The Annals of Probability
1995, Vol. 23, No. 4, 1772–1815

EXPLICIT STOCHASTIC INTEGRAL REPRESENTATIONS FOR HISTORICAL FUNCTIONALS

BY STEVEN N. EVANS[1] AND EDWIN A. PERKINS

University of California at Berkeley and University of British Columbia

It is known from previous work of the authors that any square-integrable functional of a superprocess may be represented as a constant plus a stochastic integral against the associated orthogonal martingale measure. Here we give, for a large class of such functionals, an explicit description of the integrand that is analogous to Clark's formula for the representation of certain Brownian functionals. As a consequence, we develop a partial analogue of the Wiener chaos expansion in the superprocess setting. Rather than work with superprocesses per se, our results are stated and proved in the richer and more natural context of the associated historical process.

1. Introduction. A fundamental result of Itô (1951) states that every square-integrable functional of a d-dimensional Brownian motion $B = (B^1, \dots, B^d)$ may be represented as a constant plus a sum of stochastic integrals against $B^1, \dots, B^d$. Dellacherie proved this fact as a direct consequence of Lévy's characterization of Brownian motion as the unique continuous martingale such that $B_0 = 0$ and $\langle B^i, B^j \rangle_t = \delta_{i,j} t$ [see Dellacherie and Meyer (1982), VIII.62, for an account of this proof].

Subsequent work of Jacod and Yor, culminating in Jacod (1979), greatly generalized Dellacherie's idea by showing that, roughly speaking, the existence of stochastic integral representations for functionals of a certain process and the well-posedness of a martingale problem for that process are equivalent phenomena.

In the case of Brownian motion, an even stronger result is true. Namely, any square-integrable functional has an orthogonal expansion, the *Wiener chaos* expansion, in terms of multiple stochastic integrals with deterministic integrands. Generalizing this result to other processes is a more delicate matter. For example, even the definition of multiple stochastic integrals can be difficult when the quadratic variation processes of the integrator martingale is not deterministic, and two multiple stochastic integrals of different orders will, in general, no longer be orthogonal.

Received June 1994; revised January 1995.

[1]Research supported in part by Presidential Young Investigator award and Alfred P. Sloan Fellowship.

[2]Research supported in part by an NSERC Operating Grant.

AMS 1991 *subject classifications*. 60H05, 60G57, 60J80.

Key words and phrases. Superprocess, historical process, martingale measure, stochastic integral, predictable representation, Clark's formula, Wiener chaos.

As Dellacherie's proof proceeds by showing that the nonexistence of a representation would contradict Lévy's result, it does not give "formulae" for the integrands appearing in the representation. A similar comment applies to Itô's approach. Such explicit representations of suitable functionals were first given by Clark (1971).

Haussmann (1978, 1979) significantly extended Clark's work by obtaining explicit representations for suitable functionals of finite-dimensional diffusions in terms of stochastic integrals against the driving Brownian motions. Davis (1980) indicated how the Clark–Hausmann formulae followed from potential theory considerations. Bismut (1981) provided another approach based on Girsanov's formula as part of his variational treatment of the Malliavin calculus. Ocone (1984) showed, conversely, that the Clark–Haussmann formulae may be derived fairly directly from Stroock's Wiener chaos approach to the Malliavin calculus.

As well as being intimately connected with Malliavin calculus, explicit stochastic integral representations and chaos expansions have become of central importance in fields such as control theory, filtering and mathematical finance.

In this paper we obtain explicit stochastic integral representations for suitable functionals of certain infinite-dimensional diffusions, the *Dawson–Watanabe superprocess*. We also obtain a partial analogue of the Wiener chaos expansion. (In fact, rather than consider superprocesses themselves, we work with the richer associated *historical processes*, for which the statements and proofs of our results are more natural. Although we do not do so because it would involve a significant amount of new notation, it is trivial to read off results for superprocesses as "projections" of our results.)

We will give a full account of historical processes in Section 2. To keep things simple for the sake of describing the intuitive content of our results and their proofs in this introduction, we will consider the more familiar superprocesses instead and skip over a number of details (e.g., the class of Markov processes on which we are performing the superprocess construction and what, exactly, we mean by the generator of such a process). A very useful reference for superprocesses (and much else) is Dawson (1993). One of the prime motivations for the current interest in superprocesses is their connection to certain nonlinear partial differential equations. Dynkin (1993) provides a thorough guide to the work in this area.

Suppose that Y is a Markov process with topological state space E. The superprocess over the process Y [more correctly, the $(Y, -\lambda^2/2)$ superprocess] is a continuous, Markov process S taking values in the space $M_F(E)$ of finite Borel measures on E equipped with the weak topology. There are two rather different (but equivalent) ways of defining this process. We will outline both approaches, because elements of both appear as ingredients in our explicit representation.

On the one hand, we can think of the superprocess as the solution to a certain martingale problem. Write $\mu(f)$ for $\int f\, d\mu$ when $\mu \in M_F(E)$ and f is a Borel function. Denote the generator of Y by A. The law of S is uniquely specified by requiring that, for each f in the domain of the generator, we have

$$S_t(f) = m(f) + \int_0^t S_s(Af)\, ds + Z_t^f,$$

where $m \in M_F(E)$ and Z^f is a continuous martingale such that $Z_0^f = 0$ and $\langle Z^f \rangle_t = \int_0^t S_s(f^2)\, ds$.

On the other hand, we can give a description of the law of S_t within the framework of the theory of Itô decompositions of infinitely divisible random measures. Here, when $S_0 = m \in M_F(E)$, the law of S_t is that of the finite random sum $\int_{M_F(E)} p\Pi(dp)$, where Π is a Poisson random measure on $M_F(E)$ with intensity $\int_E R_t(x, d\mu)m(dx)$. The finite measure $R_t(x, \cdot)$ is concentrated on $M_F(E) \setminus \{0\}$ and is characterized by

$$\int_{M_F(E)} [1 - \exp(-\mu(f))] R_t(x, d\mu) = v_t(x)$$

for nonnegative, bounded, Borel f, where v solves the integral equation

$$v_t(x) = P_t f(x) - \frac{1}{2}\int_0^t P_s(v_{t-s}^2)(x)\, ds,$$

with $(P_t)_{t \geq 0}$ being the semigroup of A. We note that $R_t(x, \cdot)$ has total mass $2/t \equiv r_t$, say, that is independent of x.

The intuitive content of the latter description is most clearly brought out by considering the construction of S as a limit of particle systems. Suppose that we lay down a number of particles in E according to a Poisson process with intensity Nm for some integer N. Each particle is equipped with a clock that rings after a period of time that is exponentially distributed with mean $1/N$. The clocks are independent. Each particle executes an independent copy of Y starting from its initial position until its clock rings. At that time the particle dies and gives birth to either zero or two particles, each possibility occurring with probability $\frac{1}{2}$ independently of the remainder of the population. The offspring then proceed to evolve as independent copies of their parent. That is, they too will eventually die and possibly reproduce. The total number of particles is thus a critical Galton–Watson branching process. If we construct a random measure S_t^N by associating mass $1/N$ with each particle alive at time t, then the process S^N is an $M_F(E)$-valued Markov process. If we let $N \to \infty$, then S^N converges in law to S.

The above Itô decomposition of S can then be interpreted as a decomposition of S into "clusters" descended from distinct "progenitors" present in the population at time 0. The configuration of progenitors is a Poisson process on E with intensity $r_t m$, the clusters descended from each progenitor are independent and the law of a cluster descended from a progenitor at location $x \in E$ is $R_t(x, \cdot)/r_t$. This heuristic picture may be made precise by using nonstandard analysis [cf. Dawson, Iscoe and Perkins (1989)] or by an associated enriched model that keeps track of genealogies such as the historical process of Dawson and Perkins (1991). In this introduction we will use, without comment (or inverted commas), particle process terminology, such as cluster, progenitor, descendant, subtree and so on, to give an indication of the heuristics behind our results, even though such terminology is not strictly applicable to S, but only to an enriched or approximating model.

Given that S is the solution of a well-posed martingale problem, it is natural to expect from the general results of Jacod and Yor mentioned above that there should be some kind of representation of functionals of S in terms of stochastic integrals involving the martingales Z^f. Such a result was obtained in Evans and Perkins (1994). We will recall this result in Theorem 4.7, but for the sake of the information we will give a brief outline.

There is an orthogonal martingale measure, Z, in the sense of Walsh (1986) such that

$$Z_t^f = \int_{]0,\,t]} \int_E f(x)\, dZ(s,x).$$

It is possible to define space–time stochastic integrals against Z for suitable random integrands, and for such integrands we have

$$\left\langle \int_{]0,\,\cdot\,]} \int_E \phi(s,x)\, dZ(s,x) \right\rangle_t = \int_{]0,\,t]} \int_E \phi(s,x)^2 S_s(dx)\, ds.$$

The representation result of Evans and Perkins (1994) states that if $F(S)$ is a square-integrable functional of S, then we have

$$F(S) = \mathbb{P}[F(S)] + \int_{]0,\infty[} \int_E \phi^F(s,x)\, dZ(s,x)$$

for a suitable integrand ϕ^F.

The aim of this paper is thus to find the explicit form of the integrand for a fairly general class of functionals.

Specific examples of such explicit representations already exist in the literature. For example, it is fairly easy to show directly from the martingale problem that

$$S_t(f) = m(P_t f) + \int_{]0,\infty[} \int_E P_{t-s} f(x) \mathbf{1}(s<t)\, dZ(s,x). \tag{1.1}$$

We note that if this observation is coupled with the above remark about the quadratic variation of stochastic integrals against Z, then it is straightforward to apply Itô's lemma and inductively obtain formulae for the moments of the form $\mathbb{P}[S_{t_1}(f_1) \cdots S_{t_k}(f_k)]$. Such moment formulae were obtained analytically in Dynkin (1988). They have found a host of applications [see, e.g., Adler and Lewin (1991), Evans (1990), Evans and Perkins (1991), Krone (1993), Rosen (1992) and Sugitani (1989)].

It is not hard to check that

$$P_{t-s} f(x) = \int_{M_F(E)} \mu(f) R_{t-s}(x, d\mu), \qquad s<t,$$

and so the integrand appearing in the representation (1.1) can be written as

$$\mathbb{P}\left[\int_{M_F(E)} ((S_t+\mu)(f) - S_t(f))R_{t-s}(x, d\mu)\middle| S_u, 0 \le u \le s\right]\mathbf{1}(s < t).$$

A special case of the representation we will establish is that for any suitable functional of the form $F(S) = G(S_t)$ the integrand in the representation is, roughly speaking, given by

$$\begin{aligned}\phi^F(s, x) = \mathbb{P}\Bigg[\int_{M_F(E)} [(G(S_t+\mu) - G(S_t))\\ \times R_{t-s}(x, d\mu)]\Bigg| S_u, 0 \le u \le s\Bigg]\mathbf{1}(s < t).\end{aligned} \tag{1.2}$$

(We are ignoring, for the moment, some details such as how to define the conditional expectation for all s so that we get a legitimate stochastic integrand.)

In order to understand why (1.2) should be the correct formula, it is instructive to give a caricature of the proof in which we gloss over a host of details. The basic idea is the same as that employed by Bismut (1981) for Brownian motion. An elementary exposition of this technique is given in Rogers and Williams [(1987), Chapter IV, Section 41].

Given that we know ϕ^F exists from the general representation result of Evans and Perkins (1994), our problem is to identify it. We observe that for a suitable integrand β we have

$$\begin{aligned}&\mathbb{P}\left[G(S_t)\int_{]0,t]}\int_E \beta(s, x)\, dZ(s, x)\right]\\ &\quad= \mathbb{P}\left[\left(\mathbb{P}[G(S_t)] + \int_{]0,\infty[}\int_E \phi^F(s, x)\, dZ(s, x)\right)\int_{]0,t]}\int_E \beta(s, x)\, dZ(s, x)\right]\\ &\quad= \mathbb{P}\left[\int_{]0,t]}\int_E \phi^F(s, x)\beta(s, x)S_s(dx)\, ds\right].\end{aligned}$$

We thus want to find a way of computing the leftmost expectation that gives an informative expression in the form of the rightmost expectation.

The next step is to note that

$$\begin{aligned}&\mathbb{P}\left[G(S_t)\int_{]0,t]}\int_E \beta(s, x)\, dZ(s, x)\right]\\ &\qquad= -\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\left\{\mathbb{P}\left[G(S_t)\mathscr{E}_t^{\varepsilon\beta}\right] - \mathbb{P}[G(S_t)]\right\},\end{aligned} \tag{1.3}$$

where

$$\mathscr{E}_t^{\varepsilon\beta} = \exp\left(-\int_{]0,t]}\int_E \varepsilon\beta(s, x)\, dZ(s, x) - \tfrac{1}{2}\int_{]0,t]}\int_E \varepsilon^2\beta(s, x)^2 S_s(dx)\, ds\right).$$

We have $\mathbb{P}[\mathscr{E}_t^{\varepsilon\beta}] = 1$, and so $\mathbb{P}[G(S_t)\mathscr{E}_t^{\varepsilon\beta}]$ is the expectation of $G(S_t)$ under a new probability measure that we will call $\mathbb{P}^{\varepsilon\beta}$. In order to find a useful expression for the limit in (1.3), we want to construct a process S' on our original probability space such that S' has the same law as S does under $\mathbb{P}^{\varepsilon\beta}$. That is, we want to couple S and S'. We can then write $\mathbb{P}[G(S_t)\mathscr{E}_t^{\varepsilon\beta}] - \mathbb{P}[G(S_t)] = \mathbb{P}[G(S_t') - G(S_t)]$ and investigate the limit in (1.3) using arguments that compare the sample paths of S' and S.

The analogous problem for Brownian motion is relatively simple. If we perturb the law of Brownian motion by introducing a stochastic exponential as a Radon–Nikodym factor, then Girsanov's theorem says that a process with the resulting law can be realized on the same probability space as a Brownian motion with a random drift.

In order to see what is involved in our case, first consider the case when $\beta(s,x) = b(x)$, where b is a nonnegative, deterministic function. Some straightforward stochastic calculus shows that the law of $\{S_s\colon 0 \le s \le t\}$ under $\mathbb{P}^{\varepsilon\beta}$ satisfies a martingale problem similar to that satisfied by the law of $\{S_s\colon 0 \le s \le t\}$ under $\mathbb{P}$ with the difference that the operator A is replaced by the operator $A - \varepsilon b$. Thus the law of $\{S_s\colon 0 \le s \le t\}$ under $\mathbb{P}^{\varepsilon\beta}$ is that of a process constructed in the same manner as S, with the difference that Y is replaced by Y killed according to the continuous additive functional $\varepsilon\int_{]0,\cdot]} b(Y_s)\,ds$.

The intuitive interpretation when β is allowed to be random and time varying, but still nonnegative, is obvious. We can still think that we are looking at what is essentially a cloud of branching particles moving according to the dynamics of Y, except that now individual particles are killed at a rate that may depend on the position of the particle and the whole history of the population up to the present.

In general, coupling a Markov process and a killed version of the process involves enlarging the probability space to include some extra randomness. For example, suppose that we wish to kill Y according to the continuous additive functional $\varepsilon\int_{]0,\cdot]} b(Y_s)\,ds$, where $0 \le \varepsilon b(x) \le 1$. By enlarging the probability space if necessary, we can construct a Poisson point process N on $[0,\infty[\times[0,1]$ with intensity $dt \otimes dx$. If we kill Y at the first time point, s, that N has a mark in $[0, \varepsilon b(Y_s)]$, then the resulting process is the killed process we want.

A natural approach to constructing S and S' is then as follows. First build the superprocess over the Markov process (Y, N). The process S will simply be the "E marginal" of this superprocess. To obtain S', first erase particles in a manner suggested by the above, and then proceed to take the "E marginal." We show in Section 5 that, in essence, this approach works to construct the desired coupling. There are some technical difficulties engendered by the fact that in order to implement this idea as it stands we would need to be able to track back through time the trajectory of an individual particle and its ancestors. Such trajectories are, in general, not well defined [see, however, Barlow and Perkins (1994)], so we need to introduce the extra machinery of the associated historical process to make sense of this intuition.

We know that particles alive at time $s < t$ that are progenitors of clusters present in S_t appear as a Poisson process with rate $r_{t-s}S_s$ conditional on S_s. Moreover, such a particle in position x at time s is being killed off at rate $\varepsilon\beta(s, x)$. When a particle is killed off, its entire cluster of descendants at time t is removed from S_t in order to produce S'_t. Conditional on the killing taking place at time and place (s, x), the cluster that is removed will have law $R_{t-s}(x, \cdot)/r_{t-s}$ [recall that r_{t-s} is the total mass of $R_{t-s}(x, \cdot)$]. Last, we have the fundamental Palm measure fact about Poisson random measures: if we condition any Poisson random measure Π, say, with diffuse intensity measure to have an atom at some point z, then the resulting conditional law of Π is the same as the unconditional law of $\Pi + \delta_z$, where δ_z is the unit point mass at z. In particular, the conditional law of $H(\Pi) - H(\Pi - \delta_z)$ is the same as the unconditional law of $H(\Pi + \delta_z) - H(\Pi)$ for any Borel function H defined on the space of measures in which Π takes its values.

We hope it is now reasonable at a heuristic level that the limit as $\varepsilon \downarrow 0$ of the right-hand side of (1.3) should be

$$\mathbb{P}\biggl[\int_{]0,\, t]} \int_E \int_{M_F(E)} \mathbb{P}\bigl[(G(S_t + \mu) - G(S_t))\big| S_u, 0 \le u \le s\bigr] \\ \times R_{t-s}(x, d\mu)\, \beta(s, x)\, S_s(dx)\, ds\biggr],$$

and leads directly to the identification of ϕ^F given in (1.2).

The foregoing caricature also indicates what the form of ϕ^F should be when the functional $F(S)$ is allowed to depend on the whole path and not just on S_t. It seems reasonable from the particle picture that the whole path of S can be thought of as a superposition of a Poisson process of clusters in path space. There will be infinitely many clusters, but for each $t > 0$ only finitely many clusters will live longer than t. This cluster decomposition does indeed hold, and, as we would expect, the intensity of the Poisson process of these "path" clusters plays a role in the representation of functionals of the whole path similar to that played by the intensity of the Poisson process of "one-dimensional marginal" clusters in the representation of functionals of S_t.

The plan of the remainder of the paper is as follows.

In Section 2 we recall the definition of a historical process, collect some facts from the literature and state our main results. In Section 3 we detail the construction of a "marked" process that we sketched above.

In our heuristics, the key idea was that if we kill particles using the marks, then we arrive at a process that has the same law as the law of the original process perturbed by a Girsanov-type Radon–Nikodym factor. Our actual argument will, in fact, depend on a slight variant of this idea; namely, that if we kill particles using the marks and perturb the law of the killed process by a different Girsanov-type factor, then we return to the law of our original process. This fact is proved in Section 5. In order to accomplish the proof, we develop in Section 4 a "stochastic calculus along branches" that is similar to that described in Perkins (1992, 1995), but handles certain martingales with jumps.

The proof of our explicit stochastic integral representation result (Theorem 2.5) is given in Section 7. It is an immediate consequence of a somewhat more fundamental "stochastic integration-by-parts" formula (Theorem 2.4) that we establish in Section 6. In Section 8 we prove a partial analogue of the Wiener chaos expansion (Theorem 2.7).

2. Notation and statement of results. It will be convenient to recycle some of the notation used in the Introduction. As certain symbols are reused to denote slightly different objects in the rest of the paper, the reader should forget about the usage in the Introduction.

We begin by recalling the definition of the historical process associated with a Hunt process as it is presented in Mueller and Perkins (1992).

Let E be a Polish space with Borel σ-field $\mathscr{E}$.

Let D denote $D(\mathbb{R}_+, E)$, the space of cadlag functions from $\mathbb{R}_+$ to E equipped with the Skorokhod topology induced by the metric on E. Write $\mathscr{D}$ for the Borel σ-field of D and $(\mathscr{D}_t)$ for the canonical filtration. Given $y, y' \in D$ and $t \geq 0$, define $y^t \in D$ by $y^t(s) = y(s \wedge t)$, define $y^{t-} \in D$ by

$$y^{t-}(s) = \begin{cases} y(s), & \text{if } s < t, \\ y(t-), & \text{if } s \geq t, \end{cases}$$

and define $(y/t/y') \in D$ by

$$(y/t/y')(s) = \begin{cases} y(s), & \text{if } s < t, \\ y'(s-t), & \text{if } s \geq t. \end{cases}$$

Write D^t (resp., D^{t-}) for the image of D under the map $y \mapsto y^t$ (resp., $y \mapsto y^{t-}$). For $t \geq 0$ define $Y_t\colon D \to E$ by $Y_t(y) = y(t)$.

Fix a conservative Hunt process on E, and let P^y be the probability measure on $(D, \mathscr{D})$ that is the law of this process when the process is in state $y \in E$ at time 0. (For convenience, we will, in different contexts, use the dummy variable y to denote both points in D and E.) In other words, $Y = (Y_t, P^y)$ is the canonical realization of the Hunt process.

Let $M_F(D)$ denote the space of finite measures on $(D, \mathscr{D})$ equipped with the topology of weak convergence. Given $t \geq 0$, put $M_F(D)^t = \{m \in M_F(D)\colon y = y^t \text{ for } m\text{-a.e. } y\}$.

Fix $\tau \geq 0$ and $\nu \in M_F(D)^\tau$. Define a finite measure $P^{\tau,\nu}$ on $(D, \mathscr{D})$ by

$$P^{\tau,\nu}(A) = \int_D P^{y(\tau)}\{(y/\tau/Y) \in A\}\,\nu(dy).$$

Let $\overline{\mathscr{D}}_t^\nu$ denote the σ-field generated by $\mathscr{D}_{t+}$ and the $P^{\tau,\nu}$-null subsets in D.

Define $F^{\tau,\nu}$ to be the set of functions $\phi \in b(\mathscr{B}([\tau,\infty[) \times \mathscr{D})$ that satisfy the condition $\phi(t, y) = \phi(t, y^t)$, for all $t \geq \tau$, and are such that $t \mapsto \phi(t, Y)$ is $P^{\tau,\nu}$-a.s. right continuous for $t \geq \tau$. If $\phi \in F^{\tau,\nu}$ we will sometimes write ϕ_t for the function $\phi(t,\cdot)$, $t \geq \tau$. Write $A^{\tau,\nu}$ for the subset of $F^{\tau,\nu} \times F^{\tau,\nu}$ consisting of pairs (ϕ, ψ) such that

$$\phi_t(Y) - \phi_\tau(Y) - \int_{]\tau,t]} \psi_s(Y)\,ds, \qquad t \geq \tau,$$

is a $(\bar{\mathscr{D}}_t^\nu)_{t \ge \tau}$ martingale under $P^{\tau,\nu}$. Note that if $(\phi, \psi) \in A^{\tau,\nu}$, then $\phi_t(Y)$ is cadlag $P^{\tau,\nu}$-a.s.

Suppose that $(\Omega, \mathscr{H}, (\mathscr{H}_t)_{t \ge \tau}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions. A $(Y, -\lambda^2/2)$ historical process starting at (τ, ν) is a continuous, $M_F(D)$-valued, $(\mathscr{H}_t)_{t \ge \tau}$-adapted process $(H_t)_{t \ge \tau}$ such that $\mathbb{P}$-a.s. $H_t \in M_F(D)^t$ for all $t \ge \tau$, and for all $(\phi, \psi) \in A^{\tau,\nu}$,

$$M_t^\phi = H_t(\phi_t) - \nu(\phi_\tau) - \int_{]\tau, t]} H_s(\psi_s)\, ds, \qquad t \ge \tau,$$

is a continuous $(\mathscr{H}_t)_{t \ge \tau}$ martingale for which $M_\tau^\phi = 0$ and

$$\langle M^\phi \rangle_t = \int_{]\tau, t]} \int_D \phi(s, y)^2 H_s(dy)\, ds.$$

The law of H is unique. The existence and uniqueness proof in Mueller and Perkins (1992) relies on a general result from Fitzsimmons (1988). Examining this line of reasoning, it is clear that in order to check that H is a historical process, it is possible to weaken the requirement that the processes M^ϕ are martingales to a requirement that these processes are just local martingales.

We will now record some miscellaneous facts about the process H chosen as above. Unless otherwise noted, these facts may found in either Dawson and Perkins (1991) or Mueller and Perkins (1992).

The process H is an inhomogeneous Hunt process. If $f_1, \dots, f_k \in b\mathscr{D}$, then the moment $\mathbb{P}[H_{t_1}(f_1) \cdots H_{t_k}(f_k)]$ exists and is uniformly bounded as $(t_1, \dots, t_k)$ ranges over a compact subset of $[\tau, \infty[^k$. Explicit formulae for these moments are given in Dynkin (1988). (As mentioned in the Introduction, such formulae may be derived inductively from the martingale problem using stochastic calculus.) In particular,

$$\mathbb{P}[H_t(f)] = P^{\tau,\nu}[f(Y^t)]. \tag{2.1}$$

Let $L^2(H, \mathbb{P})$ denote the class of functions g: $]\tau, \infty[\times D \times \Omega \to \mathbb{R}$ that are predictable with respect to the filtration $(\mathscr{D} \times \mathscr{H}_t)_{t \ge \tau}$ and satisfy

$$\mathbb{P}\left[\int_{]\tau, t]} \int_D g(s, y)^2 H_s(dy)\, ds\right] < \infty, \qquad t \ge \tau.$$

There is an orthogonal martingale measure M in the sense of Walsh (1986) such that

$$\int_{]\tau, t]} \int_D g(s, y)\, dM(s, y)$$

is defined and is a square-integrable, continuous $(\mathscr{H}_t)_{t \ge \tau}$ martingale under $\mathbb{P}$ for each $g \in L^2(H, \mathbb{P})$ with

$$\left\langle \int_{]\tau, \cdot]} \int_D g(s, y)\, dM(s, y) \right\rangle_t = \int_{]\tau, t]} \int_D g(s, y)^2 H_s(y)\, ds.$$

If $(M_t^{\phi})_{t\ge\tau}$ is one of the martingales appearing in the martingale problem that defines H, then

$$M_t^{\phi} = \int_{]\tau,\,t]}\int_D \phi(s,y)\,dM(s,y).$$

For $t > \tau$, the law of H_t under $\mathbb{P}$ is that of the finite random sum $\int_{M_F(D)} p\Pi(dp)$, where Π is a Poisson random measure on $M_F(D)$ with finite intensity of the form $\int_{D^\tau} R_{\tau,t;y}\nu(dy)$. The *canonical measure* $R_{\tau,t;y}$ is concentrated on $M_F(D)\setminus\{0\}$ and has total mass $r_{\tau,t} \equiv 2/(t-\tau)$, for all $y \in D^\tau$.

For $r \ge 0$ and $\zeta \in M_F(D)^r$ let $\mathbb{P}^{r;\zeta}$ denote the law of the $(Y, -\lambda^2/2)$ historical process started at (r,ζ). Thus $\mathbb{P}^{r;\zeta}$ is a probability measure on $C([r,\infty[, M_F(D))$, and, in particular, $\mathbb{P}^{\tau;\nu}$ is just the law of H under $\mathbb{P}$. We can think of $C([s,\infty[, M_F(D))$ as a semigroup equipped with the obvious pointwise addition operation inherited from $M_F(D)$. The probability measure $\mathbb{P}^{r;\zeta}$ is infinitely divisible, and we in fact have that if $\zeta = \zeta' + \zeta''$, then $\mathbb{P}^{r;\zeta}$ is the convolution of $\mathbb{P}^{r;\zeta'}$ and $\mathbb{P}^{r;\zeta''}$.

We now need to develop some observations that are not in our two general historical process references, but which follow fairly simply from ideas that are in Dawson and Perkins (1991).

For $\tau \le s < t$ and $y \in D^s$ put

$$\mathbb{Q}^{s,t;y} = \int \mathbb{P}^{t;\zeta} R_{s,t;y}(d\zeta).$$

From Proposition 3.3 and Theorem 2.2.3 of Dawson and Perkins (1991), it is not difficult to check that if $\tau \le s < t < u$ and $P_{t,u}$ is the transition kernel for H from time t to time u, then the two measures $R_{s,u;y}(\cdot)$ and $\int P_{t,u}(\xi,\cdot) R_{s,t;y}(d\xi)$ agree on $M_F(D)^u \setminus \{0\}$. Thus, if we let $\pi_{t,u}$ be the mapping from $C([t,\infty[, M_F(D))$ to $C([u,\infty[, M_F(D))$ that maps the function h: $[t,\infty[\to M_F(D)$ to its restriction to $[u,\infty[$, then we have that

$$\mathbb{Q}^{s,u;y} = \left(\mathbb{Q}^{s,t;y} \circ \pi_{t,u}^{-1}\right)_{|\{h\in C([u,\infty[,\, M_F(D)):\, h_u \ne 0\}}.$$

Equivalently,

$$\mathbb{Q}^{s,u;y} = \left(\mathbb{Q}^{s,t;y}_{|\{h\in C([t,\infty[,\, M_F(D)):\, h_v \ne 0,\, t\le v\le u\}} \circ \pi_{t,u}^{-1}\right).$$

Now write $\pi_{s+,t}$ for the mapping from $C(]s,\infty[, M_F(D))$ to $C([t,\infty[, M_F(D))$ that maps the function h: $]s,\infty[\to M_F(D)$ to its restriction to $[t,\infty[$. We conclude from the above that we may define a unique σ-finite measure $\mathbb{Q}^{s+;y}$ on $C(]s,\infty[, M_F(D))$ that puts no mass on the zero path, is Markovian with semigroup $\{P_{v,w}\}$ and satisfies

$$\mathbb{Q}^{s+;y}_{|\{h\in C(]s,\infty[,\, M_F(D)):\, h_v\ne 0,\, s<v\le u\}} \circ \pi_{s+,t}^{-1} = \mathbb{Q}^{s,t;y}_{|\{h\in C([t,\infty[,\, M_F(D)):\, h_v\ne 0,\, t\le v\le u\}},$$

for $s < t < u$.

It follows from Le Gall (1993) that under $\mathbb{Q}^{s+;y}$ the "law" of the total mass process is just the "law" of the local time process of an excursion from 0 of a rescaled reflecting Brownian motion under the Itô excursion law. Because the local time at level w converges to 0 (= the local time at level 0) as $w \downarrow 0$ a.s. under the Itô excursion law, we have

$$\mathbb{Q}^{s+;y}\left\{h \in C(]s,\infty[, M_F(D)) : \lim_{t \downarrow s} h_t(D) \neq 0\right\} = 0.$$

We may therefore define a σ-finite measure $\mathbb{Q}^{s;y}$ on $C([\tau,\infty[, M_F(D))$ such that

$$\mathbb{Q}^{s;y}\{h: \exists\, \tau \le v \le s,\ h_v \neq 0\} = 0$$

and

$$\mathbb{Q}^{s;y} \circ \pi_{s+}^{-1} = \mathbb{Q}^{s+;y},$$

where π_{s+} is the mapping from $C([\tau,\infty[, M_F(D))$ to $C(]s,\infty[, M_F(D))$ that maps the function $h: [\tau,\infty[\to M_F(D)$ onto its restriction to $]s,\infty[$.

We note that the measures $\mathbb{Q}^{s;y}$ play a role in a Poisson cluster representation of the whole path $\{H_u\}_{u \ge s}$ under $\mathbb{P}^{s;\zeta}$ similar to that played by the measures $R_{s,t;y}$ in the Poisson cluster representation of a single random measure H_u. To be more precise, we can think of $C([s,\infty[, M_F(D))$ as a semigroup equipped with the obvious pointwise addition operation inherited from $M_F(D)$. Then $\mathbb{P}^{s;\zeta}$ is the law of the superposition of clusters thrown down on $C([s,\infty[, M_F(D))$ according to the restriction to $C([s,\infty[, M_F(D))$ of a Poisson process on $C([\tau,\infty[, M_F(D))$ with intensity $\int \mathbb{Q}^{s;y}\zeta(dy)$. [Remember that we "padded out" $\mathbb{Q}^{s;y}$ so that we could think of it as a measure on $C([\tau,\infty[, M_F(D))$.] This path space type of Lévy–Khintchine representation for $P^{s;\zeta}$ is discussed at length in El-Karoui and Roelly (1991).

If $T \ge \tau$ is a bounded $(\mathscr{H}_t)_{t \ge \tau}$ stopping time, we define $\overline{\mathbb{P}}_T$, the normalized Campbell measure associated with H_T, to be the probability measure on $(D \times \Omega, \mathscr{D} \times \mathscr{H})$ given by

$$\overline{\mathbb{P}}_T(A \times B) = \mathbb{P}[H_T(A)\mathbf{1}_B]/\nu(D).$$

In Lemmas 2.2 and 2.3, we define and study a type of predictable projection under normalized Campbell measure that is necessary for defining the integrands that will appear in our representation result. First, however, we record the following observation about predictable σ-fields for product filtrations that has no doubt appeared in the literature, but we have been unable to find a reference.

LEMMA 2.1. *The $(\mathscr{D}_t \times \mathscr{H}_t)_{t \ge \tau}$-predictable σ-field on $]\tau,\infty[\times D \times \Omega$ is generated by functions of the form $(t, y, \omega) \mapsto \beta(t, y)\delta(t, \omega)$, where β is $(\mathscr{D}_t)_{t \ge \tau}$-predictable and δ is $(\mathscr{H}_t)_{t \ge \tau}$-predictable.*

PROOF. It is clear that if β is $(\mathscr{D}_t)_{t \ge \tau}$-predictable and δ is $(\mathscr{H}_t)_{t \ge \tau}$-predictable, then $(t, y, \omega) \mapsto \beta(t, y)\delta(t, \omega)$ is $(\mathscr{D}_t \times \mathscr{H}_t)_{t \ge \tau}$-predictable.

Conversely, it follows from Theorem IV.67(c) of Dellacherie and Meyer (1978) (modified slightly to take into account that we are working with the open index set $]\tau,\infty[$) that the $(\mathscr{D}_t \times \mathscr{H}_t)_{t \ge \tau}$-predictable σ-field is generated by sets of the form $]s,t] \times A$, where $\tau < s < t < \infty$ and $A \in \mathscr{D}_r \times \mathscr{H}_r$, for some

$\tau < r < s$. Hence, the $(\mathscr{D}_t \times \mathscr{H}_t)_{t \geq \tau}$-predictable σ-field is also generated by sets of the form $]s, t] \times B \times C$, where $\tau < s < t < \infty$, $B \in \mathscr{D}_r$ and $C \in \mathscr{H}_r$, for some $\tau < r < s$. The indicator function of such a set may be written as $(\mathbf{1}_{]s,t]}\mathbf{1}_B)(\mathbf{1}_{]s,t]}\mathbf{1}_C)$. Whereas $\mathbf{1}_{]s,t]}\mathbf{1}_B$ is $(\mathscr{D}_t)_{t \geq \tau}$-predictable and $\mathbf{1}_{]s,t]}\mathbf{1}_C$ is $(\mathscr{H}_t)_{t \geq \tau}$-predictable, the result follows. □

LEMMA 2.2. *Suppose that* α: $]\tau, \infty[\times D \times \Omega \to \mathbb{R}$ *is bounded below and measurable with respect to the product of the* $(\mathscr{D}_t)_{t \geq \tau}$*-predictable* σ*-field and the* σ*-field* $\mathscr{H}$. *Then there exists* $\overline{\alpha}$: $]\tau, \infty[\times D \times \Omega \to \mathbb{R}$ *that is* $(\mathscr{D}_t \times \mathscr{H}_t)_{t \geq \tau}$*-predictable and satisfies*

$$\overline{\mathbb{P}}_T[\alpha(T) | (\mathscr{D} \times \mathscr{H})_T] = \overline{\alpha}(T),$$

$\overline{\mathbb{P}}_T$*-a.s., for all bounded* $(\mathscr{H}_t)_{t \geq \tau}$*-predictable stopping times* $T > \tau$. *In particular,*

$$\mathbb{P}\left[\int_D \alpha(T, y) H_T(dy)\right] = \mathbb{P}\left[\int_D \overline{\alpha}(T, y) H_T(dy)\right].$$

Moreover, if $\hat{\alpha}$ *is any other function with the same properties as* $\overline{\alpha}$, *then* $\mathbb{P}$*-a.s., for all* $t > \tau$, $\hat{\alpha}(t, y) = \overline{\alpha}(t, y)$ *for* H_t*-a.e.* $y \in D$.

PROOF. In order to establish the existence of $\overline{\alpha}$, it suffices by a monotone class argument (just as in the construction of the ordinary predictable projection) to consider the case $\alpha(t, y, \omega) = \beta(t, y)\gamma(\omega)$, where β is bounded and $(\mathscr{D}_t)_{t \geq \tau}$-predictable and $\gamma \in b\mathscr{H}$. We claim that in this case we can take $\overline{\alpha}(t, y, \omega) = \beta(t, y)\overline{\gamma}(t, \omega)$, where $\overline{\gamma}$ is the $(\mathscr{H}_t)_{t \geq \tau}$-predictable projection of the process that takes the value γ at all times.

It follows from Lemma 2.1 that $(t, y, \omega) \mapsto \beta(t, y)\overline{\gamma}(t, \omega)$ is $(\mathscr{D}_t \times \mathscr{H}_t)_{t \geq \tau}$-predictable. It therefore remains to check that if $T > \tau$ is a bounded $(\mathscr{H}_t)_{t \geq \tau}$-predictable stopping time, then

$$\mathbb{P}\left[\int \beta(T, y)\gamma\delta(y) H_T(dy)\right] = \mathbb{P}\left[\int \beta(T, y)\overline{\gamma}(T)\delta(y) H_T(dy)\right]$$

whenever $(y, \omega) \mapsto \delta(y, \omega)$ belongs to $b(\mathscr{D} \times \mathscr{H})_T$. Lemma 3.4 of Perkins (1992) shows that $(\mathscr{D} \times \mathscr{H})_T$ is generated by functions of the form $(y, \omega) \mapsto \phi(y^{T(\omega)}) \times \psi(\omega)$, where $\phi \in b\mathscr{D}$ and $\psi \in b\mathscr{H}_T$ [the result in Perkins (1992) is for the special case in which D is replaced by a space of continuous paths, but the facts from Dellacherie and Meyer (1978) used in the proof hold in the greater generality we require]. It thus suffices by a monotone class argument to show that

$$\mathbb{P}\left[\int \beta(T, y)\phi(y^T) H_T(dy)\gamma\psi\right] = \mathbb{P}\left[\int \beta(T, y)\phi(y^T) H_T(dy)\overline{\gamma}(T)\psi\right].$$

This follows from the properties of the predictable projection since, by Proposition 2.2.b of Mueller and Perkins (1992), $(t, \omega) \mapsto \int\beta(t, y)\phi(y^t)H_t(dy)(\omega)$ is $(\mathscr{H}_t)_{t \geq \tau}$-predictable.

Turning to the uniqueness claim, let us first remark that if ρ: $]\tau,\infty[\times D\times\Omega\to\mathbb{R}$ is $(\mathscr{D}_t\times\mathscr{H}_t)_{t\geq\tau}$-predictable, then $(t,\omega)\mapsto\int\rho(t,y,\omega)H_t(dy)(\omega)$ is $(\mathscr{H}_t)_{t\geq\tau}$-predictable. This follows from Lemma 2.1, Proposition 2.2.b of Mueller and Perkins (1992) and a monotone class argument. Thus the set $\{(s,\omega)$: $s>\tau$, $H_s(\{y$: $\hat{\alpha}(s,y)>\overline{\alpha}(s,y)\})>0\}$ is $(\mathscr{H}_t)_{t\geq\tau}$-predictable. If this set is not evanescent, then it follows from the predictable section theorem that there exists a bounded $(\mathscr{H}_t)_{t\geq\tau}$-predictable stopping time T such that $\mathbb{P}[H_T(\{y$: $\hat{\alpha}(T,y)>\overline{\alpha}(T,y)\})>0]>0$, that is, $\overline{\mathbb{P}}_T\{\hat{\alpha}(T,\cdot)>\overline{\alpha}(T,\cdot)\}>0$. However, this is impossible, since $\mathbf{1}(\hat{\alpha}(T,\cdot)>\overline{\alpha}(T,\cdot))\in b(\mathscr{D}\times\mathscr{H})_T$ and so by assumption we have

$$\begin{aligned}\overline{\mathbb{P}}_T[\mathbf{1}(\hat{\alpha}(T,\cdot)>\overline{\alpha}(T,\cdot))\hat{\alpha}(T,\cdot)] &= \overline{\mathbb{P}}_T[\mathbf{1}(\hat{\alpha}(T,\cdot)>\overline{\alpha}(T,\cdot))\alpha(T,\cdot)]\\ &= \overline{\mathbb{P}}_T[\mathbf{1}(\hat{\alpha}(T,\cdot)>\overline{\alpha}(T,\cdot))\overline{\alpha}(T,\cdot)].\end{aligned}$$

Similarly, the set $\{(s,\omega)$: $s>\tau$, $H_s(\{y$: $\hat{\alpha}(s,y)<\overline{\alpha}(s,y)\})>0\}$ is evanescent and the claimed uniqueness holds. □

REMARKS. (i) An argument similar to Lemma 2.1 shows that the product of the $(\mathscr{D}_t)_{t\geq\tau}$-predictable σ-field and the σ-field $\mathscr{H}$ coincides with the $(\mathscr{D}_t\times\mathscr{H})_{t\geq\tau}$-predictable σ-field.

(ii) When dealing with the projection operation of Lemma 2.2, we will interpret $\overline{\alpha}=\overline{\beta}$ (resp., $\overline{\alpha}\geq\overline{\beta}$ and $\overline{\alpha}_n\to\overline{\alpha}$) to mean that $\overline{\alpha}(T(\omega),y)=\overline{\beta}(T(\omega),y)$ [resp., $\overline{\alpha}(T(\omega),y)\geq\overline{\beta}(T(\omega),y)$ and $\overline{\alpha}_n(T(\omega),y)\to\alpha(T(\omega),y)$] for $\mathbb{P}_T$-a.e. $(y,\omega)\in D\times\Omega$ for all bounded, $(\mathscr{H}_t)_{t\geq\tau}$-predictable stopping times T. As the proof of Lemma 2.2 shows, this definition of $\overline{\alpha}=\overline{\beta}$ is equivalent to requiring that $\mathbb{P}$-a.s., for all $t>\tau$ we have $\overline{\alpha}(t,y)=\overline{\beta}(t,y)$ for H_t-a.e. $y\in D$. Similar comments hold for $\overline{\alpha}\geq\overline{\beta}$ and $\overline{\alpha}_n\to\overline{\alpha}$. The following properties of the projection are immediate from the definition and the properties of conditional expectation.

LEMMA 2.3. (a) *If β is $(\mathscr{D}_t\times\mathscr{H}_t)_{t\geq\tau}$-predictable, then* $\overline{\alpha\beta}=\overline{\alpha}\beta$.

(b) *If $\alpha(t,y,\omega)\geq\beta(t,y,\omega)$ [resp., $\alpha(t,y,\omega)\leq\beta(t,y,\omega)$] for all (t,y,ω), then $\overline{\alpha}\geq\overline{\beta}$ (resp., $\overline{\alpha}\leq\overline{\beta}$).*

(c) *If $\alpha_n\uparrow\alpha$, then $\overline{\alpha}_n\uparrow\overline{\alpha}$.*

(d) *If $\alpha_n\to\alpha$ and $|\alpha|\leq c$ for some constant c, then $\overline{\alpha}_n\to\overline{\alpha}$.*

DEFINITION. Suppose that F: $C([\tau,\infty[,M_F(D))\to\mathbb{R}$ is a Borel function such that

$$J_{s;y}F(h)=\int(F(h+h^*)-F(h))\mathbb{Q}^{s;y^{s-}}(dh^*)$$

is defined and bounded below for all $s>\tau$, $y\in D$ and $h\in C([\tau,\infty[,M_F(D))$. Then $(s,y,\omega)\mapsto J_{s;y}F(H(\omega))$ is bounded below and measurable with respect to the product of the predictable σ-field associated with the filtration $(\mathscr{D}_t)_{t\geq\tau}$ and the σ-field $\mathscr{H}$. Let $\mathscr{J}_{s;y}F$ denote the corresponding $(\mathscr{D}_t\times\mathscr{H}_t)_{t\geq\tau}$-predictable process described in Lemma 2.2.

With these ideas in hand we can now state what might be loosely called a "stochastic integration-by-parts" formula and the explicit stochastic integral representation result that follows almost immediately from it.

THEOREM 2.4. *Suppose that* F: $C([\tau,\infty[, M_F(D)) \to \mathbb{R}$ *is continuous and such that for some compactly supported, finite measure* m *on* $[\tau,\infty[$ *we have* $|F(h+h^*) - F(h)| \le \int h_t^*(D)m(dt)$ *for all* $h, h^* \in C([\tau,\infty[, M_F(D))$. *If* β *is a bounded* $(\mathscr{D}_t \times \mathscr{H}_t)_{t\ge\tau}$*-predictable function, then, for all* $\theta > \tau$,

$$\mathbb{P}\left[F(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right]$$
$$= \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D (\mathscr{I}_{s;y}F)\beta(s,y)H_s(dy)\,ds\right].$$

THEOREM 2.5. *Suppose that* F: $C([\tau,\infty[, M_F(D)) \to \mathbb{R}$ *is continuous and such that for some compactly supported, finite measure* m *on* $[\tau,\infty[$ *we have* $|F(h+h^*) - F(h)| \le \int h_t^*(D)m(dt)$ *for all* $h, h^* \in C([\tau,\infty[, M_F(D))$. *Then*

$$F(H) = \mathbb{P}[F(H)] + \int_{]\tau,\infty[}\int_D \mathscr{I}_{s;y}F\,dM(s,y).$$

REMARKS. One class of examples of functions, F, satisfying the conditions of Theorems 2.4 and 2.5 consists of functions of the form $F(h) = f(h_{t_1}(g_1),\dots,h_{t_k}(g_k))$, where $t_i \in [\tau,\infty[$, g_i: $M_F(D) \to \mathbb{R}$ is bounded and continuous and f: $\mathbb{R}^k \to \mathbb{R}$ satisfies $|f(x) - f(y)| \le c|x-y|$, for some constant c. Another class is functions of the form $F(h) = f(\int_{]\tau,t_1]} h_s(g_1)\,ds,\dots,\int_{]\tau,t_k]} h_s(g_k)\,ds)$, where t_i, g_i and f are as before.

We will now describe our partial counterpart of the Wiener chaos expansion. First we need to define and establish the existence of the relevant multiple stochastic integrals. This needs some care. It is apparent from Perkins (1985) and Ruiz de Chavez (1985) that even defining general multiple stochastic integrals for continuous martingales with nondeterministic quadratic variation is a somewhat delicate matter. We begin with a lemma on parameterized stochastic integrals.

NOTATION. Write $\mathscr{P}$ for the $(\mathscr{H}_t)_{t\ge\tau}$-predictable σ-field of functions on $]\tau,\infty[\times\Omega$.

LEMMA 2.6. *Let* $(U,\mathscr{U})$ *be a measurable space. Suppose that* ϕ: $U \times]\tau,\infty[\times D \to \mathbb{R}$ *is bounded and* $\mathscr{U}\times\mathscr{B}(]\tau,\infty[)\times\mathscr{D}$*-measurable. Suppose that* ψ: $U\times[\tau,\infty[\times\Omega \to \mathbb{R}$ *is* $\mathscr{U}\times\mathscr{P}$*-measurable and satisfies*

$$\sup_{u\in U}\sup_{\tau<t\le\theta}\mathbb{P}\big[|\psi(u,t)|^p\big] < \infty,$$

for all $\theta > \tau$ *and* $p \geq 1$. *Then* $\int_{]\tau, t]} \int_D \phi(u, s, y) \psi(u, s)\, dM(s, y)$ *is defined for all* $u \in U$ *and* $t > \tau$ *and satisfies*

$$\sup_{u \in U} \sup_{\tau < t \leq \theta} \mathbb{P}\left[\left| \int_{]\tau, t]} \int_D \phi(u, s, y) \psi(u, s)\, dM(s, y) \right|^p \right] < \infty,$$

for all $\theta > \tau$ *and* $p \geq 1$. *Moreover, there exists a* $\mathscr{U} \times \mathscr{P}$*-measurable mapping* α *such that* $\alpha(u, \cdot, \cdot)$ *is indistinguishable from* $\int_{]\tau, \cdot]} \int_D \phi(u, s, y) \psi(u, s)\, dM(s, y)$, *for all* $u \in U$.

PROOF. We have

$$\begin{aligned} &\int_{]\tau, t]} \int_D (\phi(u, s, y) \psi(u, s))^2 H_s(dy)\, ds \\ &\qquad \leq \left(\sup_{u, s, y} |\phi(u, s, y)| \right)^2 \int_{]\tau, t]} \psi(u, s)^2 H_s(D)\, ds. \end{aligned}$$

As $\mathbb{P}[H_t(D)^q]$ is uniformly bounded on compact intervals for each $q \geq 1$, it is clear that $\int_{]\tau, t]} \int_D \phi(u, s, y) \psi(u, s)\, dM(s, y)$ is defined for all $u \in U$ and $t > \tau$ and (by the Burkholder–Davis–Gundy inequality) has moments with the stated properties. The existence of α can be proved in a manner similar to the proof of Proposition 5 in Stricker and Yor (1978). □

Suppose now that for some $m \geq 1$ we have functions $\phi_i \in b(\mathscr{B}(]\tau, \infty[)^{m-i+1} \times D)$, $i = 1, \ldots, m$.

Applying Lemma 2.6, we see that we can construct a $\mathscr{B}(]\tau, \infty[)^{m-1} \times \mathscr{P}$-measurable function α_1: $]\tau, \infty[^{m-1} \times]\tau, \infty[\times \Omega$ such that $\alpha_1(s_2, \ldots, s_m; \cdot)$ is indistinguishable from $\int_{]\tau, \cdot]} \int_D \phi_1(s_1, \ldots, s_m; y_1)\, dM(s_1, y_1)$, for all $s_2, \ldots, s_m$. Moreover,

$$\sup_{s_2, \ldots, s_m} \sup_{\tau < t \leq \theta} \mathbb{P}\left[|\alpha_1(s_2, \ldots, s_m; t)|^p \right] < \infty,$$

for all $\theta > \tau$ and $p \geq 1$.

It is easy to see that $(s_3, \ldots, s_m; s_2; \omega) \mapsto \alpha_1(s_2, s_3, \ldots, s_m; s_2; \omega)$ is $\mathscr{B}(]\tau, \infty[)^{m-2} \times \mathscr{P}$-measurable. We can therefore apply Lemma 2.6 again to construct a $\mathscr{B}(]\tau, \infty[)^{m-2} \times \mathscr{P}$-measurable function α_2: $]\tau, \infty[^{m-2} \times]\tau, \infty[\times \Omega$ such that $\alpha_2(s_3, \ldots, s_m; \cdot; \cdot)$ is indistinguishable from

$$\int_{]\tau, \cdot]} \int_D \phi_2(s_2, \ldots, s_m; y_2) \alpha_1(s_2, s_3, \ldots, s_m; s_2)\, dM(s_2, y_2),$$

for all $s_3, \ldots, s_m$, and

$$\sup_{s_3, \ldots, s_m} \sup_{\tau < t \leq \theta} \mathbb{P}\left[|\alpha_2(s_3, \ldots, s_m; t)|^p \right] < \infty,$$

for all $\theta > \tau$ and $p \geq 1$. Further, if $\tilde{\alpha}_1$ is another function with the same properties as α_1 and we construct $\tilde{\alpha}_2$ from $\tilde{\alpha}_1$ in the same manner that α_2 was constructed from α_1, then $\alpha_2(s_3, \ldots, s_m; \cdot)$ is indistinguishable from

$\tilde{\alpha}_2(s_3, \dots, s_m; \cdot)$, for all $s_3, \dots, s_m$, as

$$\int_{]\tau, t]} \int_D \phi_2(s_2, \dots, s_m; y_2)^2 \big(\alpha_1(s_2, \dots, s_m; s_2) - \tilde{\alpha}_1(s_2, \dots, s_m; s_2)\big)^2 H_{s_2}(dy_2)\, ds_2 = 0,$$

by Fubini.

Continuing in this way, we can construct successively $\alpha_3, \dots, \alpha_m$. We will write $I_m(\phi_1, \dots, \phi_m; t)$ for $\alpha_m(t)$. For $m \geq 1$ we will denote by $\mathscr{I}_m$ the set of all random variables of the form $I_m(\phi_1, \dots, \phi_m; t)$, for $\phi_1, \dots, \phi_m$ as above and $t > \tau$, and put $\mathscr{I} = \mathbb{R} \cup \bigcup_{m=1}^{\infty} \mathscr{I}_m$.

We can think of $I_m(\phi_1, \dots, \phi_m; t)$ as an attempt at giving meaning to the notation

$$\int_{\tau < s_1 < \cdots < s_m < t} \int_{(D)^m} \prod_{i=1}^{m} \phi_i(s_i, \dots, s_m; y_i)\, dM(s_1, y_1) \cdots dM(s_m, y_m).$$

Thus the linear span of $\mathscr{I}_m$ is analogous to the mth Wiener chaos (or perhaps more accurately, an L^2-dense linear subspace thereof). This analogy is not complete, because the two sets $\mathscr{I}_k$ and $\mathscr{I}_l$, $k \neq l$, are not orthogonal in $L^2(\mathbb{P})$. For example, if ϕ_1 and ϕ_2 are both constant functions taking the value 1, then

$$\mathbb{P}[I_1(\phi_1; t) I_2(\phi_1, \phi_2; t)] = \mathbb{P}\left[\int_{]\tau, t]} H_s(D)(H_s(D) - H_\tau(D))\, ds\right]$$

and the moment formulae in Dynkin (1988) show that the right-hand side is not 0. We thus cannot hope for a full analogue of the Wiener chaos expansion that involves the linear spans of the successive $\mathscr{I}_m$ in place of the mth Wiener chaos. We do, however, have the following partial analogue.

THEOREM 2.7. *The linear span of $\mathscr{I}$ is dense in $L^0(\Omega, \sigma\{H_t: t \geq \tau\}, \mathbb{P})$.*

REMARK. We should point out that our "multiple stochastic integrals" are rather different objects to the "multiple stochastic integrals" in Dynkin (1988). In particular, $t \mapsto I_m(\phi_1, \dots, \phi_m; t)$ is a martingale, a property not shared by the integrals in Dynkin (1988).

3. The marked historical process. Set $I = [0, 1]$ and $\tilde{E} = D \times I$. Let $\tilde{D}$ denote the space of cadlag functions from $\mathbb{R}_+$ to $\tilde{E}$. Write $\tilde{\mathscr{D}}$ for the Borel σ-field of $\tilde{D}$ and $(\tilde{\mathscr{D}}_t)$ for the canonical filtration. We will usually use the dummy variable x to denote a generic element of $\tilde{D}$ (or sometimes to denote a generic element of $\tilde{E}$). When we do so, we will follow the convention of denoting the E- and I-valued components of x as y and n, respectively. For $t \geq 0$, define X_t: $\tilde{D} \to \tilde{E}$ by $X_t(x) = x(t)$. With a slight abuse of notation, define Y_t: $\tilde{D} \to E$ and N_t: $\tilde{D} \to I$ by $Y_t(x) = y(t)$ and $N_t(x) = n(t)$.

Given $x \in \tilde{D}$ and $t \geq 0$, define $x^t \in \tilde{D}$ (as one would expect) by $x^t(s) = x(t \wedge s)$.

Let $M_F(\tilde{D})$ denote the space of finite measures on $\tilde{D}$ equipped with the topology of weak convergence and define $M_F(\tilde{D})^t$ by analogy with $M_F(D)^t$.

Given $x = (y, n) \in \tilde{E}$, let $\tilde{P}^{(y,n)}$ be the probability measure on $(\tilde{D}, \tilde{\mathscr{D}})$ that is the Cartesian product of P^y and the law of the Feller process on I that starts at n and has bounded generator given by

$$f \mapsto \int_I f(x)\, dx - f.$$

That is, under $\tilde{P}^{(y,n)}$ the processes Y and N are independent, with Y having the law P^y and N having the law of a pure jump Markov process on the I that starts at n, has i.i.d. mean 1 exponential interjump times and has jump sizes that are distributed according to Lebesgue measure, no matter where the jump begins.

Recall $\tau \geq 0$ and $\nu \in M_F(D)^\tau$ that appeared in the definition of the $(Y, -\lambda^2/2)$ historical process H. Let $\mu \in M_F(\tilde{D})^\tau$ be such that

$$\mu\big(\{x \in \tilde{D}: y \in A\}\big) = \nu(A),$$

for all $A \in \mathscr{D}$, and

$$\mu\big(\{x \in \tilde{D}: \{\exists\, t \geq 0: n(t) \neq 0\}\}\big) = 0.$$

Let G denote the $(X, -\lambda^2/2)$ historical process starting at (τ, μ) defined on some filtered probability space satisfying the usual conditions. That is, G is defined in a similar manner to H, except we are replacing D by $\tilde{D}$, (Y_t, P^y) by $(X_t, \tilde{P}^{(y,n)})$, ν by μ and so forth. Write $\tilde{P}^{\tau,\mu}$, $\tilde{\bar{\mathscr{D}}}_t^{\mu}$, $\tilde{F}^{\tau,\mu}$ and $\tilde{A}^{\tau,\mu}$ for the analogues of $P^{\tau,\nu}$, $\bar{\mathscr{D}}_t^{\nu}$, $F^{\tau,\nu}$ and $A^{\tau,\nu}$.

Define an $M_F(D)$-valued process $(H'_t)_{t\geq\tau}$ by

$$H'_t(A) = G_t\big(\{x \in \tilde{D}: y \in A\}\big), \qquad A \in \mathscr{D}.$$

Let $\mathscr{N}$ denote the σ-field generated by the null subsets of the probability space on which G is defined. Put

$$\mathscr{H}'_t = \bigcap_{s>t} \sigma\{H'_r: \tau \leq r \leq s\} \vee \mathscr{N}, \qquad t \geq \tau,$$

and

$$\mathscr{H}' = \bigvee_{t\geq\tau} \mathscr{H}'_t.$$

It is straightforward to check that $(H'_t)_{t\geq\tau}$ is a $(Y, -\lambda^2/2)$ historical process starting at (τ, ν) on either the filtered probability space on which G is defined or the filtered probability space with the same sample space and the "same" probability measure, but equipped with the smaller σ-fields $\mathscr{H}'$ and $(\mathscr{H}'_t)_{t\geq\tau}$.

We may therefore assume from now on that our $(X, -\lambda^2/2)$ historical process G and our $(Y, -\lambda^2/2)$ historical process H are defined on filtered probability spaces $(\Omega, \mathscr{G}, (\mathscr{G}_t)_{t\geq\tau}, \mathbb{P})$ and $(\Omega, \mathscr{H}, (\mathscr{H}_t)_{t\geq\tau}, \mathbb{P})$, respectively, that

$$\mathscr{H}_t = \bigcap_{s>t} \sigma\{H_r: \tau \leq r \leq s\} \vee \sigma\{\mathbb{P}\text{-null sets}\}, \qquad t \geq \tau,$$

$$\mathscr{H} = \bigvee_{t\geq\tau} \mathscr{H}_t$$

and

$$H_t(A) = G_t\big(\{x \in \tilde{D}: y \in A\}\big), \qquad A \in \mathscr{D}.$$

We may further assume that H is also a $(Y, -\lambda^2/2)$ historical process on the filtered probability space $(\Omega, \mathscr{G}, (\mathscr{G}_t)_{t \geq \tau}, \mathbb{P})$.

It will be clear, however, that our main results, Theorems 2.4 and 2.5, will hold whether or not H and $(\Omega, \mathscr{H}, (\mathscr{H}_t)_{t \geq \tau}, \mathbb{P})$ is of this special form.

By replacing $\mathscr{G}$ by $\bigvee_{t \geq \tau} \mathscr{G}_t$ if necessary, we may also suppose that $\mathscr{G} = \bigvee_{t \geq \tau} \mathscr{G}_t$.

Let K denote the orthogonal martingale measure that may be constructed from G in a manner analogous to that by which M is constructed from H. If T is a bounded $(\mathscr{G}_t)_{t \geq \tau}$-stopping time, then we will let $\overline{\mathbb{P}}_T$ denote the normalized Campbell measure associated with G_T, defined by the counterpart of the construction in Section 2.

4. Preliminary results.

The following simple lemma does not seem to appear explicitly in the literature.

LEMMA 4.1. *Suppose that* $\{(W_t^k)_{t \geq \tau}\}_{k=1}^{\infty}$ *is a sequence of continuous* $(\mathscr{G}_t)_{t \geq \tau}$*-martingales under* $\mathbb{P}$ *such that*

$$W_t^k = \int_{]\tau,\, t]} \int_{\tilde{D}} \phi^k(s, x)\, dK(s, x),$$

for some integrand ϕ^k *that is predictable with respect to the filtration* $(\tilde{\mathscr{D}} \times \mathscr{G}_t)_{t \geq \tau}$. *If* W_t^k *converges in* $L^1(\mathbb{P})$ *for each* $t \geq \tau$, *then the limit process has a continuous version* $(W_t)_{t \geq \tau}$, *which is a continuous* $(\mathscr{G}_\tau)_{t \geq \tau}$ *martingale. Moreover, there exists a* $(\tilde{\mathscr{D}} \times \mathscr{G}_t)_{t \geq \tau}$*-predictable integrand* ϕ *such that*

$$\lim_{k \to \infty} \int_{]\tau,\, t]} |\phi^k(s, x) - \phi(s, x)|^2 G_s(dx)\, ds = 0$$

in $\mathbb{P}$*-probability, for all* $t \geq \tau$, *and*

$$W_t = \int_{]\tau,\, t]} \int_{\tilde{D}} \phi(s, x)\, dK(s, x),$$

for all $t \geq \tau$, $\mathbb{P}$*-a.s. In particular, if* ψ *is such that* $\phi^k(s, x) \to \psi(s, x)$ *for* G_s*-a.e.* x, *Lebesgue-a.e.* s, $\mathbb{P}$*-a.s., then*

$$W_t = \int_{]\tau,\, t]} \int_{\tilde{D}} \psi(s, x)\, dK(s, x).$$

PROOF. Fix a subsequence $\{j_k\}_{k=1}^{\infty}$ of positive integers. By Doob's L^1 maximal inequality,

$$\mathbb{P}\Big\{ \sup_{\tau \leq s \leq t} |W_s^{j_k} - W_s^{j_{k'}}| \geq c \Big\} \leq \mathbb{P}\big[|W_t^{j_k} - W_t^{j_{k'}}|\big]/c,$$

for every $t \geq \tau$ and $c > 0$. Thus $\{j_k\}_{k=1}^{\infty}$ contains a further subsequence $\{l_k\}_{k=1}^{\infty}$ such that, almost surely, W^{l_k} converges uniformly on $[\tau, t]$, for each $t \geq \tau$, to a continuous process W that is clearly a martingale.

For each positive integer p define a $(\mathscr{G}_t)_{t \geq \tau}$-stopping time $T(p)$ by

$$T(p) = \inf\left\{t \geq \tau \colon \sup_k |W_t^{l_k}| \geq p\right\}.$$

Then $T(1) \leq T(2) \leq \cdots$ and $T(p) \to \infty$ $\mathbb{P}$-a.s. as $p \to \infty$. Moreover, for each p, $(W_{t \wedge T(p)}^{l_k})_{t \geq \tau}$ is a bounded martingale such that

$$W_{t \wedge T(p)}^{l_k} = \int_{]\tau, t]} \int_{\tilde{D}} \phi^{l_k}(s, x) \mathbf{1}(s \leq T(p)) \, dK(s, x)$$

and

$$\lim_{k \to \infty} \mathbb{P}\left[\left(W_{t \wedge T(p)} - W_{t \wedge T(p)}^{l_k}\right)^2\right] = 0,$$

for each $t \geq \tau$. Therefore, for each p there exists a $(\tilde{\mathscr{D}} \times \mathscr{G}_t)_{t \geq \tau}$-predictable integrand $\phi^{(p)}$ such that

$$\lim_{k \to \infty} \mathbb{P}\left[\int_\tau^t \int_{\tilde{D}} |\phi^{l_k}(s, x) \mathbf{1}(s \leq T(p)) - \phi^{(p)}(s, x)|^2 G_s(dx) \, ds\right] = 0$$

and

$$W_{t \wedge T(p)} = \int_{]\tau, t]} \int_{\tilde{D}} \phi^{(p)}(s, x) \, dK(s, x).$$

It is clear that we in fact have $\phi^{(p)}(s, x) = \phi(s, x)\mathbf{1}(s \leq T(p))$ for some $(\tilde{\mathscr{D}} \times \mathscr{G}_t)_{t \geq \tau}$-predictable integrand ϕ and

$$W_t = \int_{]\tau, t]} \int_{\tilde{D}} \phi(s, x) \, dK(s, x),$$

for all $t \geq \tau$.

Given that such a ϕ exists, it is clear that it is essentially unique and, in particular, does not depend on the subsequences used in the construction. Moreover, we have that

$$\lim_{k \to \infty} \int_{]\tau, t]} |\phi^{l_k}(s, x) - \phi(s, x)|^2 G_s(dx) \, ds = 0$$

in $\mathbb{P}$-probability for all $t \geq \tau$, and since the sequence $\{j_k\}_{k=1}^\infty$ was arbitrary, we must in fact have

$$\lim_{k \to \infty} \int_{]\tau, t]} |\phi^k(s, x) - \phi(s, x)|^2 G_s(dx) \, ds = 0$$

in $\mathbb{P}$-probability for all $t \geq \tau$.

The proof of the remainder of the lemma is straightforward and is omitted. □

Given any cadlag function, n, from $\mathbb{R}_+$ to I, we may construct a σ-finite counting measure on $\mathbb{R}_+ \times I$ by assigning an atom of mass 1 to each point (s, z) such that $n(s) - n(s-) = z \neq 0$. With a slight abuse of notation we will denote this measure by the same letter as the function from which it was constructed (in this case n). The main purpose of this section is to obtain semimartingale decompositions for processes such as $\int(\int_{]\tau,t]} f(s, z, x)n(ds, dz))G_t(dx)$, $t \geq \tau$, where f belongs to a suitable class of (possibly random) integrands. Our line of development will be similar to that followed in Perkins (1992, 1995) to produce a "stochastic calculus along branches" for historical Brownian motion, but our task is somewhat easier due to the fact that the integrals we are dealing with are essentially sums rather than Itô integrals.

LEMMA 4.2. *Suppose that f: $]\tau,\infty[\times I \times \tilde{D} \to \mathbb{R}$ is bounded and predictable with respect to the filtration $(\mathscr{B}(I) \times \tilde{\mathscr{D}}_t)_{t \geq \tau}$. Then*

$$\int_{]\tau,\, t]\times I} f(s,z,(Y^s,N^s))\,N^t(ds,dx)$$
$$-\int_{]\tau,\, t]}\int_I f(s,z,(Y^s,N^s))\,dz\,ds, \qquad t \geq \tau,$$

is a cadlag $(\bar{\tilde{\mathscr{D}}}{}_t^{\mu})_{t \geq \tau}$-martingale under $\tilde{P}^{\tau,\mu}$.

For the proof, see Section II.3 of Ikeda and Watanabe (1981).

NOTATION. Given a locally finite measure ρ on $\mathbb{R}_+ \times I$, define a locally finite signed measure $\check{\rho}$ on $\mathbb{R}_+ \times I$ by

$$\int g(s,z)\,\check{\rho}(ds,dz) = \int g(s,z)\,\rho(ds,dz) - \int g(s,z)\,ds \otimes dz.$$

LEMMA 4.3. *Suppose that f: $]\tau,\infty[\times I \times \tilde{D} \times \Omega \to \mathbb{R}$ is bounded and predictable with respect to the filtration $(\mathscr{B}(I) \times \tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t \geq \tau}$.*

(a) *We have, $\mathbb{P}$-a.s.,*

$$\int_{\tilde{D}}\left(\int_{]\tau,\, t]\times I} f(s,z,x)\,n(ds,dz)\right)G_t(dx) - \int_{]\tau,\, t]}\int_{\tilde{D}}\int_I f(s,z,x)\,dz\,G_s(dx)\,ds$$
$$= \int_{]\tau,\, t]}\int_{\tilde{D}}\left(\int_{]\tau,\, s]\times I} f(u,z,x)\,n(du,dz)\right)dK(s,x),$$

for all $t \geq \tau$, and both sides are continuous martingales.

(b) *We have, $\mathbb{P}$-a.s.,*

$$\int_{\tilde{D}}\left(\int_{]\tau,\, t]\times I} f(s,z,x)\,\check{n}(ds,dz)\right)G_t(dx)$$
$$= \int_{]\tau,\, t]}\int_{\tilde{D}}\left(\int_{]\tau,\, s]\times I} f(u,z,x)\,\check{n}(du,dz)\right)dK(s,x),$$

for all $t \geq \tau$, and both sides are continuous martingales.

(c) *If $T \geq \tau$ is a bounded $(\mathscr{G}_t)_{t \geq \tau}$-stopping time, then*

$$(t, x, \omega) \mapsto \int_{]\tau, t \wedge T(\omega)] \times I} f(s, z, x, \omega) \check{n}(ds, dz), \qquad t \geq \tau,$$

is a $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t \geq \tau}$-martingale under $\overline{\mathbb{P}}_T$.

(d) *The martingale in part* (b) *has values in $L^2(\mathbb{P})$.*

PROOF. (a) Consider first the special case in which f is replaced by g: $]\tau, \infty[\times I \times \tilde{D} \to \mathbb{R}$, where g is bounded and predictable with respect to the filtration $(\mathscr{B}(I) \times \tilde{\mathscr{D}}_t)_{t \geq \tau}$.

Observe that

$$\mathbb{P}\left[\int_{\tilde{D}} n(]\tau,t] \times I) G_t(dx)\right] = \tilde{P}^{\tau,\mu}[N^t(]\tau,t])] = \mu(\tilde{D})(t - \tau)$$

by the form of the first moment measure of G [cf. (2.1)]. Similarly,

$$\mathbb{P}\left[\int_{]\tau, t]} G_s(\tilde{D}) \, ds\right] = \mu(\tilde{D})(t - \tau).$$

In particular, $n(]\tau, t] \times I) < \infty$ for G_t-a.e. x, $\mathbb{P}$-a.s. The lemma in this special case will therefore follow from Lemma 4.1 and dominated convergence if we can show that it holds when g is replaced by $(s, z, x) \mapsto g(s, z, x)\mathbf{1}(n(]\tau, s[\times I) \leq k)$ for arbitrary $k > 0$. This, however, is clear from the definition of G and Lemma 4.2.

Now consider the special case in which f is of the form $g\mathbf{1}_{]v,w]}\gamma$, where g is as above, $\tau < v < w < \infty$ and $\gamma \in b\mathscr{G}_r$ for some $r < v$. Then, by the above,

$$\begin{aligned}
&\int_{\tilde{D}} \left(\int_{]\tau, t] \times I} f(s, z, x) n(ds, dz)\right) G_t(dx) \\
&\quad = \int_{\tilde{D}} \left(\int_{]\tau, t] \times I} g(s, z, x)\mathbf{1}_{]v, w]}(s) n(ds, dz)\right) G_t(dx)\gamma \\
&\quad = \int_{]\tau, t]} \int_{\tilde{D}} \int_I g(s, z, x)\mathbf{1}_{]v, w]}(s) \, dz \, G_s(dx) \, ds \, \gamma \\
&\qquad + \int_{]\tau, t]} \int_{\tilde{D}} \left(\int_{]\tau, s] \times I} g(u, z, x)\mathbf{1}_{]v, w]}(u) n(du, dz)\right) dK(s, x)\gamma \\
&\quad = \int_{]\tau, t]} \int_{\tilde{D}} \int_I f(s, z, x) \, dz \, G_s(dx) \, ds \\
&\qquad + \int_{]\tau, t]} \int_{\tilde{D}} \left(\int_{]\tau, s] \times I} f(u, z, x) n(du, dz)\right) dK(s, x),
\end{aligned}$$

where the last equation holds because the stochastic integrand is 0 for values of $s \leq r$. The lemma therefore also holds in this special case.

As in the proof of Lemma 2.1, functions f of this form generate the predictable σ-field for the filtration $(\mathscr{B}(I) \times \tilde{\mathscr{D}}_t \times G_t)_{t \geq \tau}$, and the general result follows from a monotone class argument and Lemma 4.1.

(b) This follows from Lemma 4.2 in essentially the same manner as part (a).

(c) Suppose that $t \geq u \geq \tau$, $\gamma \in b\tilde{\mathscr{D}}_u$ and $\xi \in b\mathscr{G}_u$. Then

$$\begin{aligned}
&\overline{\mathbb{P}}_T\left[\int_{]u\wedge T, t\wedge T]\times I} f(s,z,x)\check{n}(ds,dz)\gamma\xi\right]\mu(\tilde{D}) \\
&\quad = \mathbb{P}\left[\int_{\tilde{D}}\int_{]u\wedge T, t\wedge T]\times I} f(s,z,x)\gamma(x)\check{n}(ds,dz)G_T(dx)\,\xi\right] \\
&\quad = \mathbb{P}\left[\int_{\tilde{D}}\int_{]\tau, T]\times I} f(s,z,x)\gamma(x)\mathbf{1}_{]u\wedge T, t\wedge T]}(s)\check{n}(ds,dz)G_T(dx)\,\xi\right] \\
&\quad = \mathbb{P}\left[\int_{\tilde{D}}\int_{]\tau, u]\times I} f(s,z,x)\gamma(x)\mathbf{1}_{]u\wedge T, t\wedge T]}(s)\check{n}(ds,dz)G_u(dx)\,\xi\right] \\
&\quad = 0,
\end{aligned}$$

by part (b). A monotone class argument completes the proof.

(d) First note for $t \geq \tau$ that

$$\overline{\mathbb{P}}_t\left[\left(\int_{]\tau, t]\times I} f(s,z,x)\check{n}(ds,dz)\right)^2\right] = \overline{\mathbb{P}}_t\left[\int_{]\tau, t]\times I} f(s,z,x)^2\, ds\otimes dz\right]$$

by part (c) and formula II.3.9 of Ikeda and Watanabe (1981). Thus,

$$\begin{aligned}
&\mathbb{P}\left[\left\{\int_{]\tau, t]}\int_{\tilde{D}}\left(\int_{]\tau, s]\times I} f(u,z,x)\check{n}(du,dz)\right) dK(s,x)\right\}^2\right] \\
&\quad = \mathbb{P}\left[\int_{]\tau, t]}\int_{\tilde{D}}\left(\int_{]\tau, s]\times I} f(u,z,x)\check{n}(du,dz)\right)^2 G_s(dx)\, ds\right] \\
&\quad = \int_{]\tau, t]}\overline{\mathbb{P}}_s\left[\left(\int_{]\tau, s]\times I} f(u,z,x)\check{n}(du,dz)\right)^2\right] ds\, \mu(\tilde{D}) \\
&\quad = \int_{]\tau, t]}\overline{\mathbb{P}}_s\left[\int_{]\tau, s]\times I} f(u,z,x)^2\, du\otimes dz\right] ds\, \mu(\tilde{D}) < \infty. \qquad \square
\end{aligned}$$

LEMMA 4.4. *Suppose that* $\phi \in F^{\tau,\nu}$ *is such that* $t \mapsto \phi_t(Y)$ *is cadlag* $P^{\tau,\nu}$*-a.s. Then*

$$\int_{\tilde{D}}\left(\int_{]\tau, t]\times I}\mathbf{1}(\phi(s,y) \neq \phi(s-,y))n(ds,dz)\right)G_t(dx) = 0,$$

for all $t \geq \tau$, $\mathbb{P}$*-a.s.*

PROOF. Whereas the set of times $t > \tau$ such that $\phi(t,Y) \neq \phi(t-,Y)$ is $P^{\tau,\mu}$-a.s. countable, it follows from Fubini's theorem that $\tilde{P}^{\tau,\mu}$-a.s. there exists no time $t > \tau$ such that $\phi(t,Y) = \phi(t-,Y)$ and $N_t \neq N_{t-}$. The pair $((t,x) \mapsto \mathbf{1}(\int_{]\tau,t]\times I}\mathbf{1}(\phi(s,y) \neq \phi(s-,y))n(ds,dz) \neq 0), 0)$ thus belongs to

$\tilde{A}^{\tau,\mu}$. From the definition of G we have that

$$G_t\left(\left\{x : \mathbf{1}\left(\int_{]\tau,t]\times I} \mathbf{1}(\phi(s,y) \neq \phi(s-,y))n(ds,dz) \neq 0\right)\right\}\right)$$

is a continuous nonnegative martingale that is null at τ, and hence this process is identically zero. □

LEMMA 4.5. *Suppose that $(\phi, \psi) \in \tilde{A}^{\tau,\mu}$ and that $T \geq \tau$ is a $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t\geq\tau}$-stopping time such that the graph of T satisfies $[T] \subset \bigcup_m [U_m] \times \Omega$, where $\{U_m\}_{m=1}^{\infty}$ is a countable collection of $(\tilde{\mathscr{D}}_{t+})_{t\geq\tau}$-stopping times. Set*

$$\Phi(t,x) = \phi(t,x) - \phi(\tau,x) - \int_{]\tau,t]} \psi(s,x)\,ds, \qquad t \geq \tau.$$

Then, $\mathbb{P}$-a.s.,

$$\int_{\tilde{D}} \Phi(t \wedge T(x), x) G_t(dx) = \int_{]\tau,t]} \int_{\tilde{D}} \Phi(s \wedge T(x), x)\, dK(s,x),$$

for all $t \geq \tau$, and both sides are continuous martingales with values in $L^2(\mathbb{P})$.

PROOF. Suppose that $v \geq u \geq \tau$, $\gamma \in b\tilde{\mathscr{D}}_u$ and $\xi \in b\mathscr{G}_u$. It is immediate from the definition of G that

$$\int_{\tilde{D}} (\Phi(t \wedge v, x) - \Phi(t \wedge u, x))\gamma(x)\xi G_t(dx)$$
$$= \int_{]\tau,t]} \int_{\tilde{D}} (\Phi(s \wedge v, x) - \Phi(s \wedge u, x))\gamma(x)\xi\, dK(s,x),$$

for all $t \geq \tau$, and both sides are continuous martingales with values in $L^2(\mathbb{P})$. A monotone class argument shows that this conclusion still holds if we replace $(x,\omega) \mapsto \gamma(x)\xi(\omega)$ by $(x,\omega) \mapsto \eta(x,\omega)$ with $\eta \in b(\tilde{\mathscr{D}}_u \times \mathscr{G}_u)$.

Now consider a bounded $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t\geq\tau}$-stopping time V that can only take on finitely many values, say, $\tau = u_0 < u_1 < \cdots < u_k < \infty$. Observe that

$$\Phi(t \wedge V(x), x) = \sum_{l=0}^{k-1} (\Phi(t \wedge u_{l+1}, x) - \Phi(t \wedge u_l, x))\eta_l(x),$$

where $\eta_l(x,\omega) = \mathbf{1}(V(x,\omega) > u_l) \in b(\tilde{\mathscr{D}}_{u_l} \times \mathscr{G}_{u_l})$, and so the conclusion of the theorem holds with T replaced by V by the above.

In order to prove the theorem, it suffices to consider the case of a bounded T. Write T as the limit of a decreasing sequence $\{V_k\}$ of bounded $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)$-stopping times that each can only take on finitely many values.

We will establish below that for each m we have

$$\lim_{t \downarrow U_m(x^s)} \phi(s \wedge t, x^s) = \phi(s \wedge U_m(x^s), x^s) \tag{4.1}$$

for G_s-a.e. x, for all $s \geq \tau$, $\mathbb{P}$-a.s.

Assuming (4.1), we have

$$\begin{aligned}
&\mathbb{P}\left[\exists\, s \ge \tau,\ \int_{\tilde D} \Phi(s \wedge T(x), x) G_s(dx) \ne \lim_k \int_{\tilde D} \Phi(s \wedge V_k(x), x) G_s(dx)\right] \\
&\quad \le \sum_m \mathbb{P}\Big[\exists\, s \ge \tau,\ \int_{\tilde D} \Phi(s \wedge U_m(x), x) \mathbf{1}\Big(\lim_k V_k(x) = U_m(x)\Big) G_s(dx) \\
&\qquad\qquad \ne \lim_k \int_{\tilde D} \Phi(s \wedge V_k(x), x) \mathbf{1}\Big(\lim_k V_k(x) = U_m(x)\Big) G_s(dx)\Big] \\
&\quad \le \sum_m \mathbb{P}\Big[\exists\, s \ge \tau,\ \int_{\tilde D} \Phi(s \wedge U_m(x), x) G_s(dx) \\
&\qquad\qquad \ne \int_{\tilde D} \lim_{t \downarrow U_m(x)} \Phi(s \wedge t, x) G_s(dx)\Big] \\
&\quad = 0.
\end{aligned}$$

Similarly, we have

$$\lim_k \int_{]\tau, s]} \int_{\tilde D} \big(\Phi(u \wedge V_k(x), x) - \Phi(u \wedge T(x), x)\big)^2 G_u(dx)\, du = 0,$$

for all $s \ge \tau$, $\mathbb{P}$-a.s. Thus, for all $s \ge \tau$,

$$\sup_{\tau \le t \le s} \left| \int_{]\tau, t]} \int_{\tilde D} \Phi(u \wedge V_k(x), x)\, dK(u, x) - \int_{]\tau, t]} \int_{\tilde D} \Phi(u \wedge T(x), x)\, dK(u, x) \right|$$

converges to 0 in $\mathbb{P}$-probability as $k \to \infty$. From the above we have for each k that, $\mathbb{P}$-a.s. for all $s \ge \tau$,

$$\int_{\tilde D} \Phi(s \wedge V_k(x), x) G_s(dx) = \int_{]\tau, s]} \int_{\tilde D} \Phi(u \wedge V_k(x), x)\, dK(u, x),$$

and this establishes the desired contradiction.

It thus remains to establish (4.1). Fix m and write U for U_m. Set

$$\phi_+(s, x) = \limsup_{t \downarrow U(x^s)} \phi(s \wedge t, x^s)$$

and

$$\phi_-(s, x) = \liminf_{t \downarrow U(x^s)} \phi(s \wedge t, x^s).$$

Both ϕ_+ and ϕ_- are universally measurable [cf. Theorem 33 in Section IV of Dellacherie and Meyer (1978)] and satisfy $\phi_+(s, x) = \phi_+(s, x^s)$ and $\phi_-(s, x) = \phi_-(s, x^s)$. Also, $\phi_+(s, X) = \phi_-(s, X) = \phi(s \wedge U(X^s), X^s)$, for all $s \ge \tau$, $\tilde P^{\tau, \mu}$-a.s., by definition of $\tilde F^{\tau, \mu}$. The assumption of Borel measurability in Theorem 2.3.a of Mueller and Perkins (1992) can be weakened to universal measurability, and applying this improved result we get $\phi_+(s, x) = \phi_-(s, x) = \phi(s \wedge U(x^s), x^s)$, for G_s-a.e. x, for all $s \geq \tau$, $\mathbb{P}$-a.s. $\square$

An analogue of the following result appears as Theorem 3.8 of Perkins (1992) for the case of historical Brownian motion. The proof in our setting is identical and the statement is included here for ease of reference.

THEOREM 4.6. *Suppose that $b\colon]\tau,\infty[\times\tilde{D}\times\Omega \to \mathbb{R}$ is bounded and predictable with respect to the filtration $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t\geq\tau}$. Then*

$$\int_{\tilde{D}}\int_{]\tau,\, t]} b(s, x)\, ds\, G_t(dx) = \int_{]\tau,\, t]}\int_{\tilde{D}}\left(\int_{]\tau,\, s]} b(u, x)\, du\right) dK(s, x) + \int_{]\tau,\, t]}\int_{\tilde{D}} b(s, x) G_s(dx)\, ds,$$

for all $t \geq \tau$, $\mathbb{P}$-a.s., and each term is almost surely continuous in $t \geq \tau$.

The following general stochastic integral representation is a variant of Theorem 1.1 in Evans and Perkins (1994).

THEOREM 4.7. *Let W be a $\sigma\{H_t\colon t \geq \tau\}$-measurable random variable that belongs to $L^2(\mathbb{P})$. Then there exists a $(\mathscr{D}_t \times \mathscr{H}_t)_{t\geq\tau}$-predictable function ρ such that*

$$\mathbb{P}\left[\int_{]\tau,\infty[}\int_D \rho(s, y)^2 H_s(dy)\, ds\right] < \infty$$

and

$$W = \mathbb{P}[W] + \int_{]\tau,\infty[}\int_D \rho(s, y)\, dM(s, y).$$

PROOF. The processes $H_t(\phi_t)$, where $(\phi, \psi) \in A^{\tau,\nu}$, for some ψ, solve a well-posed martingale problem of the type studied in Jacod (1977). From Theorem 2 and Proposition 2 of Jacod (1977), it follows that for each $n \in \mathbb{N}$ there exists $(\phi_n^1, \psi_n^1), \dots, (\phi_n^{N(n)}, \psi_n^{N(n)}) \in A^{\tau,\nu}$, bounded $(\mathscr{H}_t)_{t\geq\tau}$-predictable $\xi_n^1, \dots, \xi_n^{N(n)}$ and $t_n \in]\tau,\infty[$ such that $t_n \uparrow \infty$ as $n \to \infty$ and

$$W = \mathbb{P}[W] + \lim_{n\to\infty}\int_{]\tau,\, t_n]}\int_D \sum_i \xi_n^i(s)\phi_n^i(s, y)\, dM(s, y),$$

where the convergence is in $L^2(\mathbb{P})$.

We claim that for each $(\phi, \psi) \in A^{\tau,\nu}$ and each bounded $(\mathscr{H}_t)_{t\geq\tau}$-predictable ξ there exists a bounded, $(\mathscr{D}_t)_{t\geq\tau}$-predictable function γ such that, $\mathbb{P}$-a.s.,

$$\int_{]\tau,\, t]}\int_D \xi(s)\phi(s, y)\, dM(s, y) = \int_{]\tau,\, t]}\int_D \xi(s)\gamma(s, y)\, dM(s, y),$$

for all $t > \tau$. From Theorem IV.97 of Dellacherie and Meyer (1978) it follows that any such ϕ is $(\mathscr{D}_t)_{t \geq \tau}$-optional. Thus, by Theorem IV.66 of Dellacherie and Meyer (1978), there is a bounded, $(\mathscr{D}_t)_{t \geq \tau}$-predictable function γ such that for all $y \in D$ the set $\{s > \tau: \phi(s, y) \neq \gamma(s, y)\}$ is countable. In particular, $\int_{]\tau, t]}(\phi(s, Y) - \gamma(s, Y))^2 \, ds = 0$, for all $t > \tau$, $P^{\tau, \nu}$-a.s. Hence, by (2.1), for some constant c, we have

$$\begin{aligned}
&\mathbb{P}\left[\left\{\int_{]\tau, t]} \int_D \xi(s)\phi(s, y)\, dM(s, y) - \int_{]\tau, t]} \int_D \xi(s)\gamma(s, y)\, dM(s, y)\right\}^2\right] \\
&\quad = \mathbb{P}\left[\int_{]\tau, t]} \int_D \xi(s)^2 (\phi(s, y) - \gamma(s, y))^2 H_s(dy)\, ds\right] \\
&\quad \leq c\mathbb{P}\left[\int_{]\tau, t]} \int_D (\phi(s, y) - \gamma(s, y))^2 H_s(dy)\, ds\right] \\
&\quad = 0,
\end{aligned}$$

for all $t > \tau$, and so γ has the properties we seek.

Thus there exists for each $n \in \mathbb{N}$ a bounded $(\mathscr{D}_t \times \mathscr{H}_t)_{t \geq \tau}$-predictable function ρ_n such that

$$W = \mathbb{P}[W] + \lim_{n \to \infty} \int_{]\tau, \infty[} \int_D \rho_n(s, y)\, dM(s, y),$$

where the convergence is in $L^2(\mathbb{P})$. This implies that

$$\lim_{n, n' \to \infty} \mathbb{P}\left[\int_{]\tau, \infty[} \int_D (\rho_n(s, y) - \rho_{n'}(s, y))^2 H_s(dy)\, ds\right] = 0,$$

for all $t > \tau$. Hence there exists a $(\mathscr{D}_t \times \mathscr{H}_t)_{t \geq \tau}$-predictable ρ such that

$$\mathbb{P}\left[\int_{]\tau, \infty]} \int_D \rho(s, y)^2 H_s(dy)\, ds\right] < \infty$$

and

$$\lim_{n \to \infty} \mathbb{P}\left[\int_{]\tau, \infty[} \int_D (\rho_n(s, y) - \rho(s, y))^2 H_s(dy)\, ds\right] = 0.$$

This implies that

$$\begin{aligned}
&\mathbb{P}[W] + \int_{]\tau, \infty[} \int_D \rho(s, y)\, dM(s, y) \\
&\qquad = \mathbb{P}[W] + \lim_{n \to \infty} \int_{]\tau, \infty[} \int_D \rho_n(s, y)\, dM(s, y) = W,
\end{aligned}$$

as required. □

5. Killing and resurrection.

NOTATION. Suppose that β: $]\tau, \infty[\times D \times \Omega \to I$ is predictable with respect to the filtration $(\mathscr{D}_t \times \mathscr{G}_t)_{t \geq \tau}$. Put

$$\begin{aligned} A(t, x, \omega) &= n(\{(s, z) \in]\tau, t[\times I: \beta(s, y, \omega) > z\}), \\ B(t, x, \omega) &= \mathbf{1}(A(t, x, \omega) = 0), \\ C(t, z, y, \omega) &= \mathbf{1}(\beta(t, y, \omega) > z). \end{aligned}$$

Define a new $M_F(D)$-valued process $(H_t^\beta)_{t \geq \tau}$ by

$$H_t^\beta(\Xi) = \int_{\tilde{D}} \mathbf{1}_\Xi(y) B(t, x) G_t(dx), \qquad t \geq \tau, \Xi \in \mathscr{D}.$$

Put

$$\begin{aligned} \mathscr{R}_t^\beta &= \exp\left(\int_{]\tau, t]} \int_D \beta(s, y)\, dM(s, y) - \frac{1}{2} \int_{]\tau, t]} \int_D \beta(s, y)^2 H_s(dy)\, ds \right) \\ &= \exp\left(\int_{]\tau, t]} \int_{\tilde{D}} \beta(s, y)\, dK(s, x) - \frac{1}{2} \int_{]\tau, t]} \int_{\tilde{D}} \beta(s, y)^2 G_s(dx)\, ds \right), \\ &\qquad t \geq \tau. \end{aligned}$$

It can be shown that $(\mathscr{R}_t^\beta)_{t \geq \tau}$ is a $(\mathscr{G}_t)_{t \geq \tau}$-martingale, and so we may define a new probability measure $\mathbb{P}^\beta$ on $(\Omega, \mathscr{G})$ by setting $\mathbb{P}^\beta[\Lambda] = \mathbb{P}[\Lambda \mathscr{R}_t^\beta]$, for $\Lambda \in b\mathscr{G}_t$ (recall our standing assumption that $\mathscr{G} = \bigvee_{t \geq \tau} \mathscr{G}_t$). The proof is essentially the same as that of Theorem 2.3.b in Evans and Perkins (1994) once one has a domination result analogous to Theorem 2.1 in Evans and Perkins (1994) or Theorem 5.1 in Barlow, Evans and Perkins (1991). Such a result will, in turn, be a consequence of an existence and uniqueness result for a "historical process with immigration" martingale problem that is analogous to Theorem 1.1 in Barlow, Evans and Perkins (1991).

We do not present the proof here because it is rather lengthy, and for the proof of our main results (Theorems 2.4, 2.5 and 2.7), we only need the trivial special case in which both

$$\int_{]\tau, t]} \int_D \beta(s, y)\, dM(s, y) \quad \text{and} \quad \int_{]\tau, t]} \int_D \beta(s, y)^2 H_s(dy)\, ds$$

are uniformly bounded (see the proof of Lemma 6.1). A full proof will appear in a more general situation in Barlow, Evans and Perkins (1995). We remark that the usual exponential moment criteria for establishing that stochastic exponentials are martingales [cf. Section VII.1 of Revuz and Yor (1991)] do not apply here.

THEOREM 5.1. *The law of H^β under $\mathbb{P}^\beta$ is the same as the law of H under $\mathbb{P}$.*

PROOF. Suppose that $(\phi, \psi) \in A^{\tau, \nu}$. Set

$$T(x, \omega) = \inf\{t \geq \tau: A(t, x, \omega) > 0\}.$$

From Lemma 4.3(a) we know that $\int_{\tilde{D}} n(]\tau, t] \times I) G_t(dx) - \int_{]\tau, t]} G_s(\tilde{D})\, ds$ is a continuous martingale. In particular, $n(]\tau, t] \times I)$ is finite for G_t-a.e. $x \in \tilde{D}$ for all $t \geq \tau$, $\mathbb{P}$-a.s. Thus,

$$\begin{aligned}
&\int_D \phi(t, y) H_t^\beta(dy) \\
&\quad = \int_{\tilde{D}} \phi(t \wedge T(x), y) G_t(dx) \qquad (5.1)\\
&\qquad - \int_{\tilde{D}} \int_{]\tau, t] \times I} C(s, z, y) B(s, x) \phi(s, y) n(ds, dz) G_t(dx).
\end{aligned}$$

First consider the first term on the right-hand side of (5.1). Put

$$\Phi(t, y) = \phi(t, y) - \phi(\tau, y) - \int_{]\tau, t]} \psi(s, y)\, ds.$$

For $\varepsilon > 0$ put

$$T^\varepsilon(x, \omega) = \inf\{t \geq \tau: A(t+, x, \omega) > 0, n(t) - n(t-) > \varepsilon\}$$

and

$$B^\varepsilon(t, x, \omega) = \mathbf{1}(t \in]\tau, T^\varepsilon(x, \omega)]).$$

Observe that T^ε is a $(\tilde{\mathscr{D}}_t \times \mathscr{G}_t)_{t \geq \tau}$-stopping time. We have

$$\begin{aligned}
&\int_{\tilde{D}} \phi(t \wedge T^\varepsilon(x), y) G_t(dx) \\
&\quad = \int_{\tilde{D}} \Phi(t \wedge T^\varepsilon(x), y) G_t(dx) \\
&\qquad + \int_{\tilde{D}} \phi(\tau, y) G_t(dx) + \int_{\tilde{D}} \int_{]\tau, t \wedge T^\varepsilon(x)]} \psi(s, y)\, ds\, G_t(dx) \\
&\quad = \int_{]\tau, t]} \int_{\tilde{D}} \Phi(s \wedge T^\varepsilon(x), y)\, dK(s, x) \\
&\qquad + \int_{\tilde{D}} \phi(\tau, y) G_\tau(dx) + \int_{]\tau, t]} \int_{\tilde{D}} \phi(\tau, y)\, dK(s, x) \\
&\qquad + \int_{]\tau, t]} \int_{\tilde{D}} \left(\int_{]\tau, s]} \psi(u, y) B^\varepsilon(u, x)\, du \right) dK(s, x) \\
&\qquad + \int_{]\tau, t]} \int_{\tilde{D}} \psi(s, y) B^\varepsilon(s, x) G_s(dx)\, ds \\
&\quad = \int_{\tilde{D}} \phi(\tau, y) G_\tau(Dx) + \int_{]\tau, t]} \int_{\tilde{D}} \phi(s \wedge T^\varepsilon(x), y)\, dK(s, x) \\
&\qquad + \int_{]\tau, t]} \int_{\tilde{D}} \psi(s, y) B^\varepsilon(s, x) G_s(dx)\, ds,
\end{aligned}$$

where the second equality follows from Lemma 4.5, Theorem 4.6 and the fact (which is immediate from the definition of G) that $\int_{\tilde{D}}\phi(\tau, y)G_t(dx) = \int_{\tilde{D}}\phi(\tau, y)G_\tau(dx) + \int_{]\tau,t]}\int_{\tilde{D}}\phi(\tau, y)\,dK(s, x)$.

As we recalled above, $\mathbb{P}$-a.s. for all $t \ge \tau$, we have, for G_t-a.e. $x \in \tilde{D}$, that the set $\{s: n(s) \ne n(s-)\}$ is finite, and so $T^\varepsilon(x, \cdot) = T(x, \cdot)$ and $B^\varepsilon(t, x) = B(t, x)$ when ε is sufficiently small. Thus

$$\begin{aligned}
&\int_{\tilde{D}} \phi(t \wedge T(x), y) G_t(dx) \\
&\quad = \int_{\tilde{D}} \phi(\tau, y) G_\tau(dx) + \int_{]\tau,t]} \int_{\tilde{D}} \phi(s \wedge T(x), y)\, dK(s, x) \\
&\qquad + \int_{]\tau,t]} \int_{\tilde{D}} \psi(s, y) B(s, x) G_s(dx)\, ds \\
&\quad = \int_{D} \phi(\tau, y)\nu(dy) + \int_{]\tau,t]} \int_{\tilde{D}} \phi(s \wedge T(x), y)\, dK(s, x) \\
&\qquad + \int_{]\tau,t]} \int_{D} \psi(s, y) H_s^\beta(dy)\, ds.
\end{aligned} \tag{5.2}$$

Now consider the second term on the right-hand side of (5.1). We have

$$\begin{aligned}
&\int_{\tilde{D}} \int_{]\tau,t]\times I} C(s, z, y) B(s, x) \phi(s, y) n(ds, dz) G_t(dx) \\
&\quad = \int_{\tilde{D}} \int_{]\tau,t]\times I} C(s, z, y) B(s, x) \phi(s-, y) n(ds, dz) G_t(dx) \\
&\quad = \int_{\tilde{D}} \int_{]\tau,t]\times I} C(s, z, y) B(s, x) \phi(s-, y) \check{n}(ds, dz) G_t(dx) \\
&\qquad + \int_{\tilde{D}} \int_{]\tau,t]\times I} C(s, z, y) B(s, x) \phi(s-, y)\, ds \otimes dz\, G_t(dx) \\
&\quad = \int_{]\tau,t]} \int_{\tilde{D}} \int_{]\tau,s]\times I} C(u, z, y) B(u, x) \phi(u-, y) \check{n}(du, dz)\, dK(s, x) \\
&\qquad + \int_{\tilde{D}} \int_{]\tau,t]\times I} C(s, z, y) B(s, x) \phi(s-, y)\, ds \otimes dz\, G_t(dx) \\
&\quad = \int_{]\tau,t]} \int_{\tilde{D}} \int_{]\tau,s]\times I} C(u, z, y) B(u, x) \phi(u-, y) n(du, dz)\, dK(s, x) \\
&\qquad - \int_{]\tau,t]} \int_{\tilde{D}} \int_{]\tau,s]} \beta(u, y) B(u, x) \phi(u-, y)\, du\, dK(s, x) \\
&\qquad + \int_{\tilde{D}} \int_{]\tau,t]} \beta(s, y) B(s, x) \phi(s-, x)\, ds\, G_t(dx) \\
&\quad = \int_{]\tau,t]} \int_{\tilde{D}} \int_{]\tau,s]\times I} C(u, z, y) B(u, x) \phi(u-, y) n(du, dz)\, dK(s, x)
\end{aligned} \tag{5.3}$$

$$
\begin{aligned}
&-\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]}\beta(u,y)B(u,x)\phi(u-,y)\,du\,dK(s,x)\\
&+\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]}\beta(u,y)B(u,x)\phi(u-,y)\,du\,dK(s,x)\\
&+\int_{]\tau,\,t]}\int_{\tilde D}\beta(s,y)B(s,x)\phi(s-,y)G_s(dx)\,ds\\
=&\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]\times I}C(u,z,y)B(u,x)\phi(u-,y)n(du,dz)\,dK(s,x)\\
&+\int_{]\tau,\,t]}\int_{\tilde D}\beta(s,y)B(s,x)\phi(s-,y)G_s(dx)\,ds\\
=&\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]\times I}C(u,z,y)B(u,x)\phi(u,y)n(du,dz)\,dK(s,x)\\
&+\int_{]\tau,\,t]}\int_{\tilde D}\beta(s,y)B(s,x)\phi(s,y)G_s(dx)\,ds\\
=&\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]\times I}C(u,z,y)B(u,x)\phi(u,y)n(du,dz)\,dK(s,x)\\
&+\int_{]\tau,\,t]}\int_{D}\beta(s,y)\phi(s,y)H_s^\beta(dy)\,ds,
\end{aligned}
$$

where the first, third, fifth and seventh equalities follow from Lemma 4.4, Lemma 4.3(b), Theorem 4.6 and Lemma 4.4, respectively.

Substituting (5.2) and (5.3) into (5.1), we see that

$$
\begin{aligned}
&\int_D \phi(t,y)H_t^\beta(dy)\\
&\quad=\int_D\phi(\tau,y)\nu(dy)+\int_{]\tau,\,t]}\int_{\tilde D}\phi(s\wedge T(x),y)\,dK(s,x)\\
&\qquad+\int_{]\tau,\,t]}\int_D\psi(s,y)H_s^\beta(dy)\,ds\\
&\qquad-\int_{]\tau,\,t]}\int_{\tilde D}\int_{]\tau,\,s]\times I}C(u,z,y)B(u,x)\phi(u,y)n(du,dz)\,dK(s,x)\\
&\qquad-\int_{]\tau,\,t]}\int_D\beta(s,y)\phi(s,y)H_s^\beta(dy)\,ds\\
&\quad=\int_D\phi(\tau,y)\nu(dy)+\int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)B(s,x)\,dK(s,x)\\
&\qquad+\int_{]\tau,\,t]}\int_D\psi(s,y)H_s^\beta(dy)\,ds-\int_{]\tau,\,t]}\int_D\beta(s,y)\phi(s,y)H_s^\beta(dy)\,ds.
\end{aligned}
$$

By Girsanov's theorem, under $\mathbb{P}^\beta$ there is a continuous local martingale

$(J_t^\phi)_{t\geq\tau}$ such that $J_\tau^\phi = 0$ and

$$\begin{aligned}
\int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)B(s,x)\,dK(s,x)
&= J_t^\phi + \left\langle \int_{]\tau,\cdot\,]}\int_{\tilde D}\phi(s,y)B(s,x)\,dK(s,x), \int_{]\tau,\cdot\,]}\int_{\tilde D}\beta(s,y)\,dK(s,x)\right\rangle_t\\
&= J_t^\phi + \int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)B(s,x)\beta(s,y)G_s(dx)\,ds\\
&= J_t^\phi + \int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)\beta(s,y)H_s^\beta(dy)\,ds,
\end{aligned}$$

for all $t \geq \tau$. Moreover, J^ϕ and $\int_{]\tau,\cdot]}\int_{\tilde D}\phi(s,y)B(s,x)\,dK(s,x)$ have the same quadratic variation, namely,

$$\langle J^\phi\rangle_t = \int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)^2B(s,x)G_s(dx)\,ds = \int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)^2H_s^\beta(dy)\,ds,$$

for all $t \geq \tau$.

Thus, under $\mathbb{P}^\beta$,

$$\int_D\phi(t,y)H_t^\beta(dy) = \int_D\phi(\tau,y)\nu(dy) + \int_{]\tau,\,t]}\int_D\psi(s,y)H_s^\beta(dy)\,ds + J_t^\phi,$$

for all $t \geq \tau$, where $(J_t^\phi)_{t\geq\tau}$ is a continuous local martingale such that $J_\tau^\phi = 0$ and

$$\langle J^\phi\rangle_t = \int_{]\tau,\,t]}\int_{\tilde D}\phi(s,y)^2H_s^\beta(dy)\,ds,$$

for all $t \geq \tau$. The theorem now follows from the remark we made after defining the historical process in Section 2 that it suffices to check that the appropriate processes are local martingales (rather than it being necessary to check that they are martingales). □

6. Proof of Theorem 2.4. We begin with a "finite-dimensional" special case of the result.

DEFINITION. A *cylinder function* is a function F mapping $C([\tau,\infty[, M_F(D))$ into $\mathbb{R}$ such that $F(h) = f(h_{t_1},\dots,h_{t_k})$ for some continuous function f: $(M_F(D))^k \to \mathbb{R}$, called the *representing function*, and some finite set of times $\tau < t_1 < \cdots < t_k$, called the *base*.

LEMMA 6.1. *Suppose that F is a cylinder function with representing function f and base $\tau < t_1 < \cdots < t_k$. Suppose further that for some constant c we have* $|f(\xi_1+\zeta_1,\dots,\xi_k+\zeta_k) - f(\xi_1,\dots,\xi_k)| \leq c\sum_j\zeta_j(D)$, *for all*

$\xi_1, \dots, \xi_k, \zeta_1, \dots, \zeta_k \in M_F(D)$. *If* β *is a bounded* $(\mathscr{H}_t \times \mathscr{D}_t)_{t \geq \tau}$*-predictable function, then, for all* $\theta > \tau$,

$$\mathbb{P}\left[F(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right]$$

$$= \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D (\mathscr{J}_{s;y}F)\beta(s,y)H_s(dy)\,ds\right]. \tag{6.1}$$

PROOF. Let us check that $J_{s;y}F(h)$ is defined for all s, y and h and bounded, so that $\mathscr{J}_{s;y}F$ is defined and, by Lemma 2.3, bounded. If $s \in [t_l, t_{l+1}]$ for $0 \leq l \leq (k-1)$ (where we adopt the convention that $t_0 = \tau$), then we have

$$\begin{aligned}
|J_{s;y}F(h)| &\leq \int |F(h+h^*) - F(h)| \mathbb{Q}^{s,y^{s-}}(dh^*) \\
&\leq c \sum_{j=l+1}^{k} \int h^*_{t_j}(D)\mathbb{Q}^{s,y^{s-}}(dh^*) \\
&= c \sum_{j=l+1}^{k} \int \xi(D) R_{s,t_j;y^{s-}}(d\xi) \\
&= c(k-l),
\end{aligned}$$

where the least equality follows from Lemma 3.4 of Dawson and Perkins (1991). If $s \geq t_k$, then $J_{s;y}F(h) = 0$.

We may suppose without loss of generality that $\theta \geq t_k$, because the case of $\theta < t_k$ can be deduced from this by considering β with $\beta(s,y) = 0$, for $s \geq \theta$. In that case, we may replace θ by t_k on the left-hand side of (6.1), because $F(H)$ is $\mathscr{H}_{t_k}$-measurable and the stochastic integral is a martingale. Moreover, because $J_{s;y}F(h) = J_{s;y}F(h)\mathbf{1}_{]\tau,t_k]}(s)$, we have by Lemma 2.3(a) that $\mathscr{J}_{s;y}F = \mathscr{J}_{s;y}F\mathbf{1}_{]\tau,t_k]}(s)$, and so we may also replace θ by t_k on the right-hand side of (6.1). Thus, we may suppose without loss of generality that $\theta = t_k$.

Suppose to begin with that f and β have the following extra properties:

(i) f is bounded.

(ii) $0 \leq \beta(s,y,\omega) \leq 1$.

(iii) For some $0 \leq l \leq k-1$ and some $0 < \delta < t_{l+1} - t_l$, $\beta(s,y,\omega) = 0$ unless $s \in]t_l, t_l + \delta]$ (recall that $t_0 = \tau$).

(iv) For some constant c',· $|\int_{]\tau,t]}\int_D \beta(s,y)\,dM(s,y)| \leq c'$ and $\int_{]\tau,t]}\int_D \beta(s,y)^2 H_s(dy)\,ds \leq c'$, for all $t \geq \tau$.

For $\varepsilon \in]0,1]$ we have from Theorem 5.1 that

$$\mathbb{P}[F(H)] = \mathbb{P}\left[F(H^{\varepsilon\beta})\mathscr{R}_\theta^{\varepsilon\beta}\right]$$

and so

$$
\begin{aligned}
0 &= \mathbb{P}\big[F(H)\varepsilon^{-1}\big(R_\theta^{\varepsilon\beta} - 1\big)\big] \\
&\quad + \mathbb{P}\big[\big(F(H^{\varepsilon\beta}) - F(H)\big)\varepsilon^{-1}\big(\mathscr{R}_\theta^{\varepsilon\beta} - 1\big)\big] \\
&\quad + \mathbb{P}\big[\varepsilon^{-1}\big(F(H^{\varepsilon\beta}) - F(H)\big)\big].
\end{aligned}
\tag{6.2}
$$

By assumption, $\varepsilon^{-1}(\mathscr{R}_\theta^{\varepsilon\beta} - 1)$ is bounded in $\varepsilon \in]0, 1]$ and converges almost surely to $\int_{]\tau,\theta]}\int_D \beta(s, y)\, dM(s, y)$ as $\varepsilon \downarrow 0$. Therefore, as $\varepsilon \downarrow 0$, the first term on the right-hand side of (6.2) converges to $\mathbb{P}[F(H)\int_{]\tau,\theta]}\int_D \beta(s, h)\, dM(s, y)]$ [recall that $F(H)$ is bounded by assumption (i)].

Observe that $H_t = H_t^{\varepsilon\beta} + \overline{H}_t^{\varepsilon\beta}$, where, in the notation of Section 5,

$$\overline{H}_t^{\varepsilon\beta}(\Xi) = \int_{\tilde{D}} \mathbf{1}_\Xi(y)\mathbf{1}(A(t, x) > 0)G_t(dx).$$

We remarked at the beginning of the proof of Theorem 5.1 that $n(]\tau, t] \times I)$ is finite for G_t-a.e. $x \in \tilde{D}$ for all $t \geq \tau$, $\mathbb{P}$-a.s. An argument similar to the proof of Lemma 4.4 shows that $n(]\tau, t] \times \{0\}) = 0$ for G_t-a.e. $x \in \tilde{D}$ for all $t \geq \tau$, $\mathbb{P}$-a.s. Thus $\overline{H}_t^{\varepsilon\beta}(D)\downarrow 0$ for all $t \geq \tau$, $\mathbb{P}$-a.s. Our assumption that $F(h)$ is a continuous function of $(h_{t_1}, \dots, h_{t_k})$ implies that $\lim_{\varepsilon\downarrow 0} F(H^{\varepsilon\beta}) - F(H) = 0$ almost surely. Whereas $F(H^{\varepsilon\beta}) - F(H)$ is bounded in $\varepsilon \in]0, 1]$ by assumption (i), the second term on the right-hand side of (6.2) converges to 0 as $\varepsilon \downarrow 0$.

Consider the third term on the right-hand side of (6.2).

Suppose for the moment that Y satisfies the nondegeneracy condition (3.18) of Dawson and Perkins (1991). That is, we assume that the paths of two independent copies of Y have zero probability of coinciding over any open time interval.

For $\tau \leq s < t$ define a measure $H_{s,t}^*$ on D^s by $H_{s,t}^*(\Xi) = H_t(\{y\colon y^s \in \Xi\})$. Let us recall some facts that are set out in Proposition 3.5 of Dawson and Perkins (1991). The measure $H_{s,t}^*$ is atomic with a finite set of atoms. Denote the locations of these atoms by $\mathscr{Y}_{s,t} \subset D^s$. For every $y \in D^s$ we have $R_{s,t;y}(M_F(D)) = r_{s,t}$. Conditional on $\mathscr{Y}_{s,t}$, H_t is the sum of independent nonzero clusters with laws $r_{s,t}^{-1}R_{s,t;y}$, one for each atom location $y \in \mathscr{Y}_{s,t}$. Conditional on $\mathscr{H}_s$, $\mathscr{Y}_{s,t}$ is a Poisson point process with intensity $r_{s,t}H_s$. For $\tau \leq r \leq s$ we have $\mathscr{Y}_{r,t} = (\mathscr{Y}_{s,t})^r$.

Similar comments apply to the measure $G_{s,t}^*$ on $\tilde{D}^s$ given by $G_{s,t}^*(\tilde{\Xi}) = G_t(\{x\colon x^s \in \tilde{\Xi}\})$.

Let $\mathscr{X}_{s,t} \in \tilde{D}^s$ denote the set of atom locations of $G_{s,t}^*$. From Proposition 3.7 of Dawson and Perkins (1991) it follows that each atom location in $\mathscr{Y}_{s,t}$ is the E-valued path component of one and only one atom location in $\mathscr{X}_{s,t}$. Moreover, conditional on $\mathscr{H}$, the joint law of the corresponding I-valued path components of the atom locations $\mathscr{X}_{s,t}$ is that obtained by running a branching particle system with branching structure dictated by that of the E-valued path components of $\mathscr{X}_{s,t}$ and spatial motion that of the "path process" constructed from the Markov jump process on I introduced in Section 3, with all of the particles being the constant zero path at time τ (recall our convention on μ).

Given these observations, we can describe the conditional law of $(H_{t_1}^{\varepsilon\beta}, \dots, H_{t_k}^{\varepsilon\beta})$ given $\mathscr{H}$ as follows. For each notation we will put $u = t_l$, $v = t_l + \delta$ and $w = t_{l+1}$. For $s \in]u, v]$ let η_s^β be the (random) finite measure on D that places mass $\beta(s, y)$ at each point y in $(\mathscr{Y}_{v,w})^s = \mathscr{Y}_{s,w}$, and let λ^β be the (random) measure on $]\tau, \infty[\times D$ given by

$$\int_{]\tau,\infty[\times D} \phi(s, y) \lambda^\beta(ds, dy) = \int_{]u, v] \times D} \phi(s, y) \eta_s^\beta(dy)\, ds.$$

Thus, λ^β is concentrated on $]u, v] \times D$. Put

$$\mathscr{X}^{\varepsilon\beta} = \{(s, y^s): x \in \mathscr{X}_{v,w}, n(s) \neq n(s-), n(s) < \varepsilon\beta(s, y)\}.$$

Conditional on $\mathscr{H}$, the random set $\mathscr{X}^{\varepsilon\beta}$ is a Poisson point process with intensity $\varepsilon\lambda^\beta$. We have

$$H_{t_j}^{\varepsilon\beta}(\Xi) = \begin{cases} H_{t_j}(\Xi), & \text{if } j \le l, \\ H_{t_j}(\{y \in \Xi: (s, y^s) \notin \mathscr{X}^{\varepsilon\beta}, \forall\, s \in]u, v]\}), & \text{if } j > l. \end{cases}$$

For $s \in]u, v]$ and $y \in D^s$, let $\kappa_{s;y}$ be the mapping of the set of functions $\{h: [\tau, \infty[\to M_F(D)\}$ into itself that is defined by

$$(\kappa_{s;y} h)_t(\Xi) = \begin{cases} h_t(\Xi), & \text{if } t < s, \\ h_t(\{y' \in \Xi: (y')^s \neq y\}), & \text{if } t \ge s. \end{cases}$$

We have from the above that

$$\begin{aligned} &\mathbb{P}[\varepsilon^{-1}(F(H^{\varepsilon\beta}) - F(H)) | \mathscr{H}] \\ &\quad = \exp(-\varepsilon\lambda^\beta(]\tau, \infty[\times D)) \\ &\qquad \times \int_{]\tau,\infty[\times D} (F(\kappa_{s;y} H) - F(H)) \lambda^\beta(ds, dy) + \rho(F, \varepsilon, \beta), \end{aligned}$$

where the remainder satisfies $|\rho(F, \varepsilon, \beta)| \le C\|F\|_\infty \varepsilon\lambda^\beta(]\tau, \infty[\times D)^2$, for some constant C. From the above description of the conditional law of $\mathscr{Y}_{v,w}$ given $\mathscr{H}_v$ it is clear that the cardinality of $\mathscr{X}_{v,w}$ (= the cardinality of $\mathscr{Y}_{v,w}$) has finite moments of all orders and so the same is true of the total mass of λ^β. Thus the limit as $\varepsilon \downarrow 0$ of the third term on the right-hand side of (6.2) is

$$\begin{aligned} &\mathbb{P}\left[\int_{]\tau,\infty[\times D} (F(\kappa_{s;y} H) - F(H)) \lambda^\beta(ds, dy)\right] \\ &\quad = \int_{]u, v]} \mathbb{P}\left[\int_D (F(\kappa_{s;y} H) - F(H)) \eta_s^\beta(dy)\right] ds \\ &\quad = \int_{]u, v]} \mathbb{P}\left[\sum \{(F(\kappa_{s;y} H) - F(H)) \beta(s; y): y \in \mathscr{Y}_{s,w}\}\right] ds. \end{aligned}$$

Before moving on, let us introduce some notation. For $s \in]u, v]$, let $\bar{\mathscr{H}}_s = \mathscr{H}_s \vee \sigma(\mathscr{Y}_{s,w})$. Put

$$Q_{s;y}(d\xi_{l+1}, \dots, d\xi_k) = R_{s, t_{l+1}; y}(d\xi_{l+1}) \prod_{j=l+1}^{k-1} P_{t_j, t_{j+1}}(\xi_j, d\xi_{j+1}).$$

Note that

$$\int g(\xi_{l+1},\dots,d\xi_k)Q_{s;y}(d\xi_{l+1},\dots,d\xi_k)$$

$$= \int g(h_{t_{l+1}},\dots,h_{t_k})\mathbf{1}(h_{t_{l+1}} \neq 0)\mathbb{Q}^{s;y}(dh),$$

for any bounded measurable function $g\colon (M_F(D))^{k-l} \to \mathbb{R}$. Write $\hat{Q}_{s;y} = r_{s,t_{l+1}}^{-1} Q_{s;y}$ for the normalization of $Q_{s;y}$. Given a set $\mathscr{Z} \subset D^s$ put

$$\Phi_{s;\mathscr{Z}} = \int f\left(H_{t_1},\dots,H_{t_l},\sum_{i=1}^m \xi_{l+1}^i,\dots,\sum_{i=1}^m \xi_k^i\right)\bigotimes_{i=1}^m \hat{Q}_{s;y_i}\left(d\xi_{l+1}^i,\dots,d\xi_k^i\right),$$

when $\mathscr{Z} = \{y_1,\dots,y_m\}$ is nonempty, and put

$$\Phi_{s;\varnothing} = f\left(H_{t_1},\dots,H_{t_l},0,\dots,0\right).$$

By the description of $H_{t_{l+1}}$ in terms of clusters growing from the points of $\mathscr{Y}_{s,t_{l+1}}$, the Markov property of H, the infinite divisibility of the probability measure $\mathbb{P}^{s;\zeta}$, $s \geq \tau$, $\zeta \in M_F(D)^s$ and the Poisson nature of $\mathscr{Y}_{s,t_{l+1}}$ given $\mathscr{H}_s$, we have

$$\begin{aligned}
&\mathbb{P}\Big[\sum\big\{\big(F(\kappa_{s;y}H) - F(H)\big)\beta(s,y)\colon y \in \mathscr{Y}_{s,w}\big\}\Big]\\
&\quad = \mathbb{P}\Big[\mathbb{P}\Big[\sum\big\{\mathbb{P}\big[\big(F(\kappa_{s;y}H) - F(H)\big)\beta(s,y)|\bar{\mathscr{H}}_s\big]\colon y \in \mathscr{Y}_{s,w}\big\}\Big|\mathscr{H}_s\Big]\Big]\\
&\quad = \mathbb{P}\Big[\mathbb{P}\Big[\sum\big\{\big(\Phi_{s;(\mathscr{Y}_{s,w}\setminus\{y\})} - \Phi_{s;\mathscr{Y}_{s,w}}\big)\beta(s,y)\colon y \in \mathscr{Y}_{s,w}\big\}\Big|\mathscr{H}_s\Big]\Big]\\
&\quad = \mathbb{P}\Bigg[\exp\big(-r_{s,t_{l+1}}H_s(D)\big)\sum_{m=1}^\infty \frac{1}{m!}m\int_{(D)^m}\big(\Phi_{s;\{y_1,\dots,y_{m-1}\}} - \Phi_{s;\{y_1,\dots,y_m\}}\big)\\
&\qquad\qquad \times\beta(s,y_m)r_{s,t_{l+1}}^m H_s^{\otimes m}(dy_1,\dots,dy_m)\Bigg]\\
&\quad = \mathbb{P}\Bigg[\int_D\Bigg(\exp\big(-r_{s,t_{l+1}}H_s(D)\big)\\
&\qquad\qquad \times\sum_{m=0}^\infty \frac{1}{m!}\int_{(D)^m}\big(\Phi_{s;\{y_1,\dots,y_m\}} - \Phi_{s;\{y_1,\dots,y_m,y\}}\big)\\
&\qquad\qquad \times r_{s,t_{l+1}}^m H_s^{\otimes m}(dy_1,\dots,dy_m)\Bigg)\beta(s,y)r_{s,t_{l+1}}H_s(dy)\Bigg]\\
&\quad = \mathbb{P}\Bigg[\int_D\int_{M_F(D)^{k-l}}\mathbb{P}\Big[\big(f(H_{t_1},\dots,H_{t_k})\\
&\qquad\qquad - f\big(H_{t_1},\dots,H_{t_l},H_{t_{l+1}} + \xi_{l+1},\dots,H_{t_k} + \xi_k\big)\big)\Big|\mathscr{H}_s\Big]\\
&\qquad\qquad \times r_{s,t_{l+1}}\hat{Q}_{s;y}(d\xi_{l+1},\dots,d\xi_k)\beta(s,y)H_s(dy)\Bigg].
\end{aligned}$$

Note that the last equality still holds if we replace $\hat{Q}_{s;y}$ by $\hat{Q}_{s;y^{s-}}$, since $y = y^{s-}$ for H_s-a.e. y, $\mathbb{P}$-a.s., by (2.1) and the fact that Y has no fixed discontinuities under $P^{\tau,\nu}$. It is straightforward to check that, for $s \in]u, v]$,

$$\int_{M_F(D)^{k-l}} \mathbb{P}\Big[\big(f(H_{t_1},\dots,H_{t_k}) - f(H_{t_1},\dots,H_{t_l},H_{t_{l+1}}+\xi_{l+1},\dots,H_{t_k}+\xi_k)\big)\big|\mathscr{H}_s\Big]$$
$$\times r_{s,t_{l+1}}\hat{Q}_{s;y^{s-}}(d\xi_{l+1},\dots,d\xi_k)$$
$$= -\mathscr{I}_{s;y}F,$$

$\overline{\mathbb{P}}_s$-a.e. and so

$$\mathbb{P}\left[\int_{]\tau,\infty[\times D}\big(F(\kappa_{s;y}H) - F(H)\big)\lambda^{\beta}(ds,dy)\right]$$
$$= -\mathbb{P}\left[\int_{]\tau,\theta]}\int_D \mathscr{I}_{s;y}F\beta(s,y)H_s(dy)\,ds\right].$$

Thus, if we let $\varepsilon \downarrow 0$ in (6.2) we see that (6.1) holds for the special F and β we have been considering.

Suppose now that we remove condition (iv) and replace condition (iii) by the following condition:

(iii′) $\beta(s, y, \omega) = 0$ unless $s \in]t_l, t_{l+1}]$.

In addition, leave conditions (i) and (ii) in place.

Given any β satisfying (ii) and (iii′), we can find an increasing sequence of functions $\{\beta_m\}_{m=1}^{\infty}$, such that β_m satisfies conditions (ii)–(iv) and $\beta_m(s, y, \omega)\uparrow \beta(s, y, \omega)$, for all $(s, y, \omega) \notin \{t_{l+1}\} \times D \times \Omega$. Recalling that $\mathscr{I}_{s;y}F$ is bounded, it follows from bounded convergence that

$$\mathbb{P}\left[\int_{]\tau,\theta]}\int_D (\mathscr{I}_{s;y}F)\beta(s,y)H_s(dy)\,ds\right]$$
$$= \lim_{m\to\infty}\mathbb{P}\left[\int_{]\tau,\theta]}\int_D (\mathscr{I}_{s;y}F)\beta_m(s,Y)H_s(dy)\,ds\right].$$

Also, since F is bounded by assumption (i), we have

$$\lim_{m\to\infty}\mathbb{P}\left[\left\{F(H)\int_{]\tau,\theta]}\int_D \beta_m(s,y)\,dM(s,y)\right.\right.$$
$$\left.\left. -F(H)\int_{]\tau,\theta]}\int_D \beta(s,y)\,dM(s,y)\right\}^2\right]$$
$$\le \lim_{m\to\infty}\|F\|_\infty^2\mathbb{P}\left[\int_{]\tau,\theta]}\int_D \{\beta_m(s,y) - \beta(s,y)\}^2 H_s(dy)\,ds\right] = 0.$$

Thus (6.1) holds under (i), (ii) and (iii′).

Suppose now that we remove condition (i). We have that $(F \wedge a) \vee (-a)$ satisfies the hypotheses of the lemma as well as assumptions (i), (ii) and (iii′).

An inspection of the way we bounded $|J_{s;\,y}F|$ at the beginning of the proof shows that $|J_{s;\,y}((F \wedge a) \vee (-a))(H)|$ is uniformly bounded in a. By Lemma 2.3, $|\mathscr{J}_{s;\,y}((F \wedge a) \vee (-a))|$ is uniformly bounded in a and $\mathbb{P}$-a.s., for all $s > \tau$, we have $\lim_{a \to \infty} \mathscr{J}_{s;\,y}((F \wedge a) \vee (-a)) = \mathscr{J}_{s;\,y}F$ for H_s-a.e. $y \in D$. By bounded convergence,

$$\mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D (\mathscr{J}_{s;\,y}F)\beta(s,y)H_s(dy)\,ds\right]$$
$$= \lim_{a \to \infty} \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D (\mathscr{J}_{s;\,y}(F \wedge a) \vee (-a))\beta_m(s,Y)H_s(dy)\,ds\right].$$

Also, by assumption,

$$\begin{aligned}|((F \wedge a) \vee (-a))(H)| &\le |F(H)| \le |F(0)| + |F(H) - F(0)| \\ &\le |F(0)| + c\sum_j H_{t_j}(D),\end{aligned}$$

where by 0 we mean the element of $C([\tau,\infty[, M_F(D))$ that is constant at the zero measure. Note that the rightmost random variable has finite moments of all orders. Therefore, by Cauchy–Schwarz and dominated convergence,

$$\begin{aligned}&\lim_{a \to \infty} \mathbb{P}\Bigg[\Bigg|((F \wedge a) \vee (-a))(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y) \\ &\qquad\qquad - F(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\Bigg|\Bigg] \\ &\le \lim_{a \to \infty} \mathbb{P}\left[\{((F \wedge a) \vee (-a))(H) - F(H)\}^2\right]^{1/2} \\ &\quad \times \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D \beta(s,y)^2 H_s(dy)\,ds\right]^{1/2} \\ &= 0.\end{aligned}$$

Thus (6.1) holds under (ii) and (iii′).

Linearity allows us to remove conditions (ii) and (iii′) (recall that we are assuming, as we may, that $\theta = t_k$).

Finally, we need to remove our nondegeneracy assumption that the paths of two independent copies of Y have zero probability of coinciding over any open time interval. If we form the Cartesian product of an arbitrary Y with a process that does satisfy the assumption (such as a Brownian motion), then the resulting process also satisfies the assumption. Thus the conclusion of the theorem holds for the historical process constructed over the Cartesian product. The historical process over Y and its related orthogonal martingale measure are obtained from the corresponding objects for the Cartesian product by projection, and it is straightforward to conclude that the theorem must also hold for the historical process over Y. We leave the details to the reader. □

We are now ready to proceed with the proof of Theorem 2.4.

For each $l \in \mathbb{N}$ choose a partition $\tau = t_0^l < t_1^l < \cdots < t_{k(l)}^l < \infty$ such that $\lim_{l\to\infty} \sup_i t_{i+1}^l - t_i^l = 0$ and $\lim_{l\to\infty} t_{k(l)}^l = \infty$. Define a map σ_l: $C([\tau,\infty[, M_F(D)) \to C([\tau,\infty[, M_F(D))$ by

$$\sigma_l(h)_t = \begin{cases} \dfrac{t_{i+1}^l - t}{t_{i+1}^l - t_i^l} h_{t_i^l} + \dfrac{t - t_i^l}{t_{i+1}^l - t_i^l} h_{t_{i+1}^l}, & \text{if } t \in \left[t_i^l, t_{i+1}^l\right[, \\ h_{t_{k(l)}^l}, & \text{if } t \geq t_{k(l)}^l. \end{cases}$$

Define a function L: $C([\tau,\infty[, M_F(D)) \to \mathbb{R}$ by $L(h) = \int_{[\tau,\infty[} h_t(D) m(dt)$. By assumption, $|(F \circ \sigma_l)(h + h^*) - (F \circ \sigma_l)(h)| \leq (L \circ \sigma_l)(h^*)$ and $|F(h + h^*) - F(h)| \leq L(h^*)$ for all h, h^*. Observe that $L \circ \sigma_l(h) = \int h_t(D) m_l(dt)$, where the finite measure m_l concentrates its mass on $t_0^l, t_1^l, \ldots, t_{k(l)}^l$ and is given by

$$m_l(\{t_0^l\}) = \int_{[t_0^l, t_1^l[} \frac{t_1^l - t}{t_1^l - t_0^l} m(dt),$$

$$m_l(\{t_i^l\}) = \int_{[t_{i-1}^l, t_i^l[} \frac{t - t_{i-1}^l}{t_i^l - t_{i-1}^l} m(dt) + \int_{[t_i^l, t_{i+1}^l[} \frac{t_{i+1}^l - t}{t_{i+1}^l - t_i^l} m(dt), \qquad 1 \leq i < k(l),$$

$$m_l\left(\left\{t_{k(l)}^l\right\}\right) = \int_{[t_{k(l)-1}^l, t_{k(l)}^l[} \frac{t - t_{k(l)-1}^l}{t_{k(l)}^l - t_{k(l)-1}^l} m(dt) + m\left(\left[t_{k(l)}^l, \infty\right[\right).$$

For $s < t$, we have from Lemma 3.4 of Dawson and Perkins (1991) that

$$\int h_t^*(D) \mathbb{Q}^{s;y}(dh^*) = \int \xi(D) R_{s,t;y}(d\xi) = 1.$$

So, by Fubini,

$$\int |(F \circ \sigma_l)(h + h^*) - (F \circ \sigma_l)(h)| \mathbb{Q}^{s;y}(dh^*) \leq \int (L \circ \sigma_l)(h^*) \mathbb{Q}^{s;y}(dh^*) = m_l(]s,\infty[)$$

and

$$\int |F(h + h^*) - F(h)| \mathbb{Q}^{s;y}(dh^*) \leq \int L(h^*) \mathbb{Q}^{s;y}(dh^*) = m(]s,\infty[),$$

for all s, y. In particular, $J_{s;y}(F \circ \sigma_l)(h)$ and $J_{s;y}F(h)$ are defined for all s, y, h and bounded, so that $\mathscr{J}_{s;y}(F \circ \sigma_l)$ and $\mathscr{J}_{s;y}F$ are also defined and bounded.

Now $\lim_{l\to\infty} m_l(]s,\infty[) = m(]s,\infty[)$ for all but countably many s, and so $\lim_{l\to\infty}(L\circ\sigma_l)(h) = L(h)$ (recall that m has compact support). Thus,

$$\begin{aligned}
&\mathbb{P}\left[\int_{]\tau,\theta]}\int_D\int \lim_{l\to\infty}(L\circ\sigma_l)(h^*)\mathbb{Q}^{s;y^{s-}}(dh^*)H_s(dy)\,ds\right]\\
&\quad= \mathbb{P}\left[\int_{]\tau,\theta]}\int_D\int L(h^*)\mathbb{Q}^{s;y^{s-}}(dh^*)H_s(dy)\,ds\right]\\
&\quad= \mathbb{P}\left[\int_{]\tau,\theta]}\int_D m(]s,\infty[)H_s(dy)\,ds\right]\\
&\quad= \mathbb{P}\left[\int_{]\tau,\theta]}\int_D \lim_{l\to\infty} m_l(]s,\infty[)H_s(dy)\,ds\right]\\
&\quad= \lim_{l\to\infty}\mathbb{P}\left[\int_{]\tau,\theta]}\int_D m_l(]s,\infty[)H_s(dy)\,ds\right]\\
&\quad= \lim_{l\to\infty}\mathbb{P}\left[\int_{]\tau,\theta]}\int_D\int (L\circ\sigma_l)(h^*)\mathbb{Q}^{s;y^{s-}}(dh^*)H_s(dy)\,ds\right],
\end{aligned}$$

where the fourth equality follows from the bounded convergence theorem.

By the continuity of F, $\lim_{l\to\infty}(F\circ\sigma_l)(h) = F(h)$, for all h. Thus, by a variant of the dominated convergence theorem [cf. Proposition 11.18 of Royden (1968)], we have

$$\begin{aligned}
&\mathbb{P}\left[\int_{]\tau,\theta]}\int_D (\mathscr{J}_{s;y}F)\beta(s,y)H_s(dy)\,ds\right]\\
&\quad= \mathbb{P}\left[\int_{]\tau,\theta]}\int_D \big(J_{s;y}F(H)\big)\beta(s,y)H_s(dy)\,ds\right]\\
&\quad= \mathbb{P}\left[\int_{]\tau,\theta]}\int_D\int F(H+h^*) - F(H)\mathbb{Q}^{s;y^{s-}}(dh^*)\beta(s,y)H_s(dy)\,ds\right]\\
&\quad= \lim_{l\to\infty}\mathbb{P}\Bigg[\int_{]\tau,\theta]}\int_D\int (F\circ\sigma_l)(H+h^*)\\
&\qquad\qquad -(F\circ\sigma_l)(H)\mathbb{Q}^{s;y^{s-}}(dh^*)\beta(s,y)H_s(dy)\,ds\Bigg]\\
&\quad= \lim_{l\to\infty}\mathbb{P}\left[\int_{]\tau,\theta]}\int_D \big(J_{s;y}(F\circ\sigma_l)(H)\big)\beta(s,y)H_s(dy)\,ds\right]\\
&\quad= \lim_{l\to\infty}\mathbb{P}\left[\int_{]\tau,\theta]}\int_D \big(\mathscr{J}_{s;y}(F\circ\sigma_l)\big)\beta(s,y)H_s(dy)\,ds\right].
\end{aligned}$$

Recall that, for each $p \ge 1$, $\mathbb{P}[H_t(D)^p]$ is uniformly bounded on compact intervals. One consequence of this observation is that, for each $p \ge 1$, $\mathbb{P}[(L\circ\sigma_l)(H)^p]$ is bounded in l (again recall that m had compact support). Another consequence is that $\int_{]\tau,\theta]}\int_D \beta(s,y)\,dM(s,y)$ has moments of all orders. Therefore, $\mathbb{P}[((L\circ\sigma_l)(H)\int_{]\tau,\theta]}\int_D \beta(s,y)\,dM(s,y))^2]$ is bounded in l. As $|(F\circ\sigma_l)(H)| \le |F(0)| + (L\circ\sigma_l)(H)$ and $\lim_{l\to\infty}(F\circ\sigma_l)(H) = F(H)$, a uniform integrability argument shows that

$$\lim_{l\to\infty} \mathbb{P}\left[(F\circ\sigma_l)(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right]$$
$$= \mathbb{P}\left[F(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right].$$

The proof of the theorem is completed once we observe that $F\circ\sigma_l$ satisfies the conditions of Lemma 6.1, and so

$$\mathbb{P}\left[(F\circ\sigma_l)(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right]$$
$$= \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D \mathscr{I}_{s;y}(F\circ\sigma_l)\beta(s,y)H_s(dy)\,ds\right]. \qquad \square$$

7. Proof of Theorem 2.5. Recall that $\mathbb{P}[H_t(D)^2]$ is uniformly bounded on compact intervals and so, by assumption, $\mathbb{P}[F(H)^2]$ is finite. Therefore, by Theorem 4.7, there certainly exists a $(\mathscr{D}_t\times\mathscr{H}_t)_{t\geq\tau}$-predictable process ρ such that

$$\mathbb{P}\left[\int_{]\tau,\infty[}\int_D \rho(s,y)^2 H_s(dy)\,ds\right] < \infty$$

and

$$F(H) = \mathbb{P}[F(H)] + \int_{]\tau,\infty[}\int_D \rho(s,y)\,dM(s,y).$$

From Theorem 2.4 we have for any bounded $(\mathscr{D}_t\times\mathscr{H}_t)_{t\geq\tau}$-predictable function β and for all $\theta>\tau$ that

$$\mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D \rho(s,y)\beta(s,y)H_s(dy)\,ds\right]$$
$$= \mathbb{P}\left[F(H)\int_{]\tau,\,\theta]}\int_D \beta(s,y)\,dM(s,y)\right]$$
$$= \mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D \mathscr{I}_{s;y}F\beta(s,y)H_s(dy)\,ds\right].$$

A standard Hilbert space argument now shows that

$$\mathbb{P}\left[\int_{]\tau,\,\theta]}\int_D \left(\rho(s,y)-\mathscr{I}_{s;y}F\right)^2 H_s(dy)\,ds\right] = 0.$$

Thus, if we choose θ sufficiently large that the support of m is contained in $[\tau,\theta]$, then we have

$$\int_{]\tau,\,\theta]}\int_D \rho(s,y)\,dM(s,y) = \int_{]\tau,\,\theta]}\int_D \mathscr{I}_{s;y}F\,dM(s,y)$$
$$= \int_{]\tau,\,\infty]}\int_D \mathscr{I}_{s;y}F\,dM(s,y),$$

because $\mathscr{I}_{s;y}F=0$, for H_s-a.e. y, when $s>\theta$. The result now follows. $\square$

8. Proof of Theorem 2.7. Note that $\sigma\{H_t: t \geq \tau\}$ is generated by the random variables $\exp(-\sum_{j=1}^n H_{u_j}(g_j))$, where $n \in \mathbb{N}$, $u_i \in [\tau, \infty[$ and g_i is nonnegative, bounded and continuous. A monotone class argument shows that linear combinations of such random variables are dense in $L^0(\Omega, \sigma\{H_t: t \geq \tau\}, \mathbb{P})$. If we expand out the exponential as a Taylor series in n variables, then we see that the theorem will be a consequence of Lemma 8.1 below.

NOTATION. For $k \in \mathbb{N}$ and $t > \tau$, we will denote by $\mathscr{I}_k(t)$ the set of all random variables of the form $I_k(\phi_1, \dots, \phi_k; t)$, where $\phi_i \in b(\mathscr{B}(]\tau, \infty[)^{k-i+1} \times \mathscr{D})$.

LEMMA 8.1. *Given any $t_1, \dots, t_m \in [\tau, \theta]$, $\theta > \tau$ and $f_1, \dots, f_m \in C(D, \mathbb{R})$ nonnegative and bounded, there is a sequence of random variables in the linear span of $\mathbb{R} \cup \bigcup_{j=1}^m \mathscr{I}_j(\theta)$ that converges to $\prod_{i=1}^m H_{t_i}(f_i)$ in $L^p(\mathbb{P})$, for all $1 \leq p < \infty$.*

PROOF. We will proceed by induction on m. When $m = 1$ we have

$$H_t(f) = \nu(P_{\tau,t} f) + \int_{]\tau, t]} \int_D P_{s,t} f(y)\, dM(s, y).$$

(As we mentioned in Section 1, this result is not that difficult to establish directly from the martingale problem. It follows immediately from Theorem 2.5.)

Suppose now that the inductive hypothesis holds for $1, \dots, m-1$, where $m \geq 2$. Put $F(h) = \prod_{i=1}^m h_{t_i}(f_i)$. Observe that $\int (F(h + h^*) - F(h)) \mathbb{Q}^{s; y^{s-}}(dh^*)$ is a finite sum of the terms $\phi_C(s; y) \prod_{j \notin C} h_{t_j}(f_j)$, where $\phi_C(s; y) = \int \prod_{i \in C} h^*_{t_i}(f_i) \mathbb{Q}^{s; y^{s-}}(dh^*)$ and C ranges over the nonempty subets of $\{1, \dots, m\}$. Consider one such term. For any $p \geq 1$ and $s < t$ we have

$$\int h_t^*(D)^p \mathbb{Q}^{s; y^{s-}}(dh^*) = \int \xi(D)^p R_{s,t; y^{s-}}(d\xi).$$

If we recall the description of the law of H_t under $\mathbb{P}^{s; \delta_{y^{s-}}}$ as a Poisson superposition of clusters thrown down with intensity $R_{s,t; y^{s-}}$ and apply the inequality $\sum_l a_l^p \leq (\sum_l a_l)^p$, for $a_l \geq 0$, then we find that

$$\int \xi(D)^p R_{s,t; y^{s-}}(d\xi) \leq \mathbb{P}^{s; \delta_{y^{s-}}}\left[H_t(D)^p\right].$$

For each p, the right-hand side is uniformly bounded in s, t, y satisfying $\tau < s < t \leq \theta$ and so, by Hölder's inequality, $\phi_C(s; y)$ is uniformly bounded in s, y satisfying $\tau < s \leq \theta$. Note also that $\mathbb{P}[\prod_{j \notin C} H_{t_j}(f_j)^p] < \infty$, for all $p \geq 1$.

For any $a > 0$, the function $F^{(a)}$ given by $F^{(a)}(h) = \prod_{i=1}^m (h_{t_i}(f_i) \wedge a)$ satisfies the conditions of Theorem 2.5. Let $(L_C(t))_{t \geq \tau}$ be the $(\mathscr{H}_t)_{t \geq \tau}$-predictable projection of the process that at any time takes the value $\prod_{j \notin C} H_{t_j}(f_j)$. It follows from Lemma 2.3 that $0 \leq \mathscr{I}_{s;y}(F^{(a)}) \uparrow \mathscr{I}_{s;y} F = \sum_C \phi_C(s; y) L_C(s)$ for H_s-a.e. y, $\mathbb{P}$-a.s., for each $s > \tau$. Thus

$$F(H) = \mathbb{P}[F(H)] + \int_{]\tau, \theta]} \int_D \sum_C \phi_C(s; y) L_C(s)\, dM(s, y).$$

For $k \in \mathbb{N}$ and $l = 0, 1, \dots, 2^k$, put $s_l^k = \tau + l(\theta - \tau)/2^k$. Set

$$\begin{aligned} L_C^k(s) &= \mathbb{P}\left[\prod_{j \notin C} H_{t_j}(f_j) \middle| \mathscr{H}_{s_l^k}\right] \qquad (s_l^k < s \le s_{l+1}^k) \\ &= \mathbb{P}\left[L_C(s) | \mathscr{H}_{s_l^k}\right] \qquad (s_l^k < s \le s_{l+1}^k). \end{aligned}$$

By Jensen's inequality and Hölder's inequality we see that, for each $p \ge 1$, $H_s(D)(L_C(s) - L_C^k(s))^2$ is bounded in $L^p(\mathbb{P})$ as s ranges over $]\tau, \theta]$ and k ranges over $1, 2, \dots$. Moreover, by the martingale convergence theorem, $\lim_{k \to \infty} H_s(D)(L_C(s) - L_C^k(s))^2 = 0$ in $L^p(\mathbb{P})$ for each s. Thus, by the Burkholder–Davis–Gundy inequality,

$$\lim_{k \to \infty} \int_{]\tau, \theta]} \int_D \phi_C(s; y) L_C^k(s)\, dM(s, y) = \int_{]\tau, \theta]} \int_D \phi_C(s; y) L_C(s)\, dM(s, y)$$

in $L^p(\mathbb{P})$, for all $p \ge 1$.

Now

$$\begin{aligned} &\int_{]\tau, \theta]} \int_D \phi_C(s; y) L_C^k(s)\, dM(s, y) \\ &\qquad = \sum_{l=0}^{2^k - 1} \int_{]s_l^k, s_{l+1}^k]} \int_D \phi_C(s; y)\, dM(s, y) \times L_C^k(s_l^k) \end{aligned}$$

and so, in order to establish the inductive step, it will suffice to show that each random variable $L_C^k(s_l^k)$ is the limit in $L^p(\mathbb{P})$ for all $p \ge 1$ of a sequence drawn from the linear span of $\mathbb{R} \cup \bigcup_{j=1}^{m-1} \mathscr{I}_j(s_l^k)$.

However, from the moment formulae in Dynkin (1988) and the Markov property, we see that $L_C^k(s_l^k)$ is a finite sum of terms of the form $\prod_{j \in B} H_{r_j}(e_j)$, where card $B \le$ card$(\{1, \dots, m\} \setminus C)$, each r_j is of the $t_i \wedge s_l^k$ for some $i \in \{1, \dots, m\} \setminus C$ and $e_j \in$ bp-$\mathscr{D}$. For each j we can find a uniformly bounded sequence $\{e_j^m\}$ of nonnegative continuous functions such that

$$P^{\tau, \nu}\left[|e_j^m(X^{r_j}) - e_j(X^{r_j})|\right] \to 0 \quad \text{as } m \to \infty.$$

Then $H_{r_j}(e_j^m)$ converges to $H_{r_j}(e_j)$ as $m \to \infty$ in $L^p(\mathbb{P})$, for all $p \ge 1$, and so $\prod_{j \in B} H_{r_j}(e_j^m)$ converges to $\prod_{j \in B} H_{r_j}(e_j)$ in $L^p(\mathbb{P})$, for all $p \ge 1$. We can thus apply the inductive hypothesis to complete the proof. □

Acknowledgments. We thank Martin Barlow for ongoing collaboration that has led to some of the ideas and techniques we have used here. We also thank an anonymous referee for pointing out a number of deficiencies in an earlier version of the paper.

REFERENCES

Adler, R. J. and Lewin, M. (1991). An evolution equation for the intersection local times of superprocesses. *Stochastic Analysis* (M. T. Barlow and N. H. Bingham, eds.). Cambridge Univ. Press.

Barlow, M. T., Evans, S. N. and Perkins, E. A. (1991). Collision local times and measure-valued processes. *Canad. J. Math.* **43** 897–938.

BARLOW, M. T., EVANS, S. N. and PERKINS, E. A. (1995). Historical Brownian motion with interactive killing. In preparation.

BARLOW, M. T. and PERKINS, E. A. (1994). On the filtration of historical Brownian motion. *Ann. Probab.* **22** 1273–1294.

BISMUT, J.-M. (1981). Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Z. Wahrsch. Verw. Gebiete* **56** 469–505.

CLARK, J. M. C. (1971). The representation of functionals of Brownian motion by stochastic integrals. *Ann. Math. Statist.* **41** 1281–1295.

DAVIS, M. H. A. (1980). Functionals of diffusion processes as stochastic integrals. *Math. Proc. Cambridge Philos. Soc.* **87** 157–166.

DAWSON, D. A. (1993). Measure-valued Markov processes. *Ecole d'Eté de Probabilités de Saint-Flour XXI. Lecture Notes in Math.* **1541** 1–260. Springer, Berlin.

DAWSON, D. A., ISCOE, I. and PERKINS, E. A. (1989). Super-Brownian motion: path properties and hitting probabilities. *Probab. Theory Related Fields* **83** 135–205.

DAWSON, D. A. and PERKINS, E. A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93**.

DELLACHERIE, C. and MEYER, P. A. (1978). *Probabilities and Potential. Mathematics Studies* **29**. North-Holland, Amsterdam.

DELLACHERIE, C. and MEYER, P. A. (1982). *Probabilities and Potential B. Mathematics Studies* **72**. North-Holland, Amsterdam.

DYNKIN, E. B. (1988). Representation for functionals of superprocesses by multiple stochastic integrals, with applications to self-intersection local times. Colloque Paul Lévy sur les processus stochastiques. *Astérisque* **157-158** 147–171.

DYNKIN, E. B. (1993). Superprocesses and partial differential equations. *Ann. Probab.* **21** 1185–1262.

EL-KAROUI, N. and ROELLY, S. (1991). Propriétés de martingales, explosion et représentation, de Lévy–Khintchine d'une classe de processus de branchement à valeurs mesures. *Stochastic Process. Appl.* **38** 239–266.

EVANS, S. N. (1990). Association and random measures. *Probab. Theory Related Fields* **86** 1–19.

EVANS, S. N. and PERKINS, E. A. (1991). Absolute continuity results for superprocesses with some applications. *Trans. Amer. Math. Soc.* **325** 661–681.

EVANS, S. N. and PERKINS, E. A. (1994). Measure-valued branching diffusions with singular interactions. *Canad. J. Math.* **46** 120–168.

FITZSIMMONS, P. J. (1988). Construction and regularity of measure-valued Markov branching processes. *Israel J. Math.* **64** 337–361.

HAUSSMANN, U. (1978). Functionals of Itô processes as stochastic integrals. *SIAM J. Control Optim.* **16** 252–269.

HAUSSMANN, U. (1979). On the integral representation of functionals of Itô processes. *Stochastics* **3** 17–28.

IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

ITÔ, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–169.

JACOD, J. (1977). A general theorem of representation for martingales. In *Proceedings of Symposia in Pure Mathematics* **31** 37–54. Amer. Math. Soc., Providence, RI.

JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales. Lecture Notes in Math.* **714**. Springer, Berlin.

KRONE, S. M. (1993). Local time for superdiffusions. *Ann. Probab.* **21** 1599–1623.

LE GALL, J.-F. (1993). A class of path-valued Markov processes and its applications to superprocesses. *Probab. Theory Related Fields* **95** 25–46.

MUELLER, C. and PERKINS, E. A. (1992). The compact support property for solutions to the heat equation with noise. *Probab. Theory Related Fields* **93** 325–358.

OCONE, D. (1984). Malliavin's calculus and stochastic integral representations of functionals of diffusion processes. *Stochastics* **12** 161–185.

PERKINS, E. A. (1985). Multiple stochastic integrals—a counter example. *Séminaire de Probabilités XIX. Lecture Notes in Math.* **1123**. Springer, Berlin.

PERKINS, E. A. (1992). Measure-valued branching diffusion with spatial interactions. *Probab. Theory Related Fields* **94** 189–245.

PERKINS, E. A. (1994). On the martingale problem for interactive measure-valued branching diffusions. *Mem. Amer. Math. Soc.* **115**.

REVUZ, D. and YOR, M. (1991). *Continuous Martingales and Brownian Motion*. Springer, Berlin.

ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusion, Markov Processes and Martingales* **2**. *Itô Calculus*. Wiley, New York.

ROSEN, J. (1992). Renormalization and limit theorems for self-intersections of superprocesses. *Ann. Probab.* **20** 1341–1368.

ROYDEN, H. L. (1968). *Real Analysis*, 2nd ed. Macmillan, New York.

RUIZ DE CHAVEZ, J. (1985). Sur les intégrales stochastiques multiples. *Séminaire de Probabilités XIX. Lecture Notes in Math.* **1123**. Springer, Berlin.

STRICKER, C. and YOR, M. (1978). Calcul stochastique dépendant d'un paramètre. *Z. Wahrsch. Verw. Gebiete* **45** 109–133.

SUGITANI, S. (1989). Some properties for measure-valued branching diffusion processes. *J. Math. Soc. Japan* **41** 437–462.

WALSH, J. B. (1986). An introduction to stochastic partial differential equations. *Ecole d'Eté de Probabilités de Saint-Flour, 1984. Lecture Notes in Math.* **1180**. Springer, Berlin.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CALIFORNIA 94720-3860

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
1984 MATHEMATICS ROAD
VANCOUVER, BC V6T 1Z2
CANADA