 2000. MR1768060

[19] P. Del Moral and M. E. Rio. Concentration inequalities for mean field particle models. *Ann. Appl. Probab.* **21** (3) (2010) 1017–1052. MR2830611

[20] P. Del Moral. *Mean Field Simulation for Monte Carlo Integration*. *Chapman and Hall/CRC Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013.

[21] P. Del Moral, F. Patras and S. Rubenthaler. Coalescent tree based functional representations for some Feynman–Kac particle models. *Ann. Appl. Probab.* **19** (2) (2009) 1–50. MR2521888

[22] P. Del Moral, A. Doucet and S. S. Singh. A backward particle interpretation of Feynman–Kac formulae. *M2AN Math. Model. Numer. Anal.* **44** (5) (2010) 947–976. MR2731399

[23] P. Del Moral and L. Doucet. Particle motions in absorbing medium with hard and soft obstacles. *Stoch. Anal. Appl.* **22** (2004) 1175–1207. MR2089064

[24] P. Del Moral and J. Garnier. Genealogical particle analysis of rare events. *Ann. Appl. Probab.* **15** (4) (2005) 2496–2534. MR2187302

[25] P. Del Moral, P. Hu and L. M. Wu. On the concentration properties of interacting particle processes. *Found. Trends Mach. Learn.* **3** (3–4) (2012) 225–389.

[26] P. Del Moral, P. Jacob, A. Lee, L. Murray and G. W. Peters. Feynman–Kac particle integration with geometric interacting jumps. *Stoch. Anal. Appl.* **31** (2013) 830–871. MR3175799

[27] P. Del Moral and L. Miclo. On the stability of non linear semigroup of Feynman–Kac type. *Ann. Fac. Sci. Toulouse Math. (6)* **11** (2002) 135–175.

[28] D. N. DeJong, R. Liesenfeld, G. V. Moura, J. F. Richard and H. Dharmarajan. Efficient likelihood evaluation of state-space representations. *Rev. Econ. Stud.* **80** (2) (2013) 538–567. MR3054070

[29] A. Doucet, J. F. G. de Freitas and N. J. Gordon (Eds). *Sequential Monte Carlo Methods in Practice*. Springer, New York, 2001. MR1847783

[30] A. Doucet, M. K. Pitt, G. Deligiannidis and R. Kohn. Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator. *Biometrika* **102** (2015) 295–313.

[31] M. Dowd. Estimating parameters for a stochastic dynamic marine ecological system. *Environmetrics* **22** (4) (2011) 501–515. MR2843404

[32] M. Dowd and J. Ruth. Estimating behavioral parameters in animal movement models using a state-augmented particle filter. *Ecology* **92** (3) (2011) 568–575.

[33] G. Durham and J. Geweke. Massively parallel Sequential Monte Carlo for Bayesian inference. Unpublished manuscript, 2011. Available at http://www.censoc.uts.edu.au/pdfs/geweke_papers/gp_working_9.pdf.

[34] R. G. Everitt. Bayesian parameter estimation for latent Markov random fields and social networks. *J. Comput. Graph. Statist.* **21** (4) (2012) 940–960. MR3005805

[35] T. Flury and N. Shephard. Bayesian inference based only on simulated likelihood: Particle filter analysis of dynamic economic models. *Econometric Theory* **27** (5) (2011) 933–956. MR2843833

[36] A. Golightly and D. J. Wilkinson. Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo. *Interface Focus* **1** (6) (2011) 807–820.

[37] S. Henriksen, A. Wills, T. B. Schön and B. Ninness. Parallel implementation of particle MCMC methods on a GPU. In *Proceedings of the 16th IFAC Symposium on System Identification* **16** 1143–1148. Elsevier, Amsterdam, 2012.

[38] B. Jourdain, T. Lelièvre and M. El Makrini. Diffusion Monte Carlo method: Numerical analysis in a simple case. *M2AN Math. Model. Numer. Anal.* **41** (2007) 189–213. MR2339625

[39] N. Kantas, A. Doucet, S. Singh and J. M. Maciejowski. An overview of sequential Monte Carlo methods for parameter estimation in general state space models. In *15th IFAC Symposium on System Identification* **115** 774–785, 2009.

[40] T. Lelièvre, M. Rousset and G. Stoltz. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, London, 2010. MR2681239

[41] T. Lelièvre, M. Rousset and G. Stoltz. Computation of free energy differences through non-equilibrium stochastic dynamics: The reaction coordinate case. *J. Comput. Phys.* **222** (2) (2007) 624–643. MR2313418

[42] F. Lindsten, R. Douc and E. Moulines. Uniform ergodicity of the particle Gibbs sampler. *Scand. J. Statist.* To appear, 2015.

[43] H. Moradkhani, C. M. DeChant and S. Sorooshian. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter Markov chain Monte Carlo method. *Water Resour. Res.* **48** (2012) W12520.

[44] Y. Mishchenko, J. T. T. Vogelstein and L. Paninski. A Bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data. *Ann. Appl. Stat.* **5** (2B) (2011) 1229–1261. MR2849773

[45] T. Launay, A. Philippe and S. Lamarche. On particle filters applied to electricity load forecasting. *J. SFdS* **154** (2013) 1–36. MR3120434

[46] F. Lindsten, T. Schön and M. I. Jordan. Ancestor sampling for particle Gibbs. In *Conference in Advances in Neural Information Processing Systems* **25** 2600–2608. IEEE, New York, 2012.

[47] F. Lindsten and T. Schön. On the use of backward simulation in the particle Gibbs sampler. In *37th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* 3845–3848. IEEE, New York, 2012.

[48] H. F. Lopes and R. S. Tsay. Particle filters and Bayesian inference in financial econometrics. *J. Forecast.* **30** (1) (2011) 168–209. MR2758809

[49] J. Olsson and T. Ryden. Rao–Blackwellization of particle Markov chain Monte Carlo methods using forward filtering backward sampling. *IEEE Trans. Signal Process.* **59** (10) (2011) 4606–4619. MR2882969

[50] G. W. Peters, R. G. Hosack and K. R. Hayes. Ecological non-linear state space model selection via adaptive particle Markov chain Monte Carlo (AdPMCMC). Preprint, 2010. Available at arXiv:1005.2238.

[51] M. Pitt, R. dos Santos Silva, P. Giordani and R. Kohn. On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. *J. Econometrics* **171** (2) (2012) 134–151. MR2991856

[52] A. D. Rasmussen, O. Ratmann and K. Koelle. Inference for nonlinear epidemiological models using genealogies and time series. *PLoS Comput. Biol.* **7** (8) (2011) e1002136. MR2845064

[53] D. Revuz. *Markov Chains*. North-Holland, Amsterdam, 1975. MR0758799

[54] M. Rousset. On the control of an interacting particle approximation of Schrödinger ground states. *SIAM J. Math. Anal.* **38** (2006) 824–844. MR2262944

[55] C. Vergé, P. Del Moral, C. Dubarry and E. Moulines. On parallel implementation of sequential Monte Carlo methods: The island particle model. *Stat. Comput.* **25** (2015) 243–260. MR3306704

[56] J. A. Vrugt, C. J. ter Braak, C. G. Diks and G. Schoups. Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Adv. Water Resour.* **51** (2013) 457–478.

[57] N. Whiteley, C. Andrieu and A. Doucet. Efficient Bayesian inference for switching state-space models using discrete particle Markov chain Monte Carlo methods. Preprint, 2010. Available at arXiv:1011.2437.