

## LOCAL CONDITIONING IN DAWSON–WATANABE SUPERPROCESSES

BY OLAV KALLENBERG

*Auburn University*

Consider a locally finite Dawson–Watanabe superprocess  $\xi = (\xi_t)$  in  $\mathbb{R}^d$  with  $d \geq 2$ . Our main results include some recursive formulas for the moment measures of  $\xi$ , with connections to the uniform Brownian tree, a Brownian snake representation of Palm measures, continuity properties of conditional moment densities, leading by duality to strongly continuous versions of the multivariate Palm distributions, and a local approximation of  $\xi_t$  by a stationary cluster  $\tilde{\eta}$  with nice continuity and scaling properties. This all leads up to an asymptotic description of the conditional distribution of  $\xi_t$  for a fixed  $t > 0$ , given that  $\xi_t$  charges the  $\varepsilon$ -neighborhoods of some points  $x_1, \dots, x_n \in \mathbb{R}^d$ . In the limit as  $\varepsilon \rightarrow 0$ , the restrictions to those sets are conditionally independent and given by the pseudo-random measures  $\tilde{\xi}$  or  $\tilde{\eta}$ , whereas the contribution to the exterior is given by the Palm distribution of  $\xi_t$  at  $x_1, \dots, x_n$ . Our proofs are based on the Cox cluster representations of the historical process and involve some delicate estimates of moment densities.

**1. Introduction.** This paper may be regarded as a continuation of [19], where we considered some local properties of a Dawson–Watanabe superprocess (henceforth referred to as a *DW-process*) at a fixed time  $t > 0$ . Recall that a DW-process  $\xi = (\xi_t)$  is a vaguely continuous, measure-valued diffusion process in  $\mathbb{R}^d$  with Laplace functionals  $E_\mu e^{-\xi_t f} = e^{-\mu v_t}$  for suitable functions  $f \geq 0$ , where  $v = (v_t)$  is the unique solution to the *evolution equation*  $\dot{v} = \frac{1}{2} \Delta v - v^2$  with initial condition  $v_0 = f$ . (This amounts to choosing the branching rate  $\gamma = 2$ . For general  $\gamma$ , we may reduce to this case by a suitable scaling.) We assume the initial measure  $\mu$  to be such that  $\xi_t$  is a.s. locally finite for every  $t > 0$ . (The precise criteria from [19] are quoted in Lemma 4.1.)

Our motivating result is Theorem 9.1, which describes asymptotically the conditional distribution of  $\xi_t$  for a fixed  $t > 0$ , given that  $\xi_t$  charges the  $\varepsilon$ -neighborhoods of some points  $x_1, \dots, x_n \in \mathbb{R}^d$ , where the approximation is in terms of total variation. In the limit, the restrictions to those sets are conditionally independent and given by some universal pseudo-random measures  $\tilde{\xi}$  or  $\tilde{\eta}$ , whereas the contribution to the exterior region is given by the multivariate Palm distribution of  $\xi_t$  at  $x_1, \dots, x_n$ .

---

Received February 2010; revised February 2011.

*MSC2010 subject classifications.* 60G57, 60J60, 60J80.

*Key words and phrases.* Measure-valued branching diffusions, moment measures and Palm distributions, local and global approximation, historical process, cluster representation, Brownian snake.

The present work may be regarded as part of a general research program outlined in [18], where we consider some random objects with similar local hitting and conditioning properties arising in different contexts. Examples identified so far include the simple point processes [10, 12, 15, 20, 30], local times of regenerative and related random sets [13, 14, 16], measure-valued diffusion processes [19], and intersection or self-intersection measures on random paths [24]. We are especially interested in cases where the local hitting probabilities are proportional to the appropriate moment densities, and the simple or multivariate Palm distributions can be approximated by elementary conditional distributions.

Our proofs, here as in [19], are based on the representation of each  $\xi_t$  as a countable sum of conditionally independent clusters of age  $h \in (0, t]$ , where the generating ancestors at time  $s = t - h$  form a Cox process  $\zeta_s$  directed by  $h^{-1}\xi_s$  (cf. [4, 26]). Typically we let  $h \rightarrow 0$  at a suitable rate depending on  $\varepsilon$ . In particular, the multivariate, conditional Slivnyak formula from [22] yields an explicit representation of the Palm distributions of  $\xi_t$  in terms of the Palm distributions for the individual clusters. Our arguments also rely on a detailed study of moment measures and Palm distributions, as well as on various approximation and scaling properties associated with the pseudo-processes  $\tilde{\xi}$  and  $\tilde{\eta}$ —all topics of independent interest covered by Sections 4–8. Here our analysis often goes far beyond what is needed in Section 9.

Moment measures of DW-processes play a crucial role in this paper, along with suitable versions of their densities. Thus, they appear in our asymptotic formulas for multivariate hitting probabilities, which extend the univariate results of Dawson et al. [3] and Le Gall [27]; cf. Lemma 7.2. They further form a convenient tool for the construction and analysis of multivariate Palm distributions, via the duality theory developed in [13]. Finally, they enter into a variety of technical estimates throughout the paper. In Theorem 4.2 we give a basic cluster decomposition of moment measures, along with a forward recursion (implicit in Dynkin [5]), a backward recursion and a Markov property. In Theorem 4.4 we explore the fundamental connection, first noted by Etheridge [7], between moment measures and certain uniform Brownian trees, and we provide several recursive constructions of the latter. The mentioned results enable us in Section 5 to establish some useful local estimates and continuity properties for ordinary and conditional moment densities.

Palm measures form another recurrent theme throughout the paper. After providing some general results on this topic in Section 3, we prove in Theorem 4.8 that the Palm distributions of a single DW-cluster can be obtained by ordinary conditioning from a suitably extended version of Le Gall's Brownian snake [26]. In Theorem 6.3 we use the cluster representation along with duality theory to establish some strong continuity properties of the multivariate Palm distributions.

Local approximations of DW-processes of dimension  $d \geq 3$  were studied already in [19], where we introduced a universal, stationary and scaling invariant (self-similar) pseudo-random measure  $\tilde{\xi}$ , providing a local approximation of  $\xi_t$  for

every  $t > 0$ , regardless of the initial measure  $\mu$ . (The prefix “pseudo” signifies that the underlying probability measure is not normalized and may be unbounded.) Though no such object exists for  $d = 2$ , we show in Section 8 that the stationary cluster  $\tilde{\eta}$  has similar approximation properties for all  $d \geq 2$  and satisfies some asymptotic scaling relations, which makes it a good substitute for  $\tilde{\xi}$ .

A technical complication when dealing with cluster representations is the possibility of multiple hits. More specifically, a single cluster may hit (charge) several of the  $\varepsilon$ -neighborhoods of  $x_1, \dots, x_n$ , or one of those neighborhoods may be hit by several clusters. To minimize the effect of such multiplicities, we need the cluster age  $h$  to be sufficiently small. (On the other hand, it needs to be large enough for the mentioned hitting estimates to apply to the individual clusters.) Probability estimates for multiple hits are derived in Section 7. Here we also estimate the effects of decoupling, where components of  $\xi_t$  involving possibly overlapping sets of clusters are replaced by conditionally independent measures.

Palm distributions of historical, spatial branching processes were first introduced in [11] under the name of *backward trees*, where they were used to derive criteria for persistence or extinction. The methods and ideas of [11] were extended to continuous time and more general processes in [8, 9, 29]. Further discussions of Palm distributions for superprocesses appear in [2, 4, 7, 19, 37]. In particular, a probabilistic (pathwise) description of the univariate Palm distributions of a DW-process is given in [2, 4]. More generally, there is a vast literature on conditioning in superprocesses (cf. [7], Sections 3.3–4). In particular, Salisbury and Verzani [33, 34] consider the conditional distribution of a DW-process in a bounded domain, given that the exit measure hits  $n$  given points on the boundary. However, their methods and results are entirely different from ours.

General surveys of superprocesses include the excellent monographs and lecture notes [2, 6, 7, 28, 32]. The literature on general random measures and the associated Palm kernels is vast; see [1, 12, 30] for some basic facts and further references.

For the sake of economy and readability, we are often taking slight liberties with the notation and associated terminology. Thus, for the DW-process we are often using the same symbol  $\xi$  to denote the measure-valued diffusion process itself, the associated historical process and the entire random evolution, involving complete information about the cluster structure for arbitrary times  $s < t$ . Likewise, we use  $\eta$  to denote the generic cluster of a DW-process, regarded as a measure-valued process in its own right, or the associated historical cluster, both determined (e.g.) by Le Gall’s Brownian snake based on a single Brownian excursion. Here Itô’s excursion law generates an infinite (though  $\sigma$ -finite) pseudo-distribution for  $\eta$ , here normalized such that  $P\{\eta_t \neq 0\} = t^{-1}$  for all  $t > 0$ . (This differs from the normalization in [19], which affects some formulas in subsequent sections.)

For the DW-process  $\xi$  and associated objects, we use  $P_\mu$  to denote probabilities under the assumption of initial measure  $\xi_0 = \mu$ . The associated distributions are

denoted by  $\mathcal{L}_\mu(\xi)$  or  $\mathcal{L}_\mu(\xi_t)$ . For the canonical cluster  $\eta$ , we define instead

$$P_\mu\{\eta \in \cdot\} = \mathcal{L}_\mu(\eta) = \int \mu(dx)\mathcal{L}_x(\eta) = \int \mu(dx)\mathcal{L}_0(\theta_x\eta),$$

where the shift operators  $\theta_x$  are defined by  $(\theta_x\mu)B = \mu(B - x)$  or  $(\theta_x\mu)f = \mu(f \circ \theta_x)$ , and we are writing  $\mathcal{L}_x$  instead of  $\mathcal{L}_{\delta_x}$ . When we use the notation  $P_\mu$  or  $\mathcal{L}_\mu$ , it is implicitly understood that  $\mu p_t < \infty$  for all  $t > 0$ . Let  $\mathcal{M}_d$  denote the space of locally finite measures on  $\mathbb{R}^d$ , endowed with the  $\sigma$ -field generated by all evaluation maps  $\pi_B : \mu \mapsto \mu B$  for arbitrary  $B \in \mathcal{B}^d$ , the Borel  $\sigma$ -field in  $\mathbb{R}^d$ . Write  $\hat{\mathcal{B}}^d$  or  $\hat{\mathcal{M}}_d$  for the classes of bounded Borel sets in  $\mathbb{R}^d$  or bounded measures on  $\mathbb{R}^d$ , respectively. For abstract measure spaces  $S$ , the meaning of  $\mathcal{M}_S$  or  $\hat{\mathcal{M}}_S$  is similar, except that we require the measures  $\mu \in \mathcal{M}_S$  to be *uniformly  $\sigma$ -finite*, in the sense that  $\mu B_k < \infty$  for some fixed measurable partition  $B_1, B_2, \dots$  of  $S$ .

We use double bars  $\|\cdot\|$  to denote the supremum norm when applied to functions, the operator norm when applied to matrices, and total variation when applied to signed measures. In the latter case, we define  $\|\mu\|_B = \|1_B\mu\|$ , where  $1_B\mu$  denotes the restriction of  $\mu$  to the set  $B$ . For any measure space  $\mathcal{M}_S$  and measurable set  $B \subset S$ , we consider the hitting set  $H_B = \{\mu \in \mathcal{M}_S; \mu B > 0\}$ , equipped with the  $\sigma$ -field generated by the restriction map  $\mu \mapsto 1_B\mu$ , and we often write  $\|\cdot\|_B$  instead of  $\|\cdot\|_{H_B}$  for convenience, referring to this as the total variation on  $B$ . Thus, in Section 8, we may write

$$(1) \quad \|\mathcal{L}(\tilde{\xi}) - \mathcal{L}(\tilde{\eta})\|_B = \|\mathcal{L}(1_B\tilde{\xi}) - \mathcal{L}(1_B\tilde{\eta})\|_{H_B},$$

even when the pseudo-distributions of  $\tilde{\xi}$  and  $\tilde{\eta}$  are unbounded. In Sections 7 and 9 we use a similar notation for signed measures on finite product spaces  $X_i\mathcal{M}_{S_i}$ , in which case (1) needs to be replaced by its counterpart for sequences of random measures.

For any space  $S$ , the components of  $x \in S^n$  are denoted by  $x_i$ , and we write  $S^{(n)}$  for the set of  $n$ -tuples  $x \in S^n$  with distinct components  $x_i$ . For functions  $f$ , we distinguish between ordinary powers  $f^n$  and tensor powers  $f^{\otimes n}$ , whereas for measures  $\mu$  the symbols  $\mu^{\otimes n}$  and  $\mu^{\otimes J}$  mean product measures. For point processes  $\zeta$  on  $S$ ,  $\zeta^{(n)}$  and  $\zeta^{(J)}$  denote the corresponding *factorial measures*, which for simple point processes, agree with the restrictions of  $\zeta^n$  and  $\zeta^J$  to  $S^{(n)}$  and  $S^{(J)}$ , respectively. Convolutions and convolution products are written as  $*$  and  $(*)_J$ . For suitable functions  $f$  and  $g$ ,  $f \sim g$  means  $f/g \rightarrow 1$ , whereas  $f \approx g$  means  $f - g \rightarrow 0$ , unless otherwise specified. The relation  $f \lesssim g$  means  $f \leq cg$  for some constant  $c > 0$ ,  $f \asymp g$  means  $f \lesssim g$  and  $g \lesssim f$ , and  $f \ll g$  means  $f/g \rightarrow 0$ .

Let  $\mathcal{P}_J$  be the class of partitions of the set  $J$ , and write  $\mathcal{P}_n$  when  $J = \{1, \dots, n\}$ . Define the scaling and shift operators  $S_x^r$  by  $S_x^r B = rB + x$  and put  $S_r = S_0^r$ . Thus,  $\mu S_x^r$  is the measure obtained by magnifying  $\mu$  around  $x$  by a factor  $r^{-1}$ . The open  $\varepsilon$ -ball around  $x$  is denoted by  $B_x^\varepsilon$ . Indicator functions are written as  $1\{\cdot\}$  or  $1_B$ , and  $\delta_s$  denotes the unit mass at  $s$ , so that  $\delta_s B = 1_B(s)$ . We write  $\mathbb{R}_+ = [0, \infty)$ ,  $\mathbb{Z}_+ = \{0, 1, \dots\}$  and  $\mathbb{N} = \{1, 2, \dots\}$ . The symbols  $\perp\!\!\!\perp$  and  $\perp\!\!\!\perp_\gamma$  mean

independence or conditional independence given  $\gamma$ , and we use  $\mathcal{L}(\xi)$  for the distribution of  $\xi$  and  $\stackrel{d}{=}$  for equality in distribution. We often write  $\mu f = \int f d\mu$  and  $(f \cdot \mu)B = \mu[f; B] = \int_B f d\mu$ . Conditional probabilities and distributions are written as  $P[\cdot|\cdot]$  and  $\mathcal{L}[\cdot|\cdot]$ , Palm measures and distributions as  $P[\cdot|\cdot]$  and  $\mathcal{L}[\cdot|\cdot]$ , respectively. We sometimes use  $\mathcal{L}^0$  to denote the Palm measure at 0.

**2. Gaussian, binomial and Poisson preliminaries.** Here we collect some properties of Gaussian measures and binomial or Poisson processes needed in subsequent sections. We begin with a simple exercise in linear algebra. By the *principal variances* of a random vector, we mean the positive eigenvalues of the associated covariance matrix.

LEMMA 2.1. *For any  $\pi \in \mathcal{P}_n$ , consider some uncorrelated random vectors  $\xi_J, J \in \pi$ , in  $\mathbb{R}^d$  with uncorrelated entries of variance  $\sigma^2$ , and put  $\xi_j = \xi_J$  for  $j \in J \in \pi$ . Then the array  $(\xi_1, \dots, \xi_n)$  has principal variances  $\sigma^2|J|, J \in \pi$ , each with multiplicity  $d$ .*

PROOF. By scaling we may take  $\sigma^2 = 1$ , and since each  $\xi_J$  has uncorrelated components, we may further take  $d = 1$ . Defining  $J_j$  by  $j \in J_j \in \pi$ , we get  $\text{Cov}(\xi_i, \xi_j) = \delta_{J_i, J_j}$ . It remains to note that the  $m \times m$  matrix with entries  $a_{ij} \equiv 1$  has eigenvalues  $m, 0, \dots, 0$ .  $\square$

We proceed with a simple comparison of normal distributions.

LEMMA 2.2. *Write  $\nu_\Lambda$  for the centered normal distribution on  $\mathbb{R}^n$  with covariance matrix  $\Lambda$ . Then*

$$\nu_\Lambda \leq \left( \frac{\|\Lambda\|^n}{\det \Lambda} \right)^{1/2} \nu_{\|\Lambda\|}^{\otimes n}.$$

PROOF. Let  $\Lambda$  have eigenvalues  $\lambda_1 \leq \dots \leq \lambda_n$ , and let  $x_1, \dots, x_n$  be the associated coordinates of  $x \in \mathbb{R}^n$ . Then  $\nu_\Lambda$  has density

$$\begin{aligned} \prod_{k \leq n} p_{\lambda_k}(x_k) &= \prod_{k \leq n} (2\pi \lambda_k)^{-1/2} e^{-|x_k|^2/2\lambda_k} \\ &\leq \prod_{k \leq n} (\lambda_n/\lambda_k)^{1/2} p_{\lambda_n}(x_k) \\ &= \left( \frac{\lambda_n^n}{\lambda_1 \cdots \lambda_n} \right)^{1/2} p_{\lambda_n}^{\otimes n}(x), \end{aligned}$$

and the assertion follows since  $\|\Lambda\| = \lambda_n$  and  $\det(\Lambda) = \lambda_1 \cdots \lambda_n$ .  $\square$

Now let  $p_t$  denote the continuous density of the symmetric Gaussian distribution on  $\mathbb{R}^d$  with variances  $t > 0$ .

LEMMA 2.3. *The normal densities  $p_t$  on  $\mathbb{R}^d$  satisfy*

$$p_s(x) \leq (1 \vee td|x|^{-2})^{d/2} p_t(x), \quad 0 < s \leq t, x \in \mathbb{R}^d \setminus \{0\}.$$

PROOF. For fixed  $t > 0$  and  $x \neq 0$ , the maximum of  $p_s(x)$  for  $s \in (0, t]$  occurs when  $s = (|x|^2/d) \wedge t$ . This gives  $p_s(x) \leq p_t(x)$  for  $|x|^2 \geq td$ , and for  $|x|^2 \leq td$  we have

$$\begin{aligned} p_s(x) &\leq (2\pi|x|^2/d)^{-d/2} e^{-d/2} \\ &\leq (2\pi|x|^2/d)^{-d/2} e^{-|x|^2/2t} = (td|x|^{-2})^{d/2} p_t(x). \end{aligned} \quad \square$$

For convenience, we also quote the elementary Lemma 3.1 from [19].

LEMMA 2.4. *For fixed  $d$  and  $T > 0$ , the normal densities  $p_t$  on  $\mathbb{R}^d$  satisfy*

$$p_t(x + y) \leq p_{t+h}(x), \quad x \in \mathbb{R}^d, |y| \leq h \leq t \leq T.$$

Given a measure  $\mu$  on  $\mathbb{R}^d$  and some measurable functions  $f_1, \dots, f_n \geq 0$  on  $\mathbb{R}^d$ , we introduce the convolution

$$\left(\mu * \bigotimes_{k \leq n} f_k\right)(x) = \int \mu(du) \prod_{k \leq n} f_k(x_k - u), \quad x = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n.$$

LEMMA 2.5. *Let  $\mu$  be a measure on  $\mathbb{R}^d$  with  $\mu p_t < \infty$  for all  $t > 0$ . Then for any  $n \in \mathbb{N}$ , the function  $(\mu * p_t^{\otimes n})(x)$  is finite and jointly continuous in  $(x, t) \in \mathbb{R}^{nd} \times (0, \infty)$ .*

PROOF. Letting  $t > 0$  and  $x \in \mathbb{R}^{nd}$  with  $c^{-1} < t < c$  and  $|x| < c$  for some constant  $c > 0$ , we see from Lemma 2.4 that

$$\prod_{k \leq n} p_t(x_k - u) \leq \prod_{k \leq n} p_c(x_k - u) \leq p_{2c}^n(u) \leq p_{2c/n}(u),$$

uniformly for  $u \in \mathbb{R}^d$ . Since  $\mu p_{2c/n} < \infty$ , we get  $(\mu * p_t^{\otimes n})(x) < \infty$  for any  $t > 0$ , and the asserted continuity follows by dominated convergence from the fact that  $p_t(x)$  is jointly continuous in  $(x, t) \in \mathbb{R}^d \times (0, \infty)$ .  $\square$

Given a probability measure  $\mu$  on some space  $S$ , let  $\sigma_1, \dots, \sigma_n$  be i.i.d. random elements in  $S$  with distribution  $\mu$ . Then the point process  $\xi = \sum_k \delta_{\sigma_k}$  on  $S$  (or any process with the same distribution) is called a *binomial process* based on  $\mu$ . We say that  $\xi$  is a *uniform binomial process* on an interval  $I$  if  $\mu$  is the uniform distribution on  $I$ .

We begin with a simple sampling property of binomial processes.

LEMMA 2.6. *Let  $\tau_1 < \dots < \tau_n$  form a uniform binomial process on  $[0, 1]$ , and consider an independent, uniformly distributed subset  $\varphi \subset \{1, \dots, n\}$  of fixed cardinality  $|\varphi| = k$ . Then the times  $\tau_r$  with  $r \in \varphi$  or  $r \notin \varphi$  form independent, uniform binomial processes on  $[0, 1]$  of orders  $k$  and  $n - k$ , respectively.*

PROOF. We may assume that  $\varphi = \{\pi_1, \dots, \pi_k\}$ , where  $\pi_1, \dots, \pi_n$  form a uniform permutation of  $1, \dots, n$  independent of  $\tau_1, \dots, \tau_n$ . The random variables  $\sigma_r = \tau \circ \pi_r, r = 1, \dots, n$ , are then i.i.d.  $U(0, 1)$ , and we have

$$\{\tau_r; r \in \varphi\} = \{\sigma_1, \dots, \sigma_k\}, \quad \{\tau_r; r \notin \varphi\} = \{\sigma_{k+1}, \dots, \sigma_n\}. \quad \square$$

This leads to a simple domination property for binomial processes:

LEMMA 2.7. *For each  $n \in \mathbb{Z}_+$ , let  $\xi_n$  be a uniform binomial process on  $[0, 1]$  with  $\|\xi_n\| = n$ . Then for any point process  $\eta \leq \xi_n$  with fixed  $\|\eta\| = k \leq n$ , we have*

$$\mathcal{L}(\eta) \leq \binom{n}{k} \mathcal{L}(\xi_k).$$

PROOF. Let  $\tau_1 < \dots < \tau_n$  be the points of  $\xi_n$ . Writing  $\xi_n^J = \sum_{j \in J} \delta_{\tau_j}$  when  $J \subset \{1, \dots, n\}$ , we have  $\eta = \xi_n^\varphi$  for some random subset  $\varphi \subset \{1, \dots, n\}$  with  $|\varphi| = k$  a.s. Choosing  $\psi \subset \{1, \dots, n\}$  to be independent of  $\xi_n$  and uniformly distributed with  $|\psi| = k$ , we get

$$\begin{aligned} \mathcal{L}(\eta) &= \mathcal{L}(\xi_n^\varphi) = \sum_J P\{\xi_n^J \in \cdot, \varphi = J\} \\ &\leq \sum_J \mathcal{L}(\xi_n^J) = \binom{n}{k} \mathcal{L}(\xi_n^\psi) = \binom{n}{k} \mathcal{L}(\xi_k), \end{aligned}$$

where the last equality holds by Lemma 2.6.  $\square$

We also need a conditional independence property of binomial processes.

LEMMA 2.8. *Let  $\sigma_1 < \dots < \sigma_n$  form a uniform binomial process on  $[0, t]$ , and fix any  $k \in \{1, \dots, n\}$ . Then:*

(i)

$$P\{\sigma_k \in ds\} = k \binom{n}{k} s^{k-1} (t - s)^{n-k} t^{-n} ds, \quad s \in (0, t);$$

(ii) *given  $\sigma_k$ , the times  $\sigma_1, \dots, \sigma_{k-1}$  and  $\sigma_{k+1}, \dots, \sigma_n$  form independent, uniform binomial processes on  $[0, \sigma_k]$  and  $[\sigma_k, t]$ , respectively.*

PROOF. Part (i) is elementary and classical. For part (ii) we note that, by Proposition 1.27 in [17], the times  $\sigma_1, \dots, \sigma_{k-1}$  form a uniform binomial process on  $[0, \sigma_k]$ , conditionally on  $\sigma_k, \dots, \sigma_n$ . By symmetry, the times  $\sigma_{k+1}, \dots, \sigma_n$  form a uniform binomial process on  $[\sigma_k, t]$ , conditionally on  $\sigma_1, \dots, \sigma_k$ . The conditional independence holds since the conditional distributions depend only on  $\sigma_k$ ; cf. Proposition 6.6 in [15].  $\square$

We proceed with a useful identity for homogeneous Poisson processes, stated in terms of the tetrahedral sets

$$t \Delta_n = \{s \in \mathbb{R}_+^n; s_1 < \dots < s_n < t\}, \quad t > 0, n \in \mathbb{N}.$$

LEMMA 2.9. *Let  $\tau_1 < \tau_2 < \dots$  form a Poisson process on  $\mathbb{R}_+$  with constant rate  $c > 0$ . Then for any measurable function  $f \geq 0$  on  $\mathbb{R}_+^{n+1}$  with  $n \in \mathbb{N}$ , we have*

$$(2) \quad Ef(\tau_1, \dots, \tau_{n+1}) = c^n E \int \dots \int_{\tau_1 \Delta_n} f(s_1, \dots, s_n, \tau_1) ds_1 \dots ds_n.$$

PROOF. Since

$$E\tau_1^n = \int_0^\infty t^n ce^{-ct} dt = n!c^{-n},$$

the right-hand side of (2) defines the joint distribution of some random variables  $\sigma_1, \dots, \sigma_{n+1}$ . Noting that  $\mathcal{L}(\tau_{n+1})$  has density

$$g_{n+1}(s) = \frac{c^{n+1}s^n e^{-cs}}{n!}, \quad s \geq 0,$$

we get for any measurable function  $f \geq 0$  on  $\mathbb{R}_+$

$$\begin{aligned} Ef(\tau_{n+1}) &= \int_0^\infty f(s)g_{n+1}(s) ds = \frac{c^{n+1}}{n!} \int_0^\infty s^n f(s)e^{-cs} ds \\ &= \frac{c^n}{n!} E\tau_1^n f(\tau_1) = c^n E \int \dots \int_{\tau_1 \Delta_n} f(\tau_1) ds_1 \dots ds_n, \end{aligned}$$

which shows that  $\sigma_{n+1} \stackrel{d}{=} \tau_{n+1}$ . We also see from (2) that  $\sigma_1, \dots, \sigma_n$  form a uniform binomial process on  $[0, \sigma_{n+1}]$ , conditionally on  $\sigma_{n+1}$ . Since the corresponding property holds for  $\tau_1, \dots, \tau_{n+1}$ , for example by Proposition 1.28 in [17], we obtain

$$(\sigma_1, \dots, \sigma_{n+1}) \stackrel{d}{=} (\tau_1, \dots, \tau_{n+1}),$$

as required.  $\square$

Say that a measurable space  $S$  is *additive* if it is closed under an associative, commutative and measurable operation “+” and contains a unique element 0 with

$s + 0 = s$  for all  $s \in S$ . Define  $l(s) \equiv s$ , and take  $\xi l = \int s \xi(ds)$  to be 0 when  $\xi = 0$ . We need a simple estimate for Poisson processes on an additive space. Recall that a *Borel space* is a measurable space  $S$ , that is, Borel isomorphic to a Borel set  $B \subset [0, 1]$ , so that there exists a one-to-one, bimeasurable map  $f : S \leftrightarrow B$ .

LEMMA 2.10. *On an additive Borel space  $S$ , consider a Poisson process  $\xi$  and a measurable function  $f \geq 0$ , where both  $f$  and  $E\xi$  are bounded. Then*

$$|Ef(\xi l) - E\xi f| \leq \|f\| \|E\xi\|^2.$$

PROOF. Writing  $p = \|E\xi\|$ , we get

$$\begin{aligned} E[f(\xi l); \|\xi\| > 1] &\leq \|f\| P\{\|\xi\| > 1\} \\ &= (1 - (1 + p)e^{-p})\|f\| \leq \frac{1}{2}p^2\|f\|. \end{aligned}$$

Since  $\xi$  is a mixed binomial process on  $S$  based on  $E\xi$  (cf. Proposition 1.28 in [17]), we also have

$$\begin{aligned} E[f(\xi l); \|\xi\| = 1] &= P\{\|\xi\| = 1\}E[f(\xi l)|\|\xi\| = 1] \\ &= pe^{-p} \frac{E\xi f}{\|E\xi\|} = e^{-p}E\xi f, \end{aligned}$$

and so

$$\begin{aligned} 0 &\leq E\xi f - E[f(\xi l); \|\xi\| = 1] \\ &= (1 - e^{-p})E\xi f \leq p\|E\xi\|\|f\| = p^2\|f\|. \end{aligned}$$

Noting that

$$Ef(\xi l) - E\xi f = E[f(\xi l); \|\xi\| > 1] + E[f(\xi l); \|\xi\| = 1] - E\xi f,$$

we get by combination

$$|Ef(\xi l) - E\xi f| \leq (\frac{1}{2}p^2\|f\|) \vee (p^2\|f\|) = p^2\|f\|. \quad \square$$

We conclude with an elementary inequality needed in Section 7.

LEMMA 2.11. *For any  $n \in \mathbb{N}$  and  $k_1, \dots, k_n \in \mathbb{Z}_+$ , we have*

$$2\left(\prod_{j \leq n} k_j - 1\right)_+ \leq \sum_{i \leq n} (k_i - 1) \prod_{j \neq i} k_j.$$

PROOF. Clearly

$$(hk - 1)_+ \leq h(k - 1)_+ + k(h - 1)_+, \quad h, k \in \mathbb{Z}_+.$$

Proceeding by induction, we obtain

$$\left(\prod_{j \leq n} k_j - 1\right)_+ \leq \sum_{i \leq n} (k_i - 1) \prod_{j \neq i} k_j.$$

It remains to note that  $(k - 1)_+ \leq k(k - 1)/2$  for  $k \in \mathbb{Z}_+$ .  $\square$

**3. Measure, kernel and Palm preliminaries.** Here we collect some general propositions about measures, kernels and Palm distributions, needed in subsequent sections. The first few results are easy and probably known, though no references could be found.

LEMMA 3.1. *For any measurable space  $S$ , the space  $\hat{\mathcal{M}}_S$  is complete in total variation.*

PROOF. Let  $\mu_1, \mu_2, \dots \in \hat{\mathcal{M}}_S$  with  $\|\mu_m - \mu_n\| \rightarrow 0$  as  $m, n \rightarrow \infty$ . Assuming  $\mu_n \neq 0$ , we may define  $\nu = \sum_n 2^{-n} \mu_n / \|\mu_n\|$  and choose some measurable functions  $f_1, f_2, \dots \in L^1(\nu)$  with  $\mu_n = f_n \cdot \nu$ . Then  $\nu|f_m - f_n| = \|\mu_m - \mu_n\| \rightarrow 0$ , which means that  $(f_n)$  is Cauchy in  $L^1(\nu)$ . Since  $L^1$  is complete (cf. [15], page 16), we have convergence  $f_n \rightarrow f$  in  $L^1$ , and so the measure  $\mu = f \cdot \nu$  satisfies  $\|\mu - \mu_n\| = \nu|f - f_n| \rightarrow 0$ .  $\square$

For any measure  $\mu$  on a topological space  $S$ , we define  $\text{supp } \mu$  as the intersection of all closed sets  $F \subset S$  with  $\mu F^c = 0$ .

LEMMA 3.2. *Fix a measure  $\mu$  on a Polish space  $S$ . Then  $\mu(\text{supp } \mu)^c = 0$ , and  $s \in \text{supp } \mu$  iff  $\mu G > 0$  for every neighborhood  $G$  of  $s$ .*

PROOF. Choose a countable base  $B_1, B_2, \dots$  of  $S$ , and define  $I = \{i \in \mathbb{N}; \mu B_i = 0\}$ . Any open set  $G \subset S$  can be written as  $\bigcup_{i \in J} B_i$  for some  $J \subset \mathbb{N}$ , and we note that  $\mu G = 0$  iff  $J \subset I$ . Hence,  $(\text{supp } \mu)^c = \bigcup_{i \in I} B_i$ . If  $s \notin \text{supp } \mu$ , then  $s \in B_i$  for some  $i \in I$ , and so  $\mu G = 0$  for some neighborhood  $G$  of  $s$ . Conversely, the latter condition implies  $s \in B_i$  for some  $i \in I$ , and so  $s \notin \text{supp } \mu$ .  $\square$

We continue with a simple measurability property.

LEMMA 3.3. *Let  $S$  and  $T$  be Borel spaces. For any  $\mu \in \hat{\mathcal{M}}_{S \times T}$  and  $t \in T^d$ , let  $\mu_t$  denote the restriction of  $\mu$  to  $S \times \{t_1, \dots, t_d\}^c$ . Then the mapping  $(\mu, t) \mapsto \mu_t$  is product-measurable.*

PROOF. We may take  $T = \mathbb{R}$ . Put  $I_{nj} = 2^{-n}(j - 1, j]$ ,  $n, j \in \mathbb{Z}$ , and define

$$U_n(t) = \bigcup_j \{I_{nj}; t_1, \dots, t_d \notin I_{nj}\}, \quad n \in \mathbb{N}, t \in \mathbb{R}^d.$$

Then the restriction  $\mu_t^n$  of  $\mu$  to  $S \times U_n(t)$  is product-measurable, and  $\mu_t^n \uparrow \mu_t$  by monotone convergence.  $\square$

Given two measurable spaces  $(S, \mathcal{S})$  and  $(T, \mathcal{T})$ , a *kernel* from  $S$  to  $T$  is defined as a function  $\mu \geq 0$  on  $(S, \mathcal{T})$  such that  $\mu_s B = \mu(s, B)$  is measurable in  $s \in S$  for

fixed  $B$  and a measure in  $B \in \mathcal{T}$  for fixed  $s$ . For any measure  $\nu$  on  $S$  and kernel  $\mu$  from  $S$  to  $T$ , we define the *composition*  $\nu \otimes \mu$  and *product*  $\nu\mu$  as the measures on  $S \times T$  and  $T$ , respectively, given by

$$(\nu \otimes \mu)f = \int \nu(ds) \int \mu_s(dt) f(s, t), \quad \nu\mu = (\nu \otimes \mu)(S \times \cdot).$$

Conversely, when  $T$  is Borel, any  $\sigma$ -finite measure  $M$  on  $S \times T$  admits a *disintegration*  $M = \nu \otimes \mu$  into a  $\sigma$ -finite *supporting* measure  $\nu$  on  $S$  and a kernel  $\mu$  from  $S$  to  $T$ , where the latter is again  $\sigma$ -finite, in the sense that  $\mu_s f(s, \cdot) < \infty$  for some measurable function  $f > 0$  on  $S \times T$ . When  $M(\cdot \times T)$  is  $\sigma$ -finite we may take  $\nu = M(\cdot \times T)$ , in which case  $\mu$  can be chosen to be a *probability kernel*, in the sense that  $\|\mu_s\| = 1$  for all  $s$ . In general, the measures  $\mu_s$  are unique,  $s \in S$  a.e.  $\nu$ , up to normalizations.

Some basic properties of kernels and their compositions are given in [15]. Here we first consider the total variation  $\|\nu \otimes \mu\|$ , where  $\mu$  is a *signed kernel*, defined as the difference between two a.e. bounded kernels.

LEMMA 3.4. *For any measurable space  $S$  and Borel space  $T$ , let  $\nu \in \hat{\mathcal{M}}_S$ , and consider a signed kernel  $\mu$  from  $S$  to  $T$ . Then  $\|\mu\|$  is measurable and  $\|\nu \otimes \mu\| = \nu\|\mu\|$ .*

PROOF. Assuming  $\mu = \mu' - \mu''$  for some bounded kernels  $\mu'$  and  $\mu''$  from  $S$  to  $T$ , we define  $\hat{\mu} = \mu' + \mu''$ . Since  $T$  is Borel, Proposition 7.26 in [15] yields a measurable function  $f : S \times T \rightarrow [-1, 1]$  with  $\mu = f \cdot \hat{\mu}$ . Then  $\|\mu\| = \hat{\mu}|f|$ , which is measurable by Lemma 1.41 in [15]. Furthermore,

$$\|\nu \otimes \mu\| = \|f \cdot (\nu \otimes \hat{\mu})\| = (\nu \otimes \hat{\mu})|f| = \nu(\hat{\mu}|f|) = \nu\|\mu\|. \quad \square$$

We proceed with a simple projection property.

LEMMA 3.5. *For any Borel spaces  $S, T$  and  $U$ , consider some  $\sigma$ -finite measures  $\nu$  and  $\hat{\nu}$  on  $S \times U$  and  $S$  and some signed kernels  $\mu$  and  $\hat{\mu}$  from  $S \times U$  or  $S$  to  $T$ , such that  $\nu$  and  $\nu \otimes \mu$  have projections  $\hat{\nu}$  and  $\hat{\nu} \otimes \hat{\mu}$  onto  $S$  and  $S \times T$ , respectively. Then  $\|\hat{\mu}_S\| \leq \sup_U \|\mu_{S,U}\|$  a.e.  $\hat{\nu}$ .*

PROOF. Since  $\hat{\nu}$  is  $\sigma$ -finite and  $U$  is Borel, we have  $\nu = \hat{\nu} \otimes \rho$  for some probability kernel  $\rho$  from  $S$  to  $U$ . Writing  $\pi_{S \times T}$  for projection onto  $S \times T$ , we obtain

$$\hat{\nu} \otimes \hat{\mu} = (\nu \otimes \mu) \circ \pi_{S \times T}^{-1} = (\hat{\nu} \otimes \rho \otimes \mu) \circ \pi_{S \times T}^{-1} = \hat{\nu} \otimes \rho\mu,$$

and so  $\hat{\mu} = \rho\mu$  a.e.  $\hat{\nu}$ . Hence, for any measurable function  $f$  on  $T$  with  $|f| \leq 1$ , we get a.e.

$$|\hat{\mu}f| = |(\rho\mu)f| = |\rho(\mu f)| \leq \rho|\mu f| \leq \rho\|\mu\|,$$

which implies

$$\|\hat{\mu}_s\| \leq \rho_s \|\mu_s\| \leq \sup_u \|\mu_{s,u}\| \quad \text{a.e. } \hat{\nu}. \quad \square$$

The following technical result plays a crucial role in Section 6. For any  $G_1, G_2, \dots \subset S$ , put  $\limsup_n G_n = \bigcap_n \bigcup_{k \geq n} G_k = \{s \in S; s \in G_n \text{ i.o.}\}$ .

LEMMA 3.6. *Let  $\nu$  be a kernel from  $\mathbb{R}$  to a Polish space  $S$  with  $\text{supp } \nu = S$ , let  $\mu, \mu_1, \mu_2, \dots$  be bounded kernels from  $S \times \mathbb{R}$  to a Borel space  $U$ , where each  $\mu_n$  is continuous in total variation on  $S \times G_n$  for some open set  $G_n \subset \mathbb{R}$ . Assume  $\nu\{\|\mu - \mu_n\| > h_n\} \equiv 0$  for some measurable functions  $h_n : S \times \mathbb{R} \rightarrow \mathbb{R}_+$  with  $h_n \rightarrow 0$  uniformly on bounded sets. Then  $\mu = \mu'$  a.e.  $\nu$ , where  $\mu'$  is continuous in total variation on  $S \times \limsup_n G_n$ .*

PROOF. First let the kernels  $\nu, \mu, \mu_1, \mu_2, \dots$  and functions  $h_1, h_2, \dots$  be independent of the real parameter, hence kernels or functions on  $S$ . Let  $S' \subset S$  be the set where  $\|\mu - \mu_n\| \leq h_n$ , so that  $\nu(S')^c = 0$ . For any  $t, t' \in S'$ , we have

$$\|\mu_t - \mu_{t'}\| \leq \|\mu_t - \mu_t^n\| + \|\mu_t^n - \mu_{t'}^n\| + \|\mu_{t'}^n - \mu_{t'}\|,$$

where  $\mu^n = \mu_n$ . Fixing any  $s \in S$ , we may let  $t, t' \rightarrow s$  and then  $n \rightarrow \infty$  to get  $\|\mu_t - \mu_{t'}\| \rightarrow 0$ . Hence, Lemma 3.1 yields a bounded measure  $\mu'_s$  on  $U$  with  $\|\mu_t - \mu'_s\| \rightarrow 0$ . Note that  $\mu'_s$  is well defined for every  $s \in S$  and that  $\mu'_s = \mu_s$  when  $s \in S'$ .

To prove the required continuity of  $\mu'$ , suppose that  $s_k \rightarrow s$  in  $S$ . Fixing any metrization  $d$  of  $S$ , we may choose  $t_1, t_2, \dots \in S'$  with

$$d(s_k, t_k) + \|\mu'_{s_k} - \mu_{t_k}\| < 2^{-k}, \quad k \in \mathbb{N}.$$

In particular  $t_k \rightarrow s$ , and so

$$\|\mu'_{s_k} - \mu'_s\| \leq \|\mu'_{s_k} - \mu_{t_k}\| + \|\mu_{t_k} - \mu'_s\| \rightarrow 0,$$

as desired. The continuity of  $\mu'$  implies measurability, which means that  $\mu'$  is again a locally bounded kernel from  $S$  to  $U$ . Further note that  $\|\mu_n - \mu'\| \leq h_n$  on  $S'$ , which extends by continuity to  $\|\mu_n - \mu'\| \leq h'_n$  on  $S$ , where the functions  $h'_n$  are upper semi-continuous versions of  $h_n$ , satisfying the same convergence condition.

We now allow  $\nu, \mu$  and  $\mu_1, \mu_2, \dots$  to depend on a parameter  $x \in \mathbb{R}$ . Constructing  $\mu'$  as before for each  $x$ , we get  $\|\mu_n - \mu'\| \rightarrow 0$  uniformly on bounded sets in  $S \times \mathbb{R}$ . Since each  $\mu_n$  is continuous in total variation on  $S \times G_n$ , the same continuity holds for  $\mu'$  on the set  $S \times \limsup_n G_n$ .  $\square$

A random measure  $\xi$  on a measurable space  $S$  is defined as a kernel from the basic probability space  $\Omega$  into  $S$ . The intensity  $E\xi$  is the measure on  $S$  given by  $(E\xi)f = E(\xi f)$ . For any random element  $\eta$  in a measurable space  $T$ , we define the associated Campbell measure  $M$  on  $S \times T$  by  $Mf = E \int \xi(ds) f(s, \eta)$ . When

$T$  is Borel, and  $M$  is  $\sigma$ -finite, we may form the disintegration  $M = \nu \otimes \mu$ , where  $\mu$  is a  $\sigma$ -finite kernel of *Palm measures*  $\mu_s$  on  $T$ . If  $E\xi = M(\cdot \times T)$  is  $\sigma$ -finite, we may take  $\nu = E\xi$  and choose  $\mu$  to be a probability kernel from  $S$  to  $T$ , in which case the  $\mu_s$  are called *Palm distributions* of  $\eta$  with respect to  $\xi$ . For convenience, we write

$$\mu_s = P[\eta \in \cdot \mid \xi]_s = \mathcal{L}[\eta \mid \xi]_s, \quad \mu_s f(s, \cdot) = E[f(s, \eta) \mid \xi]_s.$$

Alternatively,  $P[\cdot \mid \xi]$  may be regarded as a kernel from  $S$  to the basic probability space  $\Omega$  with  $\sigma$ -field generated by  $\eta$ . The multivariate Palm distributions are defined as the kernels  $\mathcal{L}[\eta \mid \xi^{\otimes n}]$  from  $S^n$  to  $T$ , for arbitrary  $n \in \mathbb{N}$ .

The following conditioning approach to Palm distributions (cf. [18, 21]) is often useful. On a measure space with pseudo-probability  $\tilde{P}$ , we introduce a random pair  $(\sigma, \tilde{\eta})$  in  $S \times T$  with

$$\tilde{E} f(\sigma, \tilde{\eta}) = E \int f(s, \eta) \xi(ds).$$

Then the Palm distributions of  $\eta$  with respect to  $\xi$  are given by

$$E[f(\eta) \mid \xi]_\sigma = \tilde{E}[f(\tilde{\eta}) \mid \sigma] \quad \text{a.s.}$$

The duality between Palm measures and conditional moment densities was first noted in [13]. The following versions of the main results (with subsequent clarifications) are convenient for our present purposes.

LEMMA 3.7. *On a filtered probability space  $(\Omega, \mathcal{F}, P)$ , consider a random measure  $\xi$  on a Polish space  $S$  with  $\sigma$ -finite intensity  $E\xi$  and some  $(\mathcal{F}_t \otimes \mathcal{S})$ -measurable processes  $M_t$  on  $S$ . Then:*

- (i)  $E[\xi \mid \mathcal{F}_t] = M_t \cdot E\xi$  a.s. iff  $P[\cdot \mid \xi]_s = M_{s,t} \cdot P$  a.e. on  $\mathcal{F}_t$ . In this case, the versions  $P[\cdot \mid \xi]_{s,t} = M_{s,t} \cdot P$  on  $\mathcal{F}_t$  are such that
- (ii) for fixed  $t$ , the measure  $P[\cdot \mid \xi]_{s,t}$  is continuous in  $s \in S$ , in total variation on  $\mathcal{F}_t$ , if and only if  $M_{s,t}$  is  $L^1$ -continuous in  $s$ ,
- (iii) for fixed  $s \in S$ , the measures  $P[\cdot \mid \xi]_{s,t}$  on  $\mathcal{F}_t$  are consistent in  $t$  if and only if  $M_{s,t}$  is a martingale in  $t$ ,
- (iv) if the  $\mathcal{F}_t$  are countably generated and the continuity in (ii) holds for every  $t$ , then the consistency in (iii) holds for all  $s \in \text{supp } E\xi$ .

Here  $M$  is a function on  $S \times \mathbb{R}_+ \times \Omega$  such that  $M(s, t, \omega)$  is product-measurable in  $(s, \omega)$  for each  $t$ , and we are writing  $M_t = M(\cdot, t, \cdot)$  and  $M_{s,t} = M(s, t, \cdot)$ . The Palm distributions  $P[\cdot \mid \xi]_s$  form a kernel from  $S$  to  $\Omega$ , endowed with any of the  $\sigma$ -fields  $\mathcal{F}_t$ , and in (ii)–(iv) we consider some special versions  $P[\cdot \mid \xi]_{s,t}$  of those measures on  $\mathcal{F}_t$ .

PROOF. (i) If  $E[\xi|\mathcal{F}_t] = M_t \cdot E\xi$  a.s., then for any  $A \in \mathcal{F}_t$  and  $B \in \mathcal{S}$

$$\begin{aligned} \int_B E\xi(ds)P[A|\xi]_s &= E[\xi B; A] = E[E[\xi B|\mathcal{F}_t]; A] \\ &= E\left[\int_B M_{s,t}E\xi(ds); A\right] = \int_B E\xi(ds)E[M_{s,t}; A], \end{aligned}$$

and so

$$P[A|\xi]_s = E[M_{s,t}; A] = (M_{s,t} \cdot P)A, \quad s \in S \text{ a.e. } E\xi,$$

which shows that we can choose  $P[\cdot|\xi]_s = M_{s,t} \cdot P$  on  $\mathcal{F}_t$ . Conversely, if  $P[\cdot|\xi]_s = M_{s,t} \cdot P$  a.e. on  $\mathcal{F}_t$ , then a similar calculation yields

$$E[\xi B|\mathcal{F}_t] = \int_B M_{s,t}E\xi(ds) = (M_t \cdot E\xi)B \quad \text{a.s.},$$

which means that we can choose  $E[\xi|\mathcal{F}_t] = M_t \cdot E\xi$  on  $\mathcal{S}$ .

(ii) This follows from the  $L^\infty/L^1$ -isometry

$$\|P[\cdot|\xi]_{s,t} - P[\cdot|\xi]_{s',t}\|_t = \|(M_{s,t} - M_{s',t}) \cdot P\|_t = E|M_{s,t} - M_{s',t}|,$$

where  $\|\cdot\|_t$  denotes total variation on  $\mathcal{F}_t$ .

(iii) Since  $M_{s,t}$  is  $\mathcal{F}_t$ -measurable, we get for any  $A \in \mathcal{F}_t$  and  $t \leq t'$ ,

$$\begin{aligned} P[A|\xi]_{s,t} - P[A|\xi]_{s,t'} &= E[M_{s,t} - M_{s,t'}; A] \\ &= E[M_{s,t} - E[M_{s,t'}|\mathcal{F}_t]; A], \end{aligned}$$

which vanishes for every  $A$ , if and only if  $M_{s,t} = E[M_{s,t'}|\mathcal{F}_t]$  a.s.

(iv) For any  $t \leq t'$  and  $A \in \mathcal{F}_t$ , we have  $P[A|\xi]_{s,t} = P[A|\xi]_{s,t'}$  a.e. by the uniqueness of the Palm disintegration. Since  $\mathcal{F}_t$  is countably generated, a monotone-class argument gives  $P[\cdot|\xi]_{s,t} = P[\cdot|\xi]_{s,t'}$  a.e. on  $\mathcal{F}_t$ , and so  $M_{s,t} = E[M_{s,t'}|\mathcal{F}_t]$  a.s. for  $s \in S$  a.e.  $E\xi$  as in (iii). By the  $L^1$ -continuity of  $M_{s,t}$  and  $M_{s,t'}$  and the  $L^1$ -contractivity of conditional expectations, the latter relation extends to  $\text{supp } E\xi$ , and so by (iii) the measures  $P[\cdot|\xi]_{s,t}$  are consistent for all  $s \in \text{supp } E\xi$ . □

Often in applications,  $E\xi = p \cdot \lambda$  for some  $\sigma$ -finite measure  $\lambda$  on  $S$  and continuous function  $p > 0$  on  $S$ . Assuming  $E[\xi|\mathcal{F}_t] = X_t \cdot \lambda$  a.s. for some  $(\mathcal{F}_t \otimes \mathcal{S})$ -measurable processes  $X_t$ , we may choose  $M_t = X_t/p$  in (i). Note that the  $L^1$ -continuity in (ii) and the martingale property in (iii) hold simultaneously for  $X$  and  $M$ .

A *point process* on a measure space  $T$  is defined as a random measure  $\zeta$  of the form  $\sum_i \delta_{\tau_i}$ , where the  $\tau_i$  are random elements in  $T$ . Given  $\zeta$  and a probability kernel  $\nu$  from  $T$  to a measure space  $\mathcal{M}_S$ , we may form a *cluster process*  $\xi = \sum_i \eta_i$ , where the  $\eta_i$  are conditionally independent random measures on  $S$  with distributions  $\nu_{\tau_i}$ . We assume  $\zeta$  and  $\nu$  to be such that  $\xi B_k < \infty$  a.s. for some measurable

partition  $B_1, B_2, \dots$  of  $S$ . If  $\zeta$  is Poisson or Cox, we call  $\xi$  a Poisson or Cox cluster process generated by  $\zeta$  and  $\nu$ .

We need the following representation for the Palm measures of a Cox cluster process, quoted from [22]. For Poisson processes and dimension  $n = 1$ , the result goes back to [12, 23, 31, 35], and applications to superprocesses appear in [2, 4]. When  $\xi$  is a Poisson cluster process generated by a measure  $\mu \in \mathcal{M}_T$  and a probability kernel  $\nu$  from  $T$  to  $\mathcal{M}_S$ , let  $\tilde{\xi}$  denote a random measure on  $S$  with pseudo-distribution  $\tilde{\mu} = \mu\nu$ . Given a  $\sigma$ -finite measure  $\mu = \sum_n \mu_n$ , we call  $d\mu_n/d\mu$  the relative density of  $\mu_n$  with respect to  $\mu$ . Write  $\mathcal{L}_\eta$  and  $E_\eta$  for the conditional distributions and expectations given  $\eta$ .

LEMMA 3.8. *Let  $\xi$  be a cluster process on  $S$  generated by a Cox process on  $T$  with directing measure  $\eta$ . Then*

$$(3) \quad E_\eta \xi^{\otimes n} = \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} E_\eta \tilde{\xi}^{\otimes J}, \quad n \in \mathbb{N},$$

which yields  $E\xi^{\otimes n}$  by integration with respect to  $\mathcal{L}(\eta)$ . Assuming  $E_\eta \xi^{\otimes n}$  to be a.s.  $\sigma$ -finite and writing  $p_\eta^\pi$  for the relative densities in (3), we have

$$(4) \quad \mathcal{L}_\eta[\xi \|\xi^{\otimes n}]_s = \mathcal{L}_\eta(\xi) * \sum_{\pi \in \mathcal{P}_n} p_\eta^\pi(s) \big( * \big)_{J \in \pi} \mathcal{L}_\eta[\tilde{\xi} \|\tilde{\xi}^{\otimes J}]_{s_J} \quad a.e.,$$

which yields  $\mathcal{L}[\xi \|\xi^{\otimes n}]_s$  by integration with respect to  $\mathcal{L}[\eta \|\xi^{\otimes n}]_s$ .

Here  $\mathcal{L}[\xi \|\xi^{\otimes n}]$  and  $\mathcal{L}[\eta \|\xi^{\otimes n}]$  need to be based on the same supporting measure for  $\xi^{\otimes n}$ , even when  $E\xi^{\otimes n}$  fails to be  $\sigma$ -finite. For a probabilistic interpretation in the Poisson case, consider for any  $s \in S^n$ , a random partition  $\pi_s \perp\!\!\!\perp \xi$  in  $\mathcal{P}_n$  with distribution given by the relative densities  $p_s^\pi$  in (3) and some independent Palm versions  $\tilde{\xi}_{s_J}^J$  of  $\tilde{\xi}$ . Then (4) is equivalent to

$$\xi_s \stackrel{d}{=} \xi + \sum_{J \in \pi_s} \tilde{\xi}_{s_J}^J, \quad s \in S^n \text{ a.e. } E_\mu \xi^{\otimes n}.$$

The result extends immediately to Palm measures of the form  $\mathcal{L}[\xi' \|\xi''^{\otimes n}]_s$ , where  $\xi'$  and  $\xi''$  are random measures on  $S'$  and  $S''$  such that the pair  $(\xi', \xi'')$  forms a Cox cluster process directed by  $\eta$ . Indeed, assuming  $S'$  and  $S''$  to be disjoint, we may apply Lemma 3.8 to the cluster process  $\xi = \xi' + \xi''$  on  $S = S' \cup S''$ . This more general version is needed for the proof of Lemma 6.2 below.

Next we show how the moment and Palm measures of a random measure  $\eta$  are affected by a shift by a fixed measure  $\mu$ . Here we define  $(\theta_s \mu)f = \mu(f \circ \theta_s)$ , where  $\theta_s t = s + t$ .

LEMMA 3.9. *Let  $\mathcal{L}_\mu(\eta) = \int \mu(ds) \mathcal{L}(\theta_s \eta)$  for some random measure  $\eta$  on  $\mathbb{R}^d$  and a  $\mu \in \mathcal{M}_d$  such that  $E\eta^{\otimes n} = p_n \cdot \lambda^{\otimes nd}$  with  $\mu * p_n < \infty$  a.e. Then  $E_\mu \eta^{\otimes n} =$*

$(\mu * p_n) \cdot \lambda^{\otimes nd}$  and

$$(5) \quad \mathcal{L}_\mu[\eta \parallel \eta^{\otimes n}]_s = \int \frac{p_n(s-r)\mu(dr)}{(\mu * p_n)(s)} \mathcal{L}[\theta_r \eta \parallel \eta^{\otimes n}]_{s-r} \quad a.e. E_\mu \eta^{\otimes n}.$$

PROOF. Fixing  $n$ , we may write  $p_n = p$  and  $\lambda^{\otimes nd}(ds) = ds$ , and let  $\mu_s$  denote the mixing measure in (5). Then

$$\begin{aligned} E_\mu \int \eta^{\otimes n}(ds) f(s, \eta) &= \int \mu(dr) E \int \eta^{\otimes n}(ds) f(s+r, \theta_r \eta) \\ &= \int \mu(dr) \int E \eta^n(ds) E[f(s+r, \theta_r \eta) \parallel \eta^{\otimes n}]_s \\ &= \int \mu(dr) \int p(s-r) ds E[f(s, \theta_r \eta) \parallel \eta^{\otimes n}]_{s-r} \\ &= \int E_\mu \eta^{\otimes n}(ds) \int \mu_s(dr) E[f(s, \theta_r \eta) \parallel \eta^{\otimes n}]_{s-r}, \end{aligned}$$

where the first step holds by the definition of  $P_\mu$ , the second step holds by Palm disintegration and the third step holds by the definition of  $p$  and the invariance of  $\lambda$ . This gives  $E_\mu \eta^{\otimes n} = (\mu * p) \cdot \lambda^{\otimes nd}$ , and so the fourth step holds by Fubini's theorem. Now (5) follows by the uniqueness of the Palm disintegration.  $\square$

We turn to the Palm measures of a random measure of the form  $\xi \otimes \delta_\tau$ .

LEMMA 3.10. *For any random measure  $\xi$  on  $S$  and random element  $\tau$  in a Borel space  $T$ , we have*

$$P[\tau \neq t \parallel \xi \otimes \delta_\tau]_{s,t} = 0, \quad (s, t) \in S \times T \text{ a.e. } E(\xi \otimes \delta_\tau).$$

PROOF. Since  $T$  is Borel, the diagonal  $\{(t, t); t \in T\}$  in  $T^2$  is measurable. Letting  $\nu$  be the associated supporting measure for  $\xi \otimes \delta_\tau$ , we get by Palm disintegration

$$\begin{aligned} \int \int \nu(ds dt) P[\tau \neq t \parallel \xi \otimes \delta_\tau]_{s,t} &= E \int \int (\xi \otimes \delta_\tau)(ds dt) 1\{\tau \neq t\} \\ &= E \int \xi(ds) 1\{\tau \neq \tau\} = 0, \end{aligned}$$

and the assertion follows.  $\square$

The following continuity property of Palm distributions extends Lemma 2.2 in [19]. For any random measure  $\xi$  on  $\mathbb{R}^d$ , we define the *centered Palm distributions* by  $P^s = \mathcal{L}[\theta_{-s} \xi \parallel \xi]_s$ . Recall that  $E[\xi; A] = E(\xi 1_A)$ .

LEMMA 3.11. *Let  $\xi_n$  and  $\eta_n$  be random measures on  $\mathbb{R}^d$  with centered Palm distributions  $P_n^s$  and  $Q_n^s$ , respectively, and fix any  $B \in \hat{\mathcal{B}}^d$ . Then the following conditions imply  $\sup_{s \in B} \|P_n^s - Q_n^s\| \rightarrow 0$ :*

- (i)  $E\xi_n B \asymp 1$ ;
- (ii)  $\|E[\xi_n B; \xi_n \in \cdot] - E[\eta_n B; \eta_n \in \cdot]\| \rightarrow 0$ ;
- (iii)  $\sup_{r,s \in B} \|P_n^r - P_n^s\| + \sup_{r,s \in B} \|Q_n^r - Q_n^s\| \rightarrow 0$ .

PROOF. Writing  $f_A(\mu) = (\mu B)^{-1} \int_B \mu(ds) 1_A(\theta_{-s}\mu)$ , we get

$$(6) \quad \int_B E\xi_n(ds) P_n^s A = E(\xi_n B) f_A(\xi_n) = \int E[\xi_n B; \xi_n \in d\mu] f_A(\mu),$$

and similarly for  $\eta_n$ . For any  $s \in B$ , we have, by (i),

$$\begin{aligned} \|P_n^s - Q_n^s\| &\leq E(\xi_n B) \|P_n^s - Q_n^s\| \\ &\leq \|E(\xi_n B) P_n^s - E(\eta_n B) Q_n^s\| + |E\xi_n B - E\eta_n B|. \end{aligned}$$

Here the second term on the right tends to 0 by (ii), and (6) shows that the first term is bounded by

$$\begin{aligned} &\left\| E(\xi_n B) P_n^s - \int_B E\xi_n(dr) P_n^r \right\| + \left\| E(\eta_n B) Q_n^s - \int_B E\eta_n(dr) Q_n^r \right\| \\ &+ \left\| \int_B E\xi_n(dr) P_n^r - \int_B E\eta_n(dr) Q_n^r \right\| \\ &\leq E(\xi_n B) \sup_{r,s \in B} \|P_n^r - P_n^s\| + E(\eta_n B) \sup_{r,s \in B} \|Q_n^r - Q_n^s\| \\ &+ \|E[\xi_n B; \xi_n \in \cdot] - E[\eta_n B; \eta_n \in \cdot]\|, \end{aligned}$$

which tends to 0 by (i)–(iii).  $\square$

For any diffuse and locally finite random measure  $\xi$  on  $\mathbb{R}^d$ , we say that the kernel  $\mathcal{L}[\xi \|\xi^{\otimes n}\|_x]$  is *tight* if  $E[\xi U_x^r \wedge 1 \|\xi^{\otimes n}\|_x] \rightarrow 0$  as  $r \rightarrow 0$  for every  $x \in (\mathbb{R}^d)^{(n)}$ , where  $U_x^r = \bigcup_i B_{x_i}^r$ . In particular,  $E[\xi\{x_i\} \|\xi^{\otimes n}\|_x] = 0$  for all  $i$ . For probability measures on suitable measure spaces  $\mathcal{M}_S$ , weak continuity or convergence is defined with respect to the vague topology; cf. [12, 15]. The following result extends Lemmas 3.4 and 3.5 in [13].

LEMMA 3.12. *Let  $\xi$  be a diffuse random measure on  $S = \mathbb{R}^d$ , such that:*

- (i)  $E\xi^{\otimes n}$  is locally finite on  $S^{(n)}$ ;
- (ii) for any open set  $G \subset S$ , the kernel  $\mathcal{L}[1_G c \xi \|\xi^{\otimes n}\|_x]$  has a version, that is, continuous in total variation in  $x \in G^{(n)}$ ;
- (iii) for any compact set  $K \subset S^{(n)}$ ,

$$\limsup_{r \rightarrow 0} \liminf_{x \in K} \liminf_{\varepsilon \rightarrow 0} \frac{E\xi^{\otimes n} B_x^\varepsilon(\xi U_x^r \wedge 1)}{E\xi^{\otimes n} B_x^\varepsilon} = 0.$$

Then the kernel  $\mathcal{L}[\xi \|\xi^{\otimes n}]_x$  has a tight, weakly continuous version on  $S^{(n)}$ , satisfying the property in (ii) for every  $G$ .

The result remains true with (ii) restricted to open sets  $G$  with  $G^c$  compact. It also extends with the same proof to measure-valued processes  $\xi_t$ , where (i) and (iii) hold uniformly for  $t > 0$  in compacts, and we consider joint continuity in the pair  $(x, t)$ .

PROOF. Proceeding as in [13], we may construct a version of the kernel  $\mathcal{L}[\xi \|\xi^{\otimes n}]_x$  on  $S^{(n)}$  satisfying the continuity in (ii) for any open set  $G$ . Now fix any  $x \in S^{(n)}$ , and let  $l(r)$  denote the lim inf in (iii). By Palm disintegration under condition (i), we may choose some  $x_k \rightarrow x$  in  $S^{(n)}$  such that  $E[\xi U_x^r \wedge 1 \|\xi^{\otimes n}]_{x_k} \leq 2l(r)$ . Then by (ii) we have for any open  $G \subset \mathbb{R}^d$  and  $x \in G^{(n)}$ ,

$$E[\xi(U_x^r \cap G^c) \wedge 1 \|\xi^{\otimes n}]_x = \lim_{k \rightarrow \infty} E[\xi(U_x^r \cap G^c) \wedge 1 \|\xi^{\otimes n}]_{x_k} \leq l(r).$$

Since  $G$  is arbitrary, and  $l(r) \rightarrow 0$  as  $r \rightarrow 0$ , we conclude that  $P[\xi \in \cdot \|\xi^{\otimes n}]_x$  is tight at  $x$ . Using the full strength of (iii), we get in the same way the uniform tightness

$$(7) \quad \limsup_{r \rightarrow 0, x \in K} E[\xi U_x^r \wedge 1 \|\xi^{\otimes n}]_x = 0,$$

for any compact set  $K \subset S^{(n)}$ . For open  $G$  and  $x_k \rightarrow x$  in  $G^{(n)}$ , we get, by (ii),

$$\mathcal{L}[1_{G^c} \xi \|\xi^{\otimes n}]_{x_k} \xrightarrow{w} \mathcal{L}[1_{G^c} \xi \|\xi^{\otimes n}]_x.$$

By a simple approximation based on (7) (cf. Theorem 4.28 in [15] or Theorem 4.9 in [12]), the latter convergence remains valid with  $1_{G^c} \xi$  replaced by  $\xi$ , as required.  $\square$

Next we show how the Palm distributions can be extended, with preserved continuity properties, to conditionally independent random elements.

LEMMA 3.13. *Let  $\xi$  be a random measure on a Polish space  $S$ , consider some random elements  $\alpha$  and  $\beta$  in Borel spaces and choose a kernel  $\nu$  such that  $\nu_{\alpha, \xi} = \mathcal{L}[\beta | \alpha, \xi]$  a.s. Then*

$$(8) \quad \mathcal{L}[\xi, \alpha, \beta \|\xi]_s = \mathcal{L}[\xi, \alpha \|\xi]_s \otimes \nu, \quad s \in S \text{ a.e. } E\xi.$$

When  $\beta \perp\!\!\!\perp_{\alpha} \xi$ , this simplifies to

$$(9) \quad \mathcal{L}[\alpha, \beta \|\xi]_s = \mathcal{L}[\alpha \|\xi]_s \otimes \nu, \quad s \in S \text{ a.e. } E\xi,$$

where  $\nu_{\alpha} = \mathcal{L}[\beta | \alpha]$  a.s. In the latter case, if  $\mathcal{L}[\alpha \|\xi]$  has a version, that is, continuous in total variation, then so does  $\mathcal{L}[\beta \|\xi]$ .

PROOF. First we prove (9) when  $\xi$  is  $\alpha$ -measurable and  $\nu_\alpha = \mathcal{L}[\beta|\alpha]$  a.s. Assuming  $E\xi$  to be  $\sigma$ -finite, we get

$$\begin{aligned} \int E\xi(ds)E[f(s, \alpha, \beta)\|\xi]_s &= E \int \xi(ds) f(s, \alpha, \beta) \\ &= E \int \xi(ds) \int \nu_\alpha(dt) f(s, \alpha, t) \\ &= \int E\xi(ds)E \left[ \int \nu_\alpha(dt) f(s, \alpha, t) \|\xi \right]_s, \end{aligned}$$

and so for  $s \in S$  a.e.  $E\xi$

$$\begin{aligned} \mathcal{L}[\alpha, \beta\|\xi]_s f &= E[f(\alpha, \beta)\|\xi]_s = E \left[ \int \nu_\alpha(dt) f(\alpha, t) \|\xi \right]_s \\ &= \int P[\alpha \in dr\|\xi]_s \int \nu_r(dt) f(s, t) = (\mathcal{L}[\alpha\|\xi]_s \otimes \nu) f, \end{aligned}$$

as required. To obtain (8), it suffices to replace  $\alpha$  in (9) by the pair  $(\xi, \alpha)$ . If  $\beta \perp\!\!\!\perp_\alpha \xi$ , then  $\nu_{\xi, \alpha} = \nu_\alpha$  a.s., and (9) follows from (8). In particular,

$$(10) \quad \mathcal{L}[\beta\|\xi]_s = \mathcal{L}[\nu_\alpha\|\xi]_s, \quad s \in S \text{ a.e. } E\xi.$$

Assuming  $\mathcal{L}[\alpha\|\xi]$  to be continuous in total variation and defining  $\mathcal{L}[\beta\|\xi]$  by (10), we get for any  $s, s' \in S$ ,

$$\begin{aligned} \|\mathcal{L}[\beta\|\xi]_s - \mathcal{L}[\beta\|\xi]_{s'}\| &= \|\mathcal{L}[\nu_\alpha\|\xi]_s - \mathcal{L}[\nu_\alpha\|\xi]_{s'}\| \\ &\leq \|\mathcal{L}[\alpha\|\xi]_s - \mathcal{L}[\alpha\|\xi]_{s'}\|, \end{aligned}$$

and the asserted continuity follows.  $\square$

**4. Moment measures and Palm kernels.** We are now ready to begin our study of DW-processes  $\xi$  in  $\mathbb{R}^d$ , starting from a  $\sigma$ -finite measure  $\mu$  on  $\mathbb{R}^d$ , always assumed to be such that  $\xi_t$  is a.s. locally finite for all  $t > 0$  (cf. Lemma 4.1 below). This condition is clearly stronger than requiring only  $\mu$  to be locally finite. The distribution of  $\xi$  is denoted by  $\mathcal{L}_\mu(\xi) = P_\mu\{\xi \in \cdot\}$ , and we write  $\mathcal{L}_x(\xi) = P_x\{\xi \in \cdot\}$  when  $\mu = \delta_x$ .

We will make constant use of the fact that  $\xi_t$  is infinitely divisible, hence a countable sum of conditionally independent *clusters*, equally distributed apart from shifts and rooted at the points of a Poisson process  $\zeta_0$  of *ancestors* with intensity measure  $t^{-1}\mu$ ; cf. [4, 26]. Indeed, allowing the cluster distribution to be unbounded but  $\sigma$ -finite, we obtain a similar cluster representation of the historical process, and we may introduce an associated *canonical cluster*  $\eta$  with pseudo-distributions  $\mathcal{L}_x(\eta)$ , normalized such that  $P_x\{\eta_t \neq 0\} = t^{-1}$ , where the subscript  $x$  signifies that  $\eta$  starts at  $x \in \mathbb{R}^d$ . For measures  $\mu$  on  $\mathbb{R}^d$  we write  $\mathcal{L}_\mu(\eta) = \int P_x\{\eta \in \cdot\} \mu(dx)$ , and we define  $E_\mu f(\eta)$  accordingly.

By the Markov property of  $\xi$ , we have a similar representation of  $\xi_t$  for every  $s = t - h \in (0, t)$  as a countable sum of conditionally independent  $h$ -clusters (clusters of age  $h$ ), rooted at the points of a Cox process  $\zeta_s$  directed by  $h^{-1}\xi_s$ . In other words,  $\zeta_s$  is conditionally Poisson given  $\xi_s$  with intensity measure  $h^{-1}\xi_s$ .

We first state the criteria for the random measures  $\xi_t$  to be locally finite, quoted from Lemma 3.2 in [19]. Recall that  $p_t$  denotes the continuous density of the symmetric Gaussian distribution on  $\mathbb{R}^d$  with variances  $t > 0$ .

LEMMA 4.1. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with  $\sigma$ -finite initial measure  $\mu$ . Then these conditions are equivalent:*

- (i)  $\xi_t$  is a.s. locally finite for every  $t \geq 0$ ;
- (ii)  $E_\mu \xi_t$  is locally finite for every  $t \geq 0$ ;
- (iii)  $\mu p_t < \infty$  for all  $t > 0$ ;

in which case also

- (iv)  $E_\mu \xi_t = E_\mu \eta_t$  has the continuous density  $\mu * p_t$ .

We turn to the moment measures of a DW-process  $\xi$  with canonical cluster  $\eta$ . Define  $v_t^n = E_0 \eta_t^{\otimes n}$  and  $v_t^J = E_0 \eta_t^{\otimes J}$ , and note that  $v_t = v_t^1 = p_t \cdot \lambda^{\otimes d}$ . Write  $\sum'_{I \subset J}$  for summation over all nonempty, proper subsets  $I \subset J$ . Given any elements  $i \neq j$  in  $J$ , we form a new set  $J_{ij} = J_{ji}$  by combining  $i$  and  $j$  into a single element  $\{i, j\}$ .

The moment measures of a DW-process may be obtained by an initial cluster decomposition, followed by a recursive construction for the individual clusters, as specified by the compact and suggestive formulas below. Some more explicit versions are given after the statement of the theorem.

THEOREM 4.2. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ , and write  $v_t^J = E_0 \eta_t^{\otimes J}$ . Then for any  $t > 0$  and  $\mu$ :*

- (i)  $E_\mu \xi_t^{\otimes n} = \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} (\mu * v_t^J), \quad n \in \mathbb{N};$
- (ii)  $v_t^J = \sum'_{I \subset J} \int_0^t v_s * (v_{t-s}^I \otimes v_{t-s}^{J \setminus I}) ds, \quad |J| \geq 2;$
- (iii)  $v_t^J = \sum_{i \neq j} \int_0^t (v_s^{J_{ij}} * v_{t-s}^{\otimes J}) ds, \quad |J| \geq 2;$
- (iv)  $v_{s+t}^n = \sum_{\pi \in \mathcal{P}_n} \left( v_s^\pi * \bigotimes_{J \in \pi} v_t^J \right), \quad s, t > 0, n \in \mathbb{N}.$

Note that  $*$  denotes convolution in the space variables; (ii) and (iii) also involve convolution in the time variable. To state our more explicit versions of (i)–(iv), let  $f_1, \dots, f_n$  be any nonnegative, measurable functions on  $\mathbb{R}^d$ , and write  $x_J = (x_j; j \in J) \in (\mathbb{R}^d)^J$ . For  $u_{Jij} \in (\mathbb{R}^d)^{Jij}$ , take  $u_k = u_{ij}$  when  $k \in \{i, j\}$ .

$$\begin{aligned}
 \text{(i')} \quad E_\mu \prod_{i \leq n} \xi_t f_i &= \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} \int \mu(du) \int v_t^J(dx_J) \prod_{i \in J} f_i(u + x_i); \\
 \text{(ii')} \quad v_t^J \otimes f_i &= \sum'_{I \subset J} \int_0^t ds \int v_s(du) \int v_{t-s}^I(dx_I) \int v_{t-s}^{J \setminus I}(dx_{J \setminus I}) \\
 &\quad \times \prod_{i \in J} f_i(u + x_i); \\
 \text{(iii')} \quad v_t^J \otimes f_i &= \sum_{i \neq j} \int_0^t ds \int v_s^{Jij}(du_{Jij}) \prod_{k \in J} \int v_{t-s}(dx_k) f_k(u_k + x_k); \\
 \text{(iv')} \quad v_{s+t}^n \otimes f_i &= \sum_{\pi \in \mathcal{P}_n} \int v_s^\pi(du_\pi) \prod_{J \in \pi} \int v_t^J(dx_J) \prod_{i \in J} f_i(u_J + x_i).
 \end{aligned}$$

The cluster decomposition (i) and forward recursion (ii) are implicit in Dynkin [5], who works in a very general setting, using series expansions of Laplace transforms; cf. Section 2.2 in [7]. The backward recursion (iii) and Markov property (iv) are believed to be new. First we prove (i), (ii) and (iv).

PARTIAL PROOF. (i) Use the first assertion in Lemma 3.8.

(ii) In Theorem 1.7 of [5], take  $K = \lambda$ ,  $\psi^t(z) = z^2$ , and  $\eta = \delta_0 \otimes \mu$ , and let  $\Pi_x$  be the distribution of a standard Brownian motion starting at  $x$ . Each term in (ii) appears twice, which accounts for the factor  $q_2^t = 2$  in formula 1.6.B of [5]. Alternatively, we may use a probabilistic approach based on Le Gall’s snake [28], or we may apply Itô’s formula to the martingales  $\xi_s(v_{t-s} * f) - \mu(v_t * f)$ ,  $s \in [0, t]$ , as explained for  $|J| = 2$  in [7], page 39.

(iv) Using (i) repeatedly, along with the Markov property at  $s$ , we get

$$\begin{aligned}
 \sum_{\pi \in \mathcal{P}_n} \otimes_{J \in \pi} (\mu * v_{s+t}^J) &= E_\mu \xi_{s+t}^{\otimes n} = E_\mu E_\mu[\xi_{s+t}^{\otimes n} | \xi_s] = E_\mu E_{\xi_s} \xi_t^{\otimes n} \\
 &= E_\mu \sum_{\pi \in \mathcal{P}_n} \otimes_{J \in \pi} (\xi_s * v_t^J) = \sum_{\pi \in \mathcal{P}_n} E_\mu \xi_s^{\otimes \pi} * \otimes_{J \in \pi} v_t^J \\
 &= \sum_{\pi \in \mathcal{P}_n} \sum_{\kappa \in \mathcal{P}_\pi} \otimes_{I \in \kappa} (\mu * v_s^I) * \otimes_{J \in \pi} v_t^J.
 \end{aligned}$$

Now take  $\mu = c\delta_0$  with  $c > 0$ , divide by  $c$ , and let  $c \rightarrow 0$ .  $\square$

Part (iii) will be deduced from Theorem 4.4 below, which in turn depends on the following discrete constructions. Say that a tree or branch is defined on  $[s, t]$ , if it is

rooted at time  $s$ , and all leaves extend to time  $t$ . It is also said to be *simple* if it has only one leaf, and *binary* if exactly two branches emanate from each vertex. It is further said to be *geometric* if its graph in the plane has no self-intersections. Furthermore, we say that a tree or set of trees is *marked* if distinct marks are assigned to the leaves. A random permutation is called *uniform* if it is exchangeable, and we say that the marks are *random* if they are conditionally exchangeable, given the underlying tree structure. *Siblings* are defined as leaves originating from the same vertex.

LEMMA 4.3. *There are  $n!(n - 1)!2^{1-n}$  marked, binary trees on  $[0, n]$  with distinct splitting times  $1, \dots, n - 1$ . The following constructions are equivalent and give the same probability to all such trees:*

(i) *Forward recursion: Proceed in  $n - 1$  steps, starting from a simple tree on  $[0, 1]$ . After  $k - 1$  steps, we have a binary tree on  $[0, k]$  with  $k$  leaves and distinct splitting times  $1, \dots, k - 1$ . Now divide a randomly chosen leaf into two, and extend all leaves to time  $k + 1$ . After the final step, attach random marks to the leaves.*

(ii) *Backward recursion: Proceed in  $n - 1$  steps, starting from  $n$  simple, marked trees on  $[n - 1, n]$ . After  $k - 1$  steps, we have  $n - k + 1$  binary trees on  $[n - k, n]$  with totally  $n$  leaves and distinct splitting times  $n - k + 1, \dots, n - 1$ . Now join two randomly chosen roots, and extend all roots to time  $n - k - 1$ . Continue until all roots are connected.*

(iii) *Sideways recursion: Let  $\tau_1, \dots, \tau_{n-1}$  be a uniform permutation of  $1, \dots, n - 1$ . Proceed in  $n - 1$  steps, starting from a simple tree on  $[0, n]$ . After  $k - 1$  steps, we have a binary tree on  $[0, n]$  with  $k$  leaves and distinct splitting times  $\tau_1, \dots, \tau_{k-1}$ . Now attach a new branch on  $[\tau_k, n]$  to the last available path. After the final step, attach random marks to the leaves.*

PROOF. (ii) By the obvious one-to-one correspondence between marked trees and selections of pairs, this construction gives the same probability to all possible trees. The total number of choices is clearly

$$\binom{n}{2} \binom{n-1}{2} \dots \binom{2}{2} = \frac{n(n-1)}{2} \frac{(n-1)(n-2)}{2} \dots \frac{2 \cdot 1}{2} = \frac{n!(n-1)!}{2^{n-1}}.$$

(i) Before the final marking of leaves, the resulting tree can be realized as a geometric one, which yields a one-to-one correspondence between the  $(n - 1)!$  possible constructions and the set of all geometric trees. Now any binary tree with  $n$  distinct splitting times and with  $m$  pairs of siblings can be realized as a geometric tree in  $2^{n-m-1}$  different ways. Furthermore, any geometric tree with  $n$  leaves and  $m$  pairs of siblings can be marked in  $n!2^{-m}$  nonequivalent ways. Hence, any given tree of this type has probability

$$\frac{2^{n-m-1}}{(n-1)!} \cdot \frac{2^m}{n!} = \frac{2^{n-1}}{n!(n-1)!},$$

which is independent of  $m$  and hence is the same for all trees. (Note that this agrees with the probability in (ii).)

(iii) Before the final marking, this construction yields a binary, geometric tree with distinct splitting times  $1, \dots, n$ . Conversely, any geometric tree can be realized in this way for a suitable permutation  $\tau_1, \dots, \tau_{n-1}$  of  $1, \dots, n$ . The correspondence is one-to-one, since both sets of trees have the same cardinality  $(n - 1)!$ . The proof may now be completed as in case (i).  $\square$

Given a uniform, discrete random tree, as described in Lemma 4.3, we may form a *uniform, marked, Brownian tree* in  $\mathbb{R}^d$  on the time interval  $[0, t]$  by a suitable choice of random splitting times and spatial motion. Note that this uniform tree is entirely different from the historical Brownian tree considered in [26] or in Section 3.4 of [7]. Elaborating on the insight of Etheridge [7], Sections 2.1–2, we show how the moment measures of a single cluster admit a probabilistic interpretation in terms of such a tree.

**THEOREM 4.4.** *Form a marked, binary random tree in  $\mathbb{R}^d$ , rooted at the origin at time 0, with  $n$  leaves extending to time  $t > 0$ , with branching structure as in Lemma 4.3, with splitting times  $\tau_1, \dots, \tau_{n-1}$  given by an independent, uniform binomial process on  $[0, t]$ , and with spatial motion given by independent Brownian motions along the branches. Then the joint distribution  $\mu_t^n$  of the leaves at time  $t$  and the cluster moment measure  $\nu_t^n$  in Theorem 4.2 are related by  $\nu_t^n = n!t^{n-1}\mu_t^n$ .*

**PROOF.** The assertion is obvious for  $n = 1$ . Proceeding by induction, assume that the statement holds for trees up to order  $n - 1$ , where  $n \geq 2$ , and turn to trees of order  $n$  marked by  $J = \{1, \dots, n\}$ . For any  $I \subset J$  with  $|I| = k \in [1, n)$ , Lemma 4.3 shows that the number of marked, discrete trees of order  $n$  such that  $J$  first splits into  $I$  and  $J \setminus I$  equals

$$\begin{aligned} &k!(k - 1)!2^{1-k}(n - k)!(n - k - 1)!2^{1-n+k} \binom{n - 2}{k - 1} \\ &= (n - 2)!k!(n - k)!2^{2-n}, \end{aligned}$$

where the last factor on the left arises from the choice of  $k - 1$  splitting times for the  $I$ -component, among the remaining  $n - 2$  splitting times for the original tree. Since the total number of trees is  $n!(n - 1)!2^{1-n}$ , the probability that  $J$  first splits into  $I$  and  $J \setminus I$  equals

$$\frac{(n - 2)!k!(n - k)!2^{2-n}}{n!(n - 1)!2^{1-n}} = \frac{2}{n - 1} \binom{n}{k}^{-1}.$$

Since all genealogies are equally likely, the discrete subtrees marked by  $I$  and  $J \setminus I$  are conditionally independent and uniformly distributed, and the remaining branching times  $2, \dots, n - 1$  are divided uniformly between the two trees. Since

the splitting times  $\tau_1 < \dots < \tau_{n-1}$  of the continuous tree form a uniform binomial process on  $[0, t]$ , Lemma 2.8 shows that  $\mathcal{L}(\tau_1)$  has density  $(n-1)(t-s)^{n-2}t^{1-n}$ , whereas  $\tau_2, \dots, \tau_{n-1}$  form a binomial process on  $[\tau_1, t]$ , conditionally on  $\tau_1$ . Furthermore, Lemma 2.6 shows that the splitting times of the two subtrees form independent binomial processes on  $[\tau_1, t]$ , conditionally on  $\tau_1$  and the initial split of  $J$  into  $I$  and  $J \setminus I$ . Combining these facts with the conditional independence of the spatial motion, we see that the continuous subtrees marked by  $I$  and  $J \setminus I$  are conditionally independent uniform Brownian trees on  $[\tau_1, t]$ , given the spatial motion up to time  $\tau_1$  and the split at time  $\tau_1$  of the original index set into  $I$  and  $J \setminus I$ .

Conditioning as indicated and using the induction hypothesis and Theorem 4.2(ii), we get

$$\begin{aligned} \mu_t^n &= t^{1-n} \sum'_{I \subset J} \binom{n}{k}^{-1} \int_0^t (t-s)^{n-2} \mu_s * (\mu_{t-s}^I \otimes \mu_{t-s}^{J \setminus I}) ds \\ &= t^{1-n} \sum'_{I \subset J} \binom{n}{k}^{-1} \frac{1}{k!(n-k)!} \int_0^t v_s * (v_{t-s}^I \otimes v_{t-s}^{J \setminus I}) ds \\ &= \frac{t^{1-n}}{n!} \sum'_{I \subset J} \int_0^t v_s * (v_{t-s}^I \otimes v_{t-s}^{J \setminus I}) ds = \frac{t^{1-n}}{n!} v_t^n, \end{aligned}$$

where  $|I| = k$ . Note that a factor 2 cancels out in the first step, since every partition  $\{I, J \setminus I\}$  is counted twice. This completes the induction.  $\square$

PROOF OF THEOREM 4.2(III). Let  $\tau_1 < \dots < \tau_{n-1}$  denote the splitting times of the Brownian tree in Theorem 4.4. By Lemma 2.8,  $\tau_{n-1}$  has density  $(n-1)s^{n-2}t^{1-n}$  (in  $s$  for fixed  $t$ ), and given  $\tau_{n-1}$  the remaining times  $\tau_1, \dots, \tau_{n-2}$  form a uniform binomial process on  $[0, \tau_{n-1}]$ . By Lemma 4.3 the entire structure up to time  $\tau_{n-1}$  is then conditionally a uniform Brownian tree of order  $n-1$ , independent of the last branching and the motion up to time  $t$ . Defining  $\mu_t^J$  as before and conditioning on  $\tau_{n-1}$ , we get

$$\mu_t^J = \binom{n}{2}^{-1} \sum_{\{i,j\} \subset J} \int_0^t (n-1) \frac{s^{n-2}}{t^{n-1}} (\mu_s^{J_{ij}} * \mu_{t-s}^{\otimes J}) ds.$$

By Theorem 4.4 we may substitute

$$\mu_t^J = \frac{v_t^J}{n!t^{n-1}}, \quad \mu_s^{J_{ij}} = \frac{v_s^{J_{ij}}}{(n-1)!s^{n-2}}, \quad \mu_{t-s} = v_{t-s},$$

and the assertion follows. Here again a factor 2 cancels out in the last computation, since every pair  $\{i, j\}$  appears twice in the summation  $\sum_{i,j \in J}$ .  $\square$

To describe the Palm distributions of a single cluster, we begin with some basic properties of the Brownian excursion and snake. Given a process  $X$  in a space

$S$  and some random times  $\sigma \leq \tau$ , we define the *restriction* of  $X$  to  $[\sigma, \tau]$  as the process  $Y_s = X_{\sigma+s}$  for  $s \leq \tau - \sigma$  and  $Y_s = \Delta$  for  $s > \tau - \sigma$ , where  $\Delta \notin S$ . By a *Markov time* for  $X$  we mean a random time  $\sigma$ , such that the restrictions of  $X$  to  $[0, \sigma]$  and  $[\sigma, \infty]$  are conditionally independent, given  $X_\sigma$ . We quote some distributional facts for the Brownian excursion, first noted by Williams [36]; cf. [25].

LEMMA 4.5. *Given a Brownian excursion  $X$ , conditioned to reach height  $t > 0$ , let  $\sigma$  and  $\tau$  be the first and last times that  $X$  visits  $t$ , and write  $\rho$  for the first time  $X$  attains its minimum on  $[\sigma, \tau]$ . Then  $X_\rho$  is  $U(0, t)$ , and  $\sigma, \rho$  and  $\tau$  are Markov times for  $X$ .*

Some induced properties of the Brownian snake are implicit in Le Gall [26, 28]:

LEMMA 4.6. *Given  $X, \sigma, \rho$  and  $\tau$  as in Lemma 4.5, let  $Y$  be a Brownian snake with contour process  $X$ . Then  $Y_\sigma$  and  $Y_\tau$  are Brownian motions on  $[0, t]$  extending  $Y_\rho$  on  $[0, X_\rho]$ , both are independent of  $X_\rho$ , and  $\sigma, \rho$  and  $\tau$  are Markov times for  $Y$ .*

PROOF. The corresponding properties are easily verified for the approximating discrete snake based on a simple random walk (cf. [7], Section 3.6), and they extend in the limit to the continuous snake. Alternatively, we may approximate the Brownian snake, as in [26], by a discrete tree under the Brownian excursion, for which the corresponding properties are again obvious.  $\square$

We now form an *extended Brownian excursion*, generating an *extended Brownian snake*, related to the uniform Brownian tree in Theorem 4.4. The unmarked tree, constructed as in Lemma 4.3(iii) though with Brownian spatial motion and with  $\tau_1, \dots, \tau_{n-1}$  chosen to be i.i.d.  $U(0, t)$ , is referred to below as a *discrete Brownian snake* on  $[0, t]$  of order  $n$ .

LEMMA 4.7. *Given a Brownian excursion  $X$ , conditioned to reach height  $t > 0$ , form  $X^n$  by inserting  $n$  independent copies of the path between the first and last visits to  $t$ , let  $\tau_1 < \dots < \tau_n$  be the connection times of those  $n + 1$  paths, and form a Brownian snake  $Y^n$  with contour process  $X^n$ . Then the paths  $Y_{\tau_1}^n, \dots, Y_{\tau_n}^n$  form a discrete Brownian snake on  $[0, t]$ .*

PROOF. Use Lemma 4.6 and its proof.  $\square$

The extended Brownian snake  $Y^n$  generates a measure-valued process  $\eta_n$ , in the same way as the ordinary snake  $Y$  generates a single cluster  $\eta$ ; cf. [26, 28] or [7], page 69. We show how the  $n$ th order Palm distributions with respect to  $\eta_t$  can be obtained from  $\eta_n$  by suitable conditioning. The “cluster terms” in Lemma 3.8 can then be obtained by simple averaging, based on the elementary Lemma 3.9.

**THEOREM 4.8.** *Let  $Y_n$  be an extended Brownian snake with connection times  $\tau_1, \dots, \tau_n$ , generating a measure-valued process  $\eta_n$ . Choose an independent, uniform permutation  $\pi$  of  $1, \dots, n$ , and define  $\beta_k = Y_n(\tau_{\pi_k}, t)$ ,  $k \leq n$ . Then for any initial measure  $\mu$  on  $\mathbb{R}^d$ , the  $n$ th order Palm distributions of  $\eta$  with respect to  $\eta_t$  are given a.e.  $\lambda^{\otimes nd}$  by*

$$(11) \quad \mathcal{L}_\mu[\eta \parallel \eta_t^{\otimes n}]_\beta = \mathcal{L}_\mu[\eta_n \mid \beta] \quad a.s.$$

**PROOF.** Let  $\tau_0$  and  $\tau_1$  be the first and last times that  $X$  visits  $t$ , and let  $\tau_0, \dots, \tau_{n+1}$  be the endpoints of the corresponding  $n + 1$  paths for the extended process  $X_n$ . Introduce the associated local times  $\sigma_0, \dots, \sigma_{n+1}$  of  $X_n$  at  $t$ , so that  $\sigma_0 = 0$  and the differences  $\sigma_k - \sigma_{k-1}$  are independent and exponentially distributed with rate  $c = t^{-1}$ . Writing  $\sigma \circ \pi = (\sigma_{\pi_1}, \dots, \sigma_{\pi_n})$ , we get, by Lemma 2.9,

$$(12) \quad Ef(\sigma \circ \pi, \sigma_{n+1}) = \frac{c^n}{n!} E \int_{[0, \sigma_1]^n} f(s, \sigma_1) ds.$$

By excursion theory (cf. [15], pages 432–442), the shifted path  $\theta_{\tau_0} X$  on  $[0, \tau_1 - \tau_0]$  is generated by a Poisson process  $\zeta \perp\!\!\!\perp \sigma_1$  of excursions from  $t$ , restricted to the set of paths not reaching level 0. By the strong Markov property, we can use the same process  $\zeta$  to encode the excursions of  $X_n$  on the extended interval  $[\tau_0, \tau_{n+1}]$ , provided that we choose  $\zeta \perp\!\!\!\perp (\sigma_1, \dots, \sigma_{n+1})$ . By Lemma 4.5, the restrictions of  $X$  to the intervals  $[0, \tau_0]$  and  $[\tau_1, \infty]$  are independent of the intermediate path, and the corresponding property holds for the restrictions of  $X_n$  to  $[0, \tau_0]$  and  $[\tau_{n+1}, \infty]$ , by the construction in Lemma 4.7. For convenience we may extend  $\zeta$  to a point process  $\zeta'$  on  $[0, \infty]$ , using points at 0 and  $\infty$  to encode the initial and terminal paths of  $X$  or  $X_n$ . From (12) we get, by independence,

$$(13) \quad Ef(\sigma \circ \pi, \sigma_{n+1}, \zeta') = \frac{c^n}{n!} E \int_{[0, \sigma_1]^n} f(s, \sigma_1, \zeta') ds.$$

The inverse local time processes  $T$  of  $X$  and  $T_n$  of  $X_n$  are obtained from the pairs  $(\sigma_1, \zeta')$  or  $(\sigma_{n+1}, \zeta')$ , respectively, by a common measurable construction, and we note that  $T(\sigma_k) = \tau_k$  for  $k = 0, 1$  and  $T_n(\sigma_k) = \tau_k$  for  $k = 0, \dots, n + 1$ . Furthermore,  $\xi = \lambda \circ T^{-1}$  and  $\xi_n = \lambda \circ T_n^{-1}$  are the local time random measures of  $X$  and  $X_n$ , respectively, at height  $t$ . Since the entire excursions  $X$  and  $X_n$  may be recovered from the same pairs by a common measurable mapping, we get, from (13),

$$(14) \quad Ef(\tau \circ \pi, X_n) = \frac{c^n}{n!} E \int_{[0, \sigma_1]^n} f(T \circ s, X) ds = \frac{c^n}{n!} E \int f(r, X) \xi^{\otimes n}(dr),$$

where  $T \circ s = (T_{s_1}, \dots, T_{s_n})$ , and the second equality holds by the substitution rule for Lebesgue–Stieltjes integrals.

Now introduce some random snakes  $Y$  and  $Y_n$  with contour processes  $X$  and  $X_n$ , respectively, with initial distribution  $\mu$ , and with Brownian spatial motion in  $\mathbb{R}^d$ . By Le Gall’s path-wise construction of the snake [26], or alternatively by the discrete approximation described in [7], the conditional distributions  $\mathcal{L}_\mu[Y|X]$  and

$\mathcal{L}_\mu[Y_n|X_n]$  are given by a common probability kernel. The same constructions justify the conditional independence  $Y_n \perp\!\!\!\perp_{X_n}(\tau \circ \pi)$ , and so, by (14),

$$E_\mu f(\tau \circ \pi, Y_n) = \frac{c^n}{n!} E_\mu \int f(r, Y) \xi^{\otimes n}(dr).$$

Since  $\beta_k = Y_n(\tau \circ \pi_k, t)$  for all  $k$ , and  $\eta_t$  is the image of  $\xi$  under the map  $Y(\cdot, t)$ , the substitution rule for integrals yields

$$E_\mu f(\beta, Y_n) = \frac{c^n}{n!} E_\mu \int f(Y(r, t), Y) \xi^{\otimes n}(dr) = \frac{c^n}{n!} E_\mu \int f(x, Y) \eta_t^{\otimes n}(dx).$$

Finally, the entire clusters  $\eta$  and  $\eta_n$  are generated by  $Y$  and  $Y_n$ , respectively, through a common measurable mapping, and so

$$E_\mu f(\beta, \eta_n) = \frac{c^n}{n!} E_\mu \int f(x, \eta) \eta_t^{\otimes n}(dx),$$

which extends by monotone convergence to any initial measure. The assertion now follows by direct disintegration, or by the conditioning approach to Palm distributions described in Section 3.  $\square$

**5. Moment densities.** Here we collect some technical estimates and continuity properties for the moment densities of a DW-process, useful in subsequent sections. We begin with a result for general Brownian trees, defined as random trees in  $\mathbb{R}^d$  with spatial motion given by independent Brownian motions. For  $x \in (\mathbb{R}^d)^n$ , let  $r_x$  denote the distance from  $x$  to the diagonal set  $D_n = ((\mathbb{R}^d)^n)^c$ .

LEMMA 5.1. *For any marked Brownian tree on  $[0, s]$  with  $n$  leaves and paths in  $\mathbb{R}^d$ , the joint distribution at time  $s$  has a continuous density  $q$  on  $(\mathbb{R}^d)^n$  satisfying*

$$q(x) \leq (1 \vee t d n^2 r_x^{-2})^{nd/2} p_{nt}^{\otimes nd}(x), \quad x \in (\mathbb{R}^d)^n, s \leq t.$$

PROOF. Conditionally on tree structure and splitting times, the joint distribution is a convolution of centered Gaussian distributions  $\mu_1, \dots, \mu_n$ , supported by some linear subspaces  $S_1 \subset \dots \subset S_n = \mathbb{R}^{nd}$  of dimensions  $d, 2d, \dots, nd$ . The tree structure is specified by a nested sequence of partitions  $\pi_1, \dots, \pi_n$  of the index set  $\{1, \dots, n\}$ , and we write  $h_1, \dots, h_n$  for the times between branchings. Then Lemma 2.1 shows that  $\mu_k$  has principal variances  $h_k|J|$ ,  $J \in \pi_k$ , each with multiplicity  $d$ . Writing  $v_t = p_t \cdot \lambda^{\otimes d}$  and noting that  $|J| \leq n - k + 1$  for  $J \in \pi_k$ , we get, by Lemma 2.2,

$$\mu_k \leq (n - k + 1)^{(k-1)d/2} v_{(n-k+1)h_k}^{kd} \otimes \delta_0^{\otimes (n-k)d}, \quad k \leq n.$$

Putting

$$c = \prod_{k \leq n} (n - k + 1)^{(k-1)d/2} \leq n^{n^2 d/2},$$

$$s_k = (n - k + 1)h_k, \quad t_k = s_k + \dots + s_n, \quad k \leq n,$$

we get

$$c^{-1} \underset{k \leq n}{(*)} \mu_k \leq \underset{k \leq n}{(*)} (v_{s_k}^{kd} \otimes \delta_0^{\otimes(n-k)d}) = \underset{k \leq n}{\bigotimes} \underset{j \geq k}{(*)} v_{s_j}^d = \underset{k \leq n}{\bigotimes} v_{t_k}^d.$$

Consider the orthogonal decomposition  $x = x_1 + \dots + x_n$  in  $\mathbb{R}^{nd}$  with  $x_k \in S_k \ominus S_{k-1}$ , and write  $x' = x - x_n$ . Since  $|x_n|$  equals the orthogonal distance of  $x$  to the subspace  $S_{n-1} \subset D_n$ , we get  $|x_n| \geq r_x$ . Using Lemma 2.3 and noting that  $h_n = t_n \leq t_k \leq nt$ , we see that the continuous density of  $c^{-1}(\mu_1 * \dots * \mu_n)$  at  $x$  is bounded by

$$\begin{aligned} \prod_{k \leq n} p_{t_k}^{\otimes d}(x_k) &= \prod_{k \leq n} (2\pi t_k)^{-d/2} e^{-|x_k|^2/2t_k} \\ &\leq (2\pi h_n)^{-nd/2} e^{-|x_n|^2/2h_n} \prod_{k < n} e^{-|x_k|^2/2nt} \\ &= p_{h_n}^{\otimes nd}(|x_n|) e^{-|x'|^2/2nt} \\ &\leq (1 \vee tdn^2|x_n|^{-2})^{nd/2} p_{nt}^{\otimes nd}(|x_n|) e^{-|x'|^2/2nt} \\ &\leq (1 \vee tdn^2r_x^{-2})^{nd/2} p_{nt}^{\otimes nd}(x), \end{aligned}$$

where  $|x_n|$  also denotes the vector  $(|x_n|, 0, \dots, 0)$ . Since the right-hand side is independent of branching structure and splitting times, the unconditional density  $q(x)$  has the same bound, and the desired estimate follows. The stated continuity follows by dominated convergence from the continuity of the normal density.  $\square$

This yields a useful estimate for the moment densities of a single cluster.

LEMMA 5.2. *For a DW-process in  $\mathbb{R}^d$ , the cluster moment measures  $v_t^n = E_0 \eta_t^{\otimes n}$  have densities  $q_t^n(x)$  that are jointly continuous in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  and satisfy the uniform bounds*

$$\sup_{s \leq t} q_s^n(x) \leq (1 \vee r_x^{-2}t)^{nd/2} p_{nt}^{\otimes n}(x), \quad x \in (\mathbb{R}^d)^{(n)}, t > 0.$$

PROOF. By Theorem 4.4 it is equivalent to consider the joint endpoint distribution of a uniform,  $n$ th order Brownian tree in  $\mathbb{R}^d$  on the interval  $[0, t]$ , where the stated estimate holds by Lemma 5.1. To prove the asserted continuity, we may condition on tree structure and splitting times to get a nonsingular Gaussian distribution, for which the assertion is obvious. The unconditional statement then follows by dominated convergence, based on the uniform bound in Lemma 5.1.  $\square$

We proceed with a density version of Theorem 4.2.

**THEOREM 5.3.** *For a DW-process in  $\mathbb{R}^d$  with initial measure  $\mu \neq 0$ , the moment measures  $E_\mu \xi_t^{\otimes n}$  and  $\nu_t^n = E_0 \eta_t^{\otimes n}$  have positive, jointly continuous densities on  $(\mathbb{R}^d)^{(n)} \times (0, \infty)$ , satisfying density versions of the identities in Theorem 4.2(i)–(iv).*

**PROOF.** Let  $q_t^n$  denote the jointly continuous densities of  $\nu_t^n$  obtained in Lemma 5.2. As mixtures of normal densities, they are again strictly positive. Inserting the versions  $q_t^n$  into the convolution formulas of Theorem 4.2(i), we get some strictly positive densities of the measures  $E_\mu \xi_t^{\otimes n}$ , and the joint continuity of those densities follows by extended dominated convergence (cf. [15], page 12) from the estimates in Lemma 5.2 and the joint continuity in Lemma 2.5.

Inserting the continuous densities  $q_t^n$  into the expressions on the right of Theorem 4.2(ii)–(iv), we obtain densities of the measures on the left. If the latter functions can be shown to be continuous on  $(\mathbb{R}^d)^{(J)}$  or  $(\mathbb{R}^d)^{(n)}$ , respectively, they must agree with the continuous densities  $q_t^J$  or  $q_{s+t}^n$ , and the desired identities follow. By Lemma 5.2 and extended dominated convergence, it is enough to prove the required continuity with  $q_s$  and  $q_s^n$  replaced by the normal densities  $p_t$  and  $p_{nt}^{\otimes n}$ , respectively. Hence, we need to show that the convolutions  $p_t * p_{nt}^{\otimes n}$ ,  $p_{(n-1)t}^{\otimes(n-1)} * p_{nt}^{\otimes n}$  and  $p_{nt}^{\otimes \pi} * p_{nt}^{\otimes n}$  are continuous, which is clear since they are all nonsingular Gaussian.  $\square$

We turn to the conditional moment densities.

**THEOREM 5.4.** *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with  $\xi_0 = \mu$ . Then for every  $n$  there exist some processes  $M_s^t$  on  $\mathbb{R}^{nd}$ ,  $0 \leq s < t$ , such that:*

- (i)  $E_\mu[\xi_t^{\otimes n} | \xi_s] = M_s^t \cdot \lambda^{\otimes nd}$  a.s.,  $0 \leq s < t$ ;
- (ii)  $M_s^t(x)$  is a martingale in  $s \in [0, t)$  for fixed  $x \in (\mathbb{R}^d)^{(n)}$  and  $t > 0$ ;
- (iii)  $M_s^t(x)$  is continuous, a.s. and in  $L^1$ , in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (s, \infty)$  for fixed  $s \geq 0$ .

**PROOF.** Write  $S_n = (\mathbb{R}^d)^{(n)}$ , let  $q_t^n$  denote the continuous densities in Lemma 5.2 and let  $x_J$  be the projection of  $x \in \mathbb{R}^{nd}$  onto  $(\mathbb{R}^d)^J$ . By the Markov property of  $\xi$  and Theorem 4.2(i), the random measures  $E_\mu[\xi_t^{\otimes n} | \xi_s]$  have a.s. densities

$$(15) \quad M_s^t(x) = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} (\xi_s * q_{t-s}^J)(x_J), \quad x \in S_n,$$

which are a.s. continuous in  $(x, t) \in S_n \times (s, \infty)$  for fixed  $s \geq 0$  by Theorem 5.3. Indeed, the previous theory applies with  $\mu$  replaced by  $\xi_s$ , since  $E_\mu \xi_s p_t = \mu p_{s+t} < \infty$  and hence  $\xi_s p_t < \infty$  for every  $t > 0$  a.s.

To prove the  $L^1$ -continuity in (iii), it suffices, by Lemma 1.32 in [15], to show that  $E_\mu M_s^t(x)$  is continuous in  $(x, t) \in S_n \times (s, \infty)$ . By Lemma 5.2 and extended

dominated convergence, it is then enough to prove the a.s. and  $L^1$  continuity in  $x \in S_n$  alone, for the processes in (15) with  $q_{t-s}^J$  replaced by  $p_t^{\otimes J}$ . Here the a.s. convergence holds by Lemma 2.5, and so by Theorem 4.2(i) it remains to show that  $\mu * q_s^n * p_t^{\otimes n}$  is continuous on  $S_n$  for fixed  $s, t, \mu$  and  $n$ . Since  $q_s^n * p_t^{\otimes n} = v_s^n * p_t^{\otimes n}$  is continuous on  $S_n$ , by Theorem 4.4 and Lemma 5.1, it suffices, by Lemma 5.2 and extended dominated convergence, to show that  $\mu * p_t^{\otimes n}$  is continuous on  $\mathbb{R}^{nd}$  for fixed  $t, \mu$  and  $n$ , which holds by Lemma 2.5.

To prove (ii), let  $B \subset \mathbb{R}^{nd}$  be measurable, and note that

$$\begin{aligned} \lambda^{\otimes nd}[M_0^t; B] &= E_\mu \xi_t^{\otimes n} B = E_\mu E_\mu[\xi_t^{\otimes n} B | \xi_s] \\ &= E_\mu \lambda^{\otimes nd}[M_s^t; B] = \lambda^{\otimes nd}[E_\mu M_s^t; B], \end{aligned}$$

which implies  $M_0^t = E_\mu M_s^t$  a.e. Since both sides are continuous on  $S_n$ , they agree identically on the same set, and so by (15)

$$E_\mu \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} (\xi_s * q_{t-s}^J)(x_J) = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} (\mu * q_t^J)(x_J), \quad s < t.$$

Replacing  $\mu$  by  $\xi_r$  for arbitrary  $r > 0$  and using the Markov property at  $r$ , we obtain

$$E_\mu[M_{r+s}^{r+t}(x) | \xi_r] = M_r^{r+t}(x) \quad \text{a.s., } x \in S_n, r > 0, 0 \leq s < t,$$

which yields the martingale property in (ii).  $\square$

We turn to a simple truncation property of the conditional densities.

LEMMA 5.5. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with initial measure  $\mu$ , fix some disjoint, open sets  $B_1, \dots, B_n \subset \mathbb{R}^d$  and put  $B = \bigcup_k B_k$  and  $U = \bigcup_k B_k$ . Then as  $h \rightarrow 0$  we have, uniformly for  $(x, t) \in B \times (0, \infty)$  in compacts,*

$$E_\mu \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} (1_{U^c} \xi_{t-h} * q_h^J)(x_J) \rightarrow 0.$$

PROOF. Writing  $t = s + h$  and using the notation and results of Theorem 5.4, we see that the left-hand side is bounded by  $E_\mu M_s^t(x) = M_0^t(x)$ . By Lemma 5.2, this is locally bounded by a sum of products of convolutions  $\mu * p_t^{\otimes J}(x_J)$ , and Lemma 2.4 yields a similar uniform bound, valid in some neighborhood of every fixed pair  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$ . Letting  $\mu \downarrow 0$  locally and using Lemma 2.5 and dominated convergence, we get  $E_\mu M_s^t(x) \rightarrow 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. This reduces the proof to the case of bounded  $\mu$ . We may then estimate the expression on the left by

$$\sum_{\pi \in \mathcal{P}_n} \|E_\mu \xi_{t-h}^{\otimes \pi}\| \prod_{J \in \pi} \sup_{u \in U^c} q_h^J(x_J - u).$$

By Theorems 4.2(i) and 4.4 the norms  $\|E_\mu \xi_{t-h}^{\otimes \pi}\|$  are bounded for bounded  $t - h$ . Furthermore, Lemma 5.2 shows that the functions  $q_h^J$  may be estimated by the corresponding normal densities  $p_h^{\otimes J}$ , for which the desired uniform convergence is obvious.  $\square$

We need some more precise estimates of the moment densities near the diagonals. Here  $q_{\mu,t}^n$  denotes the continuous density of  $E_\mu \xi_t^{\otimes n}$  in Theorem 5.3.

LEMMA 5.6. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ . Then  $E_\mu \xi_s^{\otimes n} * p_h^{\otimes n}$  is continuous on  $\mathbb{R}^{nd}$  and such that for fixed  $t > 0$ , uniformly on  $\mathbb{R}^{nd}$  and in  $s \leq t, h > 0$  and  $\mu$ ,*

$$E_\mu \xi_s^{\otimes n} * p_h^{\otimes n} \leq (1 \vee h^{-1}t)^{nd/2} \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} (\mu * p_{nt+h}^{\otimes J}) < \infty.$$

Furthermore,  $E_\mu \xi_s^{\otimes n} * p_h^{\otimes n} \rightarrow q_{\mu,t}^n$  on  $(\mathbb{R}^d)^{(n)}$  as  $s \rightarrow t$  and  $h \rightarrow 0$ .

PROOF. By Theorem 4.2(i) it suffices to show that  $v_s^n * p_h^{\otimes n} \leq (1 \vee h^{-1}t) \times p_{nt+h}^{\otimes n}$ , uniformly for  $s \leq t$ . By Theorem 4.4 we may replace  $v_s^n$  by the distribution of the endpoint vector  $\gamma_s^n$  of a uniform Brownian tree. Conditioning on tree structure and splitting times, we see from Lemma 2.1 that  $\gamma_s^n$  becomes centered Gaussian with principal variances bounded by  $nt$ . Convolving with  $p_h^{\otimes n}$  gives a centered Gaussian density with principal variances in  $[h, nt + h]$ , and Lemma 2.2 yields the required bound for the latter density in terms of the rotationally symmetric version  $p_{nt+h}^{\otimes n}$ . Taking expected values yields the corresponding unconditional bound. The asserted continuity may now be proved as in case of Lemma 2.5.

To prove the last assertion, consider first the corresponding statement for a single cluster. Here both sides are mixtures of similar normal densities, obtained by conditioning on splitting times and branching structure in the equivalent Brownian trees of Theorem 4.4, and the statement results from an elementary approximation of the uniform binomial process on  $[0, t]$  by a similar process on  $[0, s]$ . The general result now follows by dominated convergence from the density version of Theorem 4.2(i) established in Theorem 5.3.  $\square$

To state the next result, we use for  $x = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n$  and  $k \in [1, n]$  the notation  $x^k = (x_1, \dots, x_k)$ .

LEMMA 5.7. *For any  $\mu$  and  $1 \leq k \leq n$  we have, uniformly for  $0 < h \leq r \leq (t \wedge \frac{1}{2})$  and  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts,*

$$(E_\mu \xi_t^{\otimes(n+k)} * (p_h^{\otimes n} \otimes p_r^{\otimes k}))(x, x^k) \leq \begin{cases} r^{k(1-d/2)}, & d \geq 3, \\ |\log r|^k, & d = 2. \end{cases}$$

PROOF. First we prove a similar estimate for the moment measures  $v_t^{n+k}$  of a single cluster. By Theorem 4.4 it is equivalent to consider the distribution  $\mu_t^{n+k}$  for the endpoint vector  $(\gamma_1, \dots, \gamma_{n+k})$  of a uniform Brownian tree on  $[0, t]$ . Then let  $\tau_i$  and  $\alpha_i$  be the time and place where leaf number  $n + i$  is attached, and put  $\tau = (\tau_i)$  and  $\alpha = (\alpha_i)$ . Let  $\mu_{t|\tau}^n$  and  $\mu_{t|\tau,\alpha}^n$  denote the conditional distributions of  $(\gamma_1, \dots, \gamma_n)$ , given  $\tau$  or  $(\tau, \alpha)$ , respectively, and put  $u = t + r$ . Then we have, uniformly for  $h$  and  $r$  as above and  $x \in (\mathbb{R}^d)^{(n)}$ ,

$$\begin{aligned} (\mu_t^{n+k} * (p_h^{\otimes n} \otimes p_r^{\otimes k}))(x, x^k) &= E(\mu_{t|\tau,\alpha}^n * p_h^{\otimes n})(x) \prod_{i \leq k} p_{u-\tau_i}(x_i - \alpha_i) \\ &\leq E(\mu_{t|\tau,\alpha}^n * p_h^{\otimes n})(x) \prod_{i \leq k} (u - \tau_i)^{-d/2} \\ &\leq q_u^n(x) \left( \int_r^u s^{-d/2} ds \right)^k \leq q_u^n(x) \begin{cases} r^{k(1-d/2)}, \\ |\log r|^k, \end{cases} \end{aligned}$$

when  $d \geq 3$  or  $d = 2$ , respectively. Here the first equality holds since  $(\gamma_1, \dots, \gamma_n)$  and  $\gamma_{n+1}, \dots, \gamma_{n+k}$  are conditionally independent, given  $\tau$  and  $\alpha$ . The second relation holds since  $\|p_r\| \leq r^{-d/2}$ . We may now use the chain rule for conditional expectations to replace  $\mu_{t|\tau,\alpha}^n$  by  $\mu_{t|\tau}^n$ . Next we apply Lemma 2.7 twice, first to replace  $\mu_{t|\tau}^n$  by  $\mu_t^n$ , then to replace  $\tau_{n+1}, \dots, \tau_{n+k}$  by a uniform binomial process on  $[0, t]$ . We also note that  $\mu_t^n * v_h^{\otimes n} \leq \mu_u^n$  by the corresponding property of the binomial process. This implies  $\mu_t^n * p_h^{\otimes n} \leq q_u^n$  since both sides have continuous densities outside the diagonals by Lemma 5.2, justifying the third step. The last step is elementary calculus.

Since the previous estimate is uniform in  $x \in (\mathbb{R}^d)^{(n)}$ , it remains valid for the measures  $\mu * v_t^{n+k}$  with  $q_u^n$  replaced by the convolution  $\mu * q_u^n$ , which is bounded on compacts in  $(\mathbb{R}^d)^{(n)} \times (0, \infty)$  by Lemma 5.2. Finally, Theorem 4.2(i) shows, as before, that  $\mu * v_t^n * p_h^{\otimes n}(x) \leq \mu * q_u^n(x)$  for  $h \leq t$ , which is again locally bounded.  $\square$

We conclude with two technical results, needed in the next section.

LEMMA 5.8. For any initial measure  $\mu \in \mathcal{M}_d$  and a  $\pi \in \mathcal{P}_n$  with  $|\pi| < n$ , we have

$$\mu * v_t^\pi * \bigotimes_{J \in \pi} v_h^J = q_{t,h} \cdot \lambda^{\otimes nd}, \quad t, h > 0,$$

where  $q_{t,h}(x) \rightarrow 0$  as  $h \rightarrow 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts.

PROOF. For  $x = (x_1, \dots, x_n) \in (\mathbb{R}^d)^{(n)}$ , write  $\Delta = \min_{i \neq j} |x_i - x_j|$ , and note that

$$\inf_{u \in (\mathbb{R}^d)^\pi} \sum_{i \in J \in \pi} |x_i - u_J|^2 \geq (\Delta/2)^2 \sum_{J \in \pi} (|J| - 1) \geq \Delta^2/4.$$

Letting  $q_h$  denote the continuous density of  $\otimes_{J \in \pi} v_h^J$  on  $(\mathbb{R}^d)^{(n)}$  and using Lemma 5.2, we get as  $h \rightarrow 0$

$$\sup_{u \in (\mathbb{R}^d)^\pi} q_h(x - u) \leq \sup_{u \in (\mathbb{R}^d)^\pi} \prod_{i \in J \in \pi} p_{nh}(x_i - u_J) \leq h^{-nd/2} e^{-\Delta^2/8nhd} \rightarrow 0,$$

uniformly for  $x \in (\mathbb{R}^d)^{(n)}$  in compacts. Since  $\|v_t^\pi\| = |\pi|!t^{|\pi|-1}$  by Theorem 4.4, we conclude that

$$\sup_{u \in (\mathbb{R}^d)^\pi} (v_t^\pi * q_h)(x - u) \rightarrow 0, \quad h \rightarrow 0,$$

uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times \mathbb{R}_+$  in compacts.

Since the densities  $q_t^n$  of  $v_t^n$  satisfy  $v_t^\pi * \otimes_{J \in \pi} q_h^J \leq q_{t+h}^n$  by Theorems 4.2(iv) and 5.3, we have  $v_t^\pi * q_h \leq q_{t+h}^n$ . Using Lemmas 2.4 and 5.2 and writing  $\bar{u} = (u, \dots, u)$ , we get

$$(v_t^\pi * q_h)(x - u) \leq p_b^{\otimes n}(-\bar{u}) = p_b^{\otimes n}(\bar{u}), \quad u \in \mathbb{R}^d,$$

for some constant  $b > 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. Here  $\int p_b^{\otimes n}(\bar{u})\mu(du) < \infty$  since  $\mu p_t < \infty$  for all  $t > 0$ . Letting  $h = h_n \rightarrow 0$  and restricting  $(x, t) = (x_n, t_n)$  to a compact subset of  $(\mathbb{R}^d)^{(n)} \times (0, \infty)$ , we get by dominated convergence

$$q_{t,h}(x) = (\mu * v_t^\pi * q_h)(x) = \int \mu(du)(v_t^\pi * q_h)(x - u) \rightarrow 0,$$

which yields the required uniform convergence.  $\square$

For the clusters  $\eta$  of a DW-process in  $\mathbb{R}^d$ , we define

$$v_{h,\varepsilon}(dx) = E_0[\eta_h(dx); \sup_t \eta_t(B_x^\varepsilon)^c > 0], \quad h, \varepsilon > 0, x \in \mathbb{R}^d.$$

LEMMA 5.9. *For any initial measure  $\mu \in \mathcal{M}_d$ , we have*

$$\mu * v_t^n * (v_{h,\varepsilon} \otimes v_h^{\otimes(n-1)}) = q_{t,h}^\varepsilon \cdot \lambda^{\otimes nd}, \quad t, h, \varepsilon > 0,$$

where  $q_{t,h}^\varepsilon(x) \rightarrow 0$  as  $\varepsilon^{2+r} \geq h \rightarrow 0$  for some  $r > 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts.

PROOF. Let  $\rho$  be the span of  $\eta$  from 0, and put  $T(r) = P_0[\rho > r \mid \eta_1]_0$  and  $h' = h^{1/2}$ . By Palm disintegration and Lemma 8.1(iv),  $v_{h,\varepsilon}$  has a density bounded by

$$p_h(x)T\left(\frac{\varepsilon - |x|}{h'}\right) \leq \begin{cases} p_h(x)T(\varepsilon/2h'), & |x| \leq \varepsilon/2, \\ p_h(x), & |x| > \varepsilon/2. \end{cases}$$

Letting  $v'_{h,\varepsilon}$  and  $v''_{h,\varepsilon}$  denote the restrictions of  $v_{h,\varepsilon}$  to  $B_0^{\varepsilon/2}$  and  $(B_0^{\varepsilon/2})^c$ , respectively, we conclude that  $v'_{h,\varepsilon} \otimes v_h^{\otimes(n-1)}$  has a density  $\leq T(\varepsilon/2h')p_h^{\otimes n}(x)$ .

Hence, for  $0 < h \leq t$ , Theorem 4.4 yields a density  $\leq T(\varepsilon/2h')q_{t+h}^n(x)$  of  $v_t^n * (v'_{h,\varepsilon} \otimes v_h^{\otimes(n-1)})$ . Here  $T(\varepsilon/2h') \rightarrow 0$  as  $h/\varepsilon^2 \rightarrow 0$  since  $\rho < \infty$  a.s., and Lemma 5.2 gives  $\sup_u q_{t+h}^n(x-u) \leq 1$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts.

Next we note that  $v''_{h,\varepsilon} \otimes v_h^{\otimes(n-1)}$  has a density bounded by

$$p_h^{\otimes n}(x)1\{|x_1| > \varepsilon/2\} \leq h^{-nd/2}e^{-\varepsilon^2/8h} \rightarrow 0.$$

Since  $\|v_t^n\| \leq 1$  for bounded  $t > 0$  by Theorem 4.4, even  $v_t^n * (v''_{h,\varepsilon} \otimes v_h^{\otimes(n-1)})$  has a density that tends to 0, uniformly for  $x \in \mathbb{R}^{nd}$  and bounded  $t > 0$ . Combining the results for  $v'_{h,\varepsilon}$  and  $v''_{h,\varepsilon}$ , we conclude that  $v_t^n * (v_{h,\varepsilon} \otimes v_h^{\otimes(n-1)})$  has a density  $q_{t,h}^\varepsilon$  satisfying  $\sup_u q_{t,h}^\varepsilon(x-u) \rightarrow 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. To deduce the stated result for general  $\mu$ , we may argue as in the previous proof, using dominated convergence based on the relations  $v_{h,\varepsilon} \leq v_h$  and  $v_t^n * v_h^{\otimes n} \leq v_{t+h}^n$  with associated density versions, valid by Theorems 4.2(iv) and 5.3.  $\square$

**6. Palm continuity and approximation.** Here we establish some continuity and approximation properties for the multivariate Palm distributions of a DW-process, needed in Section 9. We begin with a continuity property of the transition kernels, which might be known, though no reference could be found.

LEMMA 6.1. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ , and fix any  $\mu$  and  $B \in \mathcal{B}^d$ , where either  $\mu$  or  $B$  is bounded. Then  $\mathcal{L}_\mu(1_B \xi_t)$  is continuous in total variation in  $t > 0$ .*

PROOF. First let  $\|\mu\| < \infty$ . For any  $t > 0$ , the ancestors of  $\xi_t$  at time 0 form a Poisson process  $\zeta_0$  with intensity  $t^{-1}\mu$ . By Lemma 4.5 the ancestors splitting before time  $s \in (0, t)$  form a Poisson process with intensity  $st^{-2}\mu$ , and so such a split occurs with probability  $1 - \exp(-st^{-2}\|\mu\|) \leq st^{-2}\|\mu\|$ . Hence, the process  $\zeta_s$  of ancestors at time  $s$  agrees, up to a set of probability  $st^{-2}\|\mu\|$ , with a Poisson process with intensity  $t^{-1}\mu * p_s$ .

Replacing  $s$  and  $t$  by  $s+h$  and  $t+h$  where  $|h| < s$ , we see that the process  $\zeta'_{s+h}$  of ancestors of  $\xi_{t+h}$  at time  $s+h$  agrees up to probability  $(s+h)(t+h)^{-2}\|\mu\|$  with a Poisson process with intensity  $(t+h)^{-1}\mu * p_{s+h}$ . Since  $\xi_t$  and  $\xi_{t+h}$  are both Cox cluster processes with the same cluster kernel, given by the normalized distribution of a  $(t-s)$ -cluster, the total variation distance between their distributions is bounded by the corresponding distance for the two ancestral processes. Noting that  $\|\mathcal{L}(\eta_1) - \mathcal{L}(\eta_2)\| \leq \|E\eta_1 - E\eta_2\|$  for any Poisson processes  $\eta_1$  and  $\eta_2$  on the same space, we obtain a total bound of the order  $\|\mu\|$  times

$$\frac{s}{t^2} + \frac{s+|h|}{(t+h)^2} + \left| \frac{1}{t} - \frac{1}{t+h} \right| + \frac{\|p_s - p_{s+h}\|_1}{t} \leq \frac{s+|h|}{t^2} + \frac{|h|}{st}.$$

Choosing  $s = |h|^{1/2}$ , we get convergence to 0 as  $h \rightarrow 0$ , uniformly for  $t \in (0, \infty)$  in compacts, which proves the continuity in  $t$ .

Now let  $\mu$  be arbitrary, and assume instead that  $B$  is bounded. Let  $\mu_r$  and  $\mu'_r$  denote the restrictions of  $\mu$  to  $B_0^r$  and  $(B_0^r)^c$ . Then  $P_{\mu'_r}\{\xi_t B > 0\} \leq (\mu'_r * \nu_t)B < \infty$ , uniformly for  $t \in (0, \infty)$  in compacts (cf. Lemma 7.2 below). As  $r \rightarrow \infty$ , we get  $P_{\mu'_r}\{\xi_t B > 0\} \rightarrow 0$  by dominated convergence, in the same uniform sense. Finally, by the version for bounded  $\mu$ ,  $\mathcal{L}_{\mu_r}(1_B \xi_t)$  is continuous in total variation in  $t > 0$  for fixed  $r > 0$ .  $\square$

A similar argument based on the estimate  $\|\delta_x * p_s - p_s\|_1 \leq |x|s^{-1/2}$  yields continuity in the same sense even under spatial shifts. We proceed with a uniform bound for the associated Palm distributions.

LEMMA 6.2. *For any  $\mu$ ,  $t > 0$ , and open  $G \subset \mathbb{R}^d$ , there exist some functions  $p_h$  on  $G^{(n)}$  with  $p_h \rightarrow 0$  as  $h \rightarrow 0$ , uniformly for  $(x, t) \in G^{(n)} \times (0, \infty)$  in compacts, such that a.e.  $E_\mu \xi_s^{\otimes n}$  on  $G^{(n)}$  and for  $r < s \leq t$  with  $2s > t + r$*

$$\|\mathcal{L}_\mu[1_{G^c} \xi_t \|\xi_s^{\otimes n}] - E_\mu[\mathcal{L}_{\xi_r}(1_{G^c} \xi_{t-r}) \|\xi_s^{\otimes n}]\| \leq p_{t-r}.$$

PROOF. The random measures  $\xi_s$  and  $\xi_t$  may be regarded as Cox cluster processes generated by the random measure  $h^{-1} \xi_r$  and the probability kernel  $h \mathcal{L}_u(\eta_h)$  from  $\mathbb{R}^d$  to  $\mathcal{M}_d$ , where  $h = s - r$  or  $h = t - r$ , respectively. To keep track of the cluster structure, we introduce some marked versions  $\tilde{\xi}_s$  or  $\tilde{\xi}_t$  on  $\mathbb{R}^d \times [0, 1]$ , where each cluster  $\xi_i$  is replaced by  $\tilde{\xi}_i = \xi_i \otimes \delta_{\sigma_i}$  for some i.i.d.  $U(0, 1)$  random variables  $\sigma_i$  independent of the  $\xi_i$ . Note that  $\tilde{\xi}_s$  and  $\tilde{\xi}_t$  are again cluster processes with generating kernels  $\tilde{v}_h = \mathcal{L}(\xi_h \otimes \delta_\sigma)$  from  $\mathbb{R}^d$  to  $\mathcal{M}(\mathbb{R}^d \times [0, 1])$ , where  $\mathcal{L}(\xi_h, \sigma) = v_h \otimes \lambda$ . By the transfer theorem (cf. [15], page 112), we may assume that  $\tilde{\xi}_t(\cdot \times [0, 1]) = \xi_t$  a.s.

For any  $v = (v_1, \dots, v_n) \in [0, 1]^n$ , we may write  $\tilde{\xi}_t = \tilde{\xi}_{t,v} + \tilde{\xi}'_{t,v}$ , where  $\tilde{\xi}_{t,v}$  denotes the restriction of  $\tilde{\xi}_t$  to  $\mathbb{R}^d \times \{v_1, \dots, v_n\}^c$ , which is product-measurable in  $(\omega, v)$  by Lemma 3.3. Writing  $D = ([0, 1]^{(n)})^c$ , we get for any  $B \in \hat{\mathcal{B}}^{nd}$  and for measurable functions  $f : (\mathbb{R}^d \times [0, 1])^n \times \mathcal{M}(\mathbb{R}^d \times [0, 1]) \rightarrow [0, 1]$  with  $f_{x,v} = 0$  for  $x \in B^c$

$$\begin{aligned} & \left| \int \int E_\mu \tilde{\xi}_s^{\otimes n}(dx dv) (E_\mu[f_{x,v}(1_{G^c} \tilde{\xi}_t) \|\tilde{\xi}_s^{\otimes n}]_{x,v} - E_\mu[f_{x,v}(1_{G^c} \tilde{\xi}_{t,v}) \|\tilde{\xi}_s^{\otimes n}]_{x,v}) \right| \\ & \leq E_\mu \int \int 1_B(x) \tilde{\xi}_s^{\otimes n}(dx dv) |f_{x,v}(1_{G^c} \tilde{\xi}_t) - f_{x,v}(1_{G^c} \tilde{\xi}_{t,v})| \\ & \leq E_\mu \int \int 1_B(x) \tilde{\xi}_s^{\otimes n}(dx dv) 1_{\{\tilde{\xi}'_{t,v} G^c > 0\}} \\ & \leq E_\mu \tilde{\xi}_s^{\otimes n}(B \times D) + E_\mu \int \int_{B \times D^c} \tilde{\xi}_s^{\otimes n}(dx dv) \sum_{i \leq n} 1_{\{\tilde{\xi}'_{t,v_i} G^c > 0\}}. \end{aligned}$$

To estimate the first term on the right, we define  $\mathcal{P}'_n = \{\pi \in \mathcal{P}_n; |\pi| < n\}$ . For any  $\kappa, \pi \in \mathcal{P}_n$ , write  $\kappa < \pi$  to mean that every set in  $\kappa$  is a union of sets in  $\pi$ , and put  $\pi I = \{J \in \pi; J \subset I\}$ . Let  $\zeta_r$  be the Cox process of ancestors to  $\xi_s$  at time  $r = s - h$ . Using the definition of  $\tilde{\xi}_s$ , the conditional independence of the clusters  $\eta_u$  and Theorem 4.2(i), we get

$$\begin{aligned} E_\mu \tilde{\xi}_s^{\otimes n}(\cdot \times D) &= \sum_{\pi \in \mathcal{P}'_n} E_\mu \int \zeta_r^{(\pi)}(du) \bigotimes_{J \in \pi} \eta_{h,u_J}^{\otimes J} = \sum_{\pi \in \mathcal{P}'_n} E_\mu \xi_s^{\otimes \pi} * \bigotimes_{J \in \pi} v_h^J \\ &= \sum_{\pi \in \mathcal{P}'_n} \sum_{\kappa < \pi} \bigotimes_{I \in \kappa} (\mu * v_s^{\pi I}) * \bigotimes_{J \in \pi I} v_h^J, \end{aligned}$$

and similarly for the associated densities, where  $\zeta_s^{(\pi)}$  denotes the factorial measure of  $\zeta_s$  on  $(\mathbb{R}^d)^{(\pi)}$ . For each term on the right, we have  $|\pi I| < |I|$  for at least one  $I \in \kappa$ , and then Lemma 5.8 yields a corresponding density that tends to 0 as  $h \rightarrow 0$ , uniformly for  $(x, r) \in (\mathbb{R}^d)^{(I)} \times (0, \infty)$  in compacts. The remaining factors have locally bounded densities on  $(\mathbb{R}^d)^{(I)} \times (0, \infty)$ , for example, by Lemma 5.6. Hence, by combination,  $E_\mu \tilde{\xi}_s^{\otimes n}(\cdot \times D)$  has a density that tends to 0 as  $h \rightarrow 0$ , uniformly for  $(x, s) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts.

Turning to the second term on the right, let  $B = X_i B_i \subset G^{(n)}$  be compact, and write  $B_J = X_{i \in J} B_i$  for  $J \subset \{1, \dots, n\}$ . Using the previous notation and defining  $v_{h,\varepsilon}$  as in Lemma 5.9, we get for  $\varepsilon > 0$  small enough,

$$\begin{aligned} E_\mu \int \int_{B \times D^c} \tilde{\xi}_s^{\otimes n}(dx dv) 1\{\tilde{\xi}'_{t,v_1} G^c > 0\} \\ &= E_\mu \int \zeta_r^{(n)}(du) \int_{B_1} \eta_h^{u_1}(dx_1) 1\{\eta_h^{u_1} G^c > 0\} \prod_{i>1} \eta_h^{u_i} B_i \\ &\leq E_\mu \xi_r^{\otimes n} * (v_{h,\varepsilon} \otimes v_h^{\otimes(n-1)}) B \\ &= \sum_{\pi \in \mathcal{P}_n} (\mu * v_r^{J_1} * (v_{h,\varepsilon} \otimes v_h^{\otimes J'_1}) B_{J_1}) \prod_{J \in \pi'} (\mu * v_r^J * v_h^{\otimes J}) B_J, \end{aligned}$$

where  $1 \in J_1 \in \pi$ ,  $J'_1 = J_1 \setminus \{1\}$  and  $\pi' = \pi \setminus \{J_1\}$ . For each term on the right, Lemma 5.9 yields a density of the first factor that tends to 0 as  $h \rightarrow 0$  for fixed  $\varepsilon > 0$ , uniformly for  $(x_{J_1}, t) \in (\mathbb{R}^d)^{J_1} \times (0, \infty)$  in compacts. Since the remaining factors have locally bounded densities on the sets  $(\mathbb{R}^d)^J \times (0, \infty)$  by Lemma 5.6, the entire sum has a density that tends to 0 as  $h \rightarrow 0$  for fixed  $\varepsilon > 0$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. Combining the previous estimates and using Lemma 3.4, we obtain

$$(16) \quad \|\mathcal{L}_\mu[1_{G^c} \tilde{\xi}_t \| \tilde{\xi}_s^{\otimes n}]_{x,v} - \mathcal{L}_\mu[1_{G^c} \tilde{\xi}_{t,v} \| \tilde{\xi}_s^{\otimes n}]_{x,v}\| \leq p_h \quad \text{a.e. } E_\mu \tilde{\xi}_s^{\otimes n},$$

for some measurable functions  $p_h$  on  $(\mathbb{R}^d)^{(n)}$  with  $p_h \rightarrow 0$  as  $h \rightarrow 0$ , uniformly on compacts.

We now apply the probabilistic form of Lemma 3.8 to the pair  $(\tilde{\xi}_s, \tilde{\xi}_t)$ , regarded as a Cox cluster process generated by  $\xi_r$ . Under  $\mathcal{L}_\mu[\tilde{\xi}_t \|\tilde{\xi}_s^{\otimes n}]_{x,v}$ , the leading term agrees with the nondiagonal component  $\tilde{\xi}_{t,v}$ . To see this, we first condition on  $\xi_r$ , so that  $\tilde{\xi}_t$  becomes a Poisson cluster process generated by a nonrandom measure at time  $r$ . By Fubini's theorem, the leading term is a.e. restricted to  $\mathbb{R}^d \times \{v_1, \dots, v_n\}^c$ , whereas by Lemma 3.10 the remaining terms are a.e. restricted to  $\mathbb{R}^d \times \{v_1, \dots, v_n\}$ . Hence, Lemma 3.8 yields a.e.

$$\mathcal{L}_\mu[\tilde{\xi}_{t,v} \|\tilde{\xi}_s^{\otimes n}]_{x,v} = E_\mu[\mathcal{L}_{\xi_r}(\tilde{\xi}_{t-r}) \|\tilde{\xi}_s^{\otimes n}]_{x,v}, \quad x \in (\mathbb{R}^d)^{(n)}, v \in [0, 1]^n,$$

and so, by (16),

$$\|\mathcal{L}_\mu[1_{G^c} \tilde{\xi}_t \|\tilde{\xi}_s^{\otimes n}]_{x,v} - E_\mu[\mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t-r}) \|\tilde{\xi}_s^{\otimes n}]_{x,v}\| \leq p_h \quad \text{a.e. } E_\mu \tilde{\xi}_s^{\otimes n}.$$

The assertion now follows by Lemma 3.5.  $\square$

We may now establish some basic regularity properties for the Palm distributions of a DW-process. Here again, weak continuity is defined with respect to the vague topology.

**THEOREM 6.3.** *For a DW-process  $\xi$  in  $\mathbb{R}^d$ , there exist versions of the Palm kernels  $\mathcal{L}_\mu[\xi_t \|\xi_t^{\otimes n}]_x$ , such that:*

- (i)  $\mathcal{L}_\mu[\xi_t \|\xi_t^{\otimes n}]_x$  is tight and weakly continuous in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$ ;
- (ii) for any  $t > 0$  and open  $G \subset \mathbb{R}^d$ ,  $\mathcal{L}_\mu[1_{G^c} \xi_t \|\xi_t^{\otimes n}]_x$  is continuous in total variation in  $x \in G^{(n)}$ ;
- (iii) for any open  $G$  and bounded  $\mu$  or  $G^c$ ,  $\mathcal{L}_\mu[1_{G^c} \xi_t \|\xi_t^{\otimes n}]_x$  is continuous in total variation in  $(x, t) \in G^{(n)} \times (0, \infty)$ .

**PROOF.** (ii) For the conditional moment measure  $E_\mu[\xi_t^{\otimes n} | \xi_r]$  with  $r \geq 0$  fixed, Theorem 5.4 yields a Lebesgue density, that is,  $L^1$ -continuous in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (r, \infty)$ . Since the continuous density of  $E_\mu \xi_t^{\otimes n}$  is even strictly positive by Theorem 5.3, the  $L^1$ -continuity extends to the density of  $E_\mu[\xi_t^{\otimes n} | \xi_r]$  with respect to  $E_\mu \xi_t^{\otimes n}$ . Hence, by Lemma 3.7 the Palm kernels  $\mathcal{L}_\mu[\xi_r \|\xi_t^{\otimes n}]_x$  have versions that are continuous in total variation in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (r, \infty)$  for fixed  $r \geq 0$ . Fixing any  $t > r \geq 0$  and  $G \subset \mathbb{R}^d$ , we see in particular that the kernel  $E_\mu[\mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t-r}) \|\xi_t^{\otimes n}]_x$  has a version, that is, continuous in total variation in  $x \in (\mathbb{R}^d)^{(n)}$ . Choosing arbitrary  $r_1, r_2, \dots \in (0, t)$  with  $r_k \rightarrow t$ , using Lemma 6.2 with  $r = r_k$  and  $s = t$ , and invoking Lemma 3.6, we obtain a similar continuity property for the kernel  $\mathcal{L}_\mu[1_{G^c} \xi_t \|\xi_t^{\otimes n}]_x$ .

(iii) Let  $\mu$  or  $G^c$  be bounded, and fix any  $r \geq 0$ . As before, we may choose the kernels  $\mathcal{L}_\mu[\xi_r \|\xi_t^{\otimes n}]_x$  to be continuous in total variation in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (r, \infty)$ . For any  $x, x' \in (\mathbb{R}^d)^{(n)}$  and  $t, t' > r$ , write

$$\begin{aligned} & \|E_\mu[\mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t-r}) \|\xi_t^{\otimes n}]_x - E_\mu[\mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t'-r}) \|\xi_{t'}^{\otimes n}]_{x'}\| \\ & \leq E_\mu[\|\mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t-r}) - \mathcal{L}_{\xi_r}(1_{G^c} \tilde{\xi}_{t'-r})\| \|\xi_t^{\otimes n}\|_x \\ & \quad + \|\mathcal{L}_\mu[\xi_r \|\xi_t^{\otimes n}]_x - \mathcal{L}_\mu[\xi_r \|\xi_{t'}^{\otimes n}]_{x'}\|. \end{aligned}$$

As  $x' \rightarrow x$  and  $t' \rightarrow t$  for fixed  $r$ , the first term on the right tends to 0 in total variation by Lemma 6.1 and dominated convergence, whereas the second term tends to 0 in the same sense by the continuous choice of kernels. This shows that  $E_\mu[\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r})\|\xi_t^{\otimes n}]_x$  is continuous in total variation in  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (r, \infty)$  for fixed  $r, \mu$ , and  $G$ .

Now choose  $h_1, h_2, \dots > 0$  to be rationally independent<sup>1</sup> with  $h_n \rightarrow 0$ , and define

$$r_k(t) = h_k[h_k^{-1}t-], \quad t > 0, k \in \mathbb{N}.$$

Then Lemma 6.2 applies with  $r = r_k(t)$  and  $s = t$  for some functions  $p_k$  with  $p_k \rightarrow 0$ , uniformly for  $(x, t) \in G^{(n)} \times (0, \infty)$  in compacts. Since the sets  $U_k = h_k \cup_j (j - 1, j)$  satisfy  $\limsup_k U_k = (0, \infty)$ , Lemma 3.6 yields a version of the kernel  $\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_t^{\otimes n}]_x$ , that is, continuous in total variation in  $(x, t) \in G^{(n)} \times (0, \infty)$ .

(i) Writing  $U_x^r = \cup_i B_{x_i}^r$  and using Theorem 5.3 and Lemma 5.7, we get

$$\frac{E_\mu \xi_t^{\otimes n} B_x^\varepsilon(\xi_t U_x^r \wedge 1)}{E_\mu \xi_t^{\otimes n} B_x^\varepsilon} \leq r^d \sum_{i \leq n} (E_\mu \xi_t^{\otimes(n+1)} * (p_{\varepsilon^2}^{\otimes n} \otimes p_{r^2}))(x, x_i) \leq \begin{cases} r^2, \\ r^2 |\log r|, \end{cases}$$

uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. Now use part (iii), along with a uniform version of Lemma 3.12 for random measures  $\xi_t$ .  $\square$

The last result yields a similar continuity property for the forward Palm kernels  $\mathcal{L}_\mu[\xi_t\|\xi_s^{\otimes n}]_x$  with  $s < t$ .

**COROLLARY 6.4.** *For fixed  $t > s > 0$ ,  $\mathcal{L}_\mu[\xi_t\|\xi_s^{\otimes n}]_x$  has a version, that is, continuous in total variation in  $x \in (\mathbb{R}^d)^{(n)}$ .*

**PROOF.** Let  $\zeta_s$  denote the ancestral process of  $\xi_t$  at time  $s = t - h$ . Since  $\xi_t \perp\!\!\!\perp_{\zeta_s} \xi_s^{\otimes n}$ , it suffices by Lemma 3.13 to prove the continuity in total variation of  $\mathcal{L}_\mu[\zeta_s\|\xi_s^{\otimes n}]_x$ . Since  $\zeta_s$  is a Cox process directed by  $h^{-1}\xi_s$ , we see from [22] that  $\mathcal{L}_\mu[\zeta_s\|\xi_s^{\otimes n}]_x$  is a.e. the distribution of a Cox process directed by  $\mathcal{L}_\mu[h^{-1}\xi_s\|\xi_s^{\otimes n}]_x$ . Hence, for any  $G \subset \mathbb{R}^d$ ,

$$\begin{aligned} & \| \mathcal{L}_\mu[\zeta_s\|\xi_s^{\otimes n}]_x - \mathcal{L}_\mu[1_{G^c}\zeta_s\|\xi_s^{\otimes n}]_x \| \\ & \leq P_\mu[\zeta_s G > 0\|\xi_s^{\otimes n}]_x \\ & = E_\mu[1 - e^{-h^{-1}\xi_s G}\|\xi_s^{\otimes n}]_x \leq E_\mu[h^{-1}\xi_s G \wedge 1\|\xi_s^{\otimes n}]_x. \end{aligned}$$

Choosing versions of  $\mathcal{L}_\mu[\xi_s\|\xi_s^{\otimes n}]_x$  as in Theorem 6.3 and using part (i) of that result, we conclude that  $\mathcal{L}_\mu[\zeta_s\|\xi_s^{\otimes n}]_x$  can be approximated in total variation by

---

<sup>1</sup>Meaning that no nontrivial linear combination with integral coefficients exists.

kernels  $\mathcal{L}_\mu[1_{G^c}\zeta_s\|\xi_s^{\otimes n}]_x$  with open  $G \subset \mathbb{R}^d$ , uniformly for  $x \in G^{(n)}$  in compacts. It is then enough to choose the latter kernel to be continuous in total variation on  $G^{(n)}$ . Since  $1_{G^c}\zeta_s \perp\!\!\!\perp 1_{G^c}\xi_s$ , such a version exists by Lemma 3.13, given the corresponding property of  $\mathcal{L}_\mu[1_{G^c}\xi_s\|\xi_s^{\otimes n}]_x$  from Theorem 6.3(ii).  $\square$

The following approximation plays a crucial role in Section 9. Here the Palm kernels are assumed to be continuous, in the sense of Theorem 6.3 and Corollary 6.4.

LEMMA 6.5. *Fix any  $\mu, t > 0$  and open  $G \subset \mathbb{R}^d$ . Then as  $s \uparrow t$  and  $u \rightarrow x \in G^{(n)}$ , we have*

$$\|\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_s^{\otimes n}]_u - \mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_t^{\otimes n}]_x\| \rightarrow 0.$$

PROOF. Letting  $r < s \leq t$ , we write

$$\begin{aligned} & \|\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_s^{\otimes n}]_x - \mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_t^{\otimes n}]_x\| \\ (17) \quad & \leq \|\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_s^{\otimes n}]_x - E_\mu[\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r})\|\xi_s^{\otimes n}]_x\| \\ & \quad + \|\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_t^{\otimes n}]_x - E_\mu[\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r})\|\xi_t^{\otimes n}]_x\| \\ & \quad + \|\mathcal{L}_\mu[\xi_r\|\xi_s^{\otimes n}]_x - \mathcal{L}_\mu[\xi_r\|\xi_t^{\otimes n}]_x\|. \end{aligned}$$

By Lemma 3.7 and Theorems 5.3 and 5.4, the kernels  $\mathcal{L}_\mu[\xi_r\|\xi_s^{\otimes n}]_x$  and  $\mathcal{L}_\mu[\xi_r\|\xi_t^{\otimes n}]_x$  have versions that are continuous in total variation in  $x \in (\mathbb{R}^d)^{(n)}$ . With such choices and for  $2s > t + r$ , Lemma 6.2 shows that the first two terms on the right of (17) are bounded by some functions  $p_{t-r}$ , where  $p_h \downarrow 0$  as  $h \rightarrow 0$ , uniformly for  $(x, t) \in G^{(n)} \times (0, \infty)$  in compacts. Next, by Lemma 3.7 and Theorem 5.4, the last term in (17) tends to 0 as  $s \rightarrow t$  for fixed  $r$  and  $t$ , uniformly for  $(x, t) \in (\mathbb{R}^d)^{(n)} \times (0, \infty)$  in compacts. Letting  $s \rightarrow t$  and then  $r \rightarrow t$ , we conclude that the left-hand side of (17) tends to 0 as  $s \uparrow t$ , uniformly for  $x \in G^{(n)}$  in compacts. Since  $\mathcal{L}_\mu[1_{G^c}\xi_t\|\xi_t^{\otimes n}]_x$  is continuous in total variation in  $x \in G^{(n)}$  by Theorem 6.3(ii), we obtain the required joint convergence as  $s \uparrow t$  and  $u \rightarrow x$ .  $\square$

We conclude with a continuity property of the one-dimensional Palm distributions, quoted from Lemma 3.5 in [19] and its proof. Here  $\mathcal{L}_\mu[\xi_t\|\xi_t]_x$  and  $\mathcal{L}_\mu[\eta_t\|\eta_t]_x$  denote the continuous versions of the Palm distributions of  $\xi_t$  and  $\eta_t$ , as constructed explicitly in [2, 4].

LEMMA 6.6. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ . Then for fixed  $t > 0$  and  $\mu$ , the shifted Palm distributions  $\mathcal{L}_\mu[\theta_{-x}\xi_t\|\xi_t]_x$  and  $\mathcal{L}_\mu[\theta_{-x}\eta_t\|\eta_t]_x$  are continuous in  $x \in \mathbb{R}^d$ , in total variation on any compact set  $B \subset \mathbb{R}^d$ . When  $\|\mu\| < \infty$  we may even take  $B = \mathbb{R}^d$ .*

**7. Hitting, multiplicities, and decoupling.** Here we derive some estimates of hitting probabilities needed in subsequent sections. We begin with the basic hitting estimates, quoted in the form of Lemma 4.2 in [19]. The statement also defines the function  $t(\varepsilon)$  that occurs frequently below. Note that the definitions differ for  $d \geq 3$  and  $d = 2$ .

LEMMA 7.1. *Let  $\eta$  be the canonical cluster of a DW-process in  $\mathbb{R}^d$ . Then:*

(i) *for  $d \geq 3$ , we have with  $t(\varepsilon) = t + \varepsilon^2$ , uniformly in  $\mu, t$  and  $\varepsilon$  with  $0 < \varepsilon \leq \sqrt{t}$ ,*

$$\mu p_t \leq \varepsilon^{2-d} P_\mu\{\eta_t B_0^\varepsilon > 0\} \leq \mu p_{t(\varepsilon)};$$

(ii) *for  $d = 2$ , we may choose  $t(\varepsilon) = tl(\varepsilon/\sqrt{t})$  with  $0 \leq l_\varepsilon - 1 \leq |\log \varepsilon|^{-1/2}$  such that, uniformly for  $\mu, t$  and  $\varepsilon$  with  $0 < 2\varepsilon < \sqrt{t}$ ,*

$$\mu p_t \leq \log(t/\varepsilon^2) P_\mu\{\eta_t B_0^\varepsilon > 0\} \leq \mu p_{t(\varepsilon)}.$$

We proceed with a uniform limit theorem for hitting probabilities, quoted from Lemma 5.1 and Theorem 5.3 in [19]. Define

$$c_d = \lim_{\varepsilon \rightarrow 0} \varepsilon^{2-d} P_0\{\xi_t B_0^\varepsilon > 0\} / p_t(0), \quad d \geq 3,$$

$$m(\varepsilon) = |\log \varepsilon| P_{\lambda^2}\{\eta_1 B_0^\varepsilon > 0\}, \quad d = 2,$$

where the constants  $c_d$  exist by [3], and the function  $\log m(\varepsilon)$  is bounded for  $\varepsilon \ll 1$  by Lemma 5.1 in [19].

LEMMA 7.2. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ . Then as  $\varepsilon \rightarrow 0$  for fixed  $t > 0, \mu$  and bounded  $B$ :*

- (i)  $\|\varepsilon^{2-d} P_\mu\{\xi_t B_{(\cdot)}^\varepsilon > 0\} - c_d(\mu * p_t)\|_B \rightarrow 0, \quad d \geq 3;$
- (ii)  $\|\log \varepsilon |P_\mu\{\xi_t B_{(\cdot)}^\varepsilon > 0\} - m(\varepsilon)(\mu * p_t)\|_B \rightarrow 0, \quad d = 2,$

*and similarly with  $\xi_t$  replaced by  $\eta_t$ . When  $\mu$  is bounded, we may take  $B = \mathbb{R}^d$ .*

We also quote from Lemma 4.4 in [19] some estimates for multiple hits, later to be extended in Lemma 7.5. Let  $\kappa_h^\varepsilon$  denote the number of  $h$ -clusters of  $\xi_t$  hitting  $B_0^\varepsilon$  at time  $t$ .

LEMMA 7.3. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ . Then:*

(i) *for  $d \geq 3$ , we have with  $t_\varepsilon = t + \varepsilon^2$  as  $\varepsilon^2 \ll h \leq t$*

$$E_\mu \kappa_h^\varepsilon (\kappa_h^\varepsilon - 1) \leq \varepsilon^{2(d-2)} \{h^{1-d/2} \mu p_t + (\mu p_{t(\varepsilon)})^2\};$$

(ii) for  $d = 2$ , there exists a function  $t_{h,\varepsilon} > t$  with  $0 < t_{h,\varepsilon} - t \leq h|\log \varepsilon|^{-1/2}$ , such that as  $\varepsilon \ll h \leq t$

$$E_\mu \kappa_h^\varepsilon (\kappa_h^\varepsilon - 1) \leq |\log \varepsilon|^{-2} \{ \log(t/h) \mu p_t + (\mu p_{t(h,\varepsilon)})^2 \}.$$

For a DW-process  $\xi$  in  $\mathbb{R}^d$  and for any  $t > h > 0$ , let  $\eta_h^1, \eta_h^2, \dots$  denote the  $h$ -clusters in  $\xi_t$ , and write  $\zeta_s$  for the ancestral process of  $\xi_t$  at time  $s = t - h$ . When  $\zeta_s = \sum_i \delta_{u_i}$ , we also write  $\eta_h^{u_i}$  for the  $h$ -cluster rooted at  $u_i$ . Put  $(\mathbb{N}^n)' = \mathbb{N}^n \setminus \mathbb{N}^{(n)}$ . Define  $h(\varepsilon)$  as in Lemma 7.1, but with  $t$  replaced by  $h$ .

First we estimate the probability for a single  $h$ -cluster in  $\xi_t$  to hit several  $\varepsilon$ -balls around  $x_1, \dots, x_n \in \mathbb{R}^d$ . We will refer repeatedly to the conditions

$$(18) \quad \begin{aligned} \varepsilon^2 \ll h \leq \varepsilon, \quad d \geq 3, \\ h \leq |\log \varepsilon|^{-1} \ll |\log h|^{-1}, \quad d = 2. \end{aligned}$$

LEMMA 7.4. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ , and fix any  $\mu, t > 0$ , and  $x \in (\mathbb{R}^d)^{(n)}$ . Then as  $\varepsilon, h \rightarrow 0$ , subject to (18),*

$$P_\mu \bigcup_{k \in (\mathbb{N}^n)'} \bigcap_{j \leq n} \{ \eta_h^{k_j} B_{x_j}^\varepsilon > 0 \} \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

PROOF. We need to show that for any  $i \neq j$  in  $\{1, \dots, n\}$ ,

$$P_\mu \bigcup_{k \in \mathbb{N}} \{ \eta_h^k B_{x_i}^\varepsilon \wedge \eta_h^k B_{x_j}^\varepsilon > 0 \} \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

Writing  $\bar{x} = \frac{1}{2}(x_i + x_j)$  and  $\Delta x = |x_i - x_j|$ , and using Cauchy's inequality, Lemmas 4.1 and 7.1(i) and the parallelogram identity, we get for  $d \geq 3$ ,

$$\begin{aligned} & P_\mu \bigcup_{k \in \mathbb{N}} \{ \eta_h^k B_{x_i}^\varepsilon \wedge \eta_h^k B_{x_j}^\varepsilon > 0 \} \\ & \leq E_\mu \sum_{k \in \mathbb{N}} 1 \{ \eta_h^k B_{x_i}^\varepsilon \wedge \eta_h^k B_{x_j}^\varepsilon > 0 \} \\ & = E_\mu \int \zeta_s(du) 1 \{ \eta_h^u B_{x_i}^\varepsilon \wedge \eta_h^u B_{x_j}^\varepsilon > 0 \} \\ & = \int E_\mu \xi_s(du) P_u \{ \eta_h B_{x_i}^\varepsilon \wedge \eta_h B_{x_j}^\varepsilon > 0 \} \\ & \leq \int (\mu * p_s)_u du (P_u \{ \eta_h B_{x_i}^\varepsilon > 0 \} P_u \{ \eta_h B_{x_j}^\varepsilon > 0 \})^{1/2} \\ & \leq \varepsilon^{d-2} \int (\mu * p_s)_u du (p_{h\varepsilon}(x_i - u) p_{h\varepsilon}(x_j - u))^{1/2} \\ & = \varepsilon^{d-2} \int (\mu * p_s)_u p_{h\varepsilon}(\bar{x} - u) du e^{-|\Delta x|^2/8h'} \\ & \leq \varepsilon^{d-2} (\mu * p_{2t})(\bar{x}) e^{-|\Delta x|^2/8h'}, \end{aligned}$$

which tends to 0 faster than any power of  $\varepsilon$ . If instead  $d = 2$ , we get, from Lemma 7.1(ii), the bound

$$(\log(h/\varepsilon^2))^{-1}(\mu * p_{t_\varepsilon})(\bar{x})e^{-|\Delta x|^2/8h'} \lesssim |\log \varepsilon|^{-1}(\mu * p_{2t})(\bar{x})\varepsilon^{|\Delta x|^2/16},$$

which tends to 0 faster than any power of  $|\log \varepsilon|^{-1}$ .  $\square$

We turn to the possibility for a single ball  $B_{x_j}^\varepsilon$  to be hit by several  $h$ -clusters of  $\xi_t$ , thus providing a multivariate version of Lemma 7.3:

LEMMA 7.5. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ , and fix any  $\mu, t > 0$  and  $x \in (\mathbb{R}^d)^{(n)}$ . Then as  $\varepsilon, h \rightarrow 0$ , subject to (18),*

$$E_\mu \left( \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

To compare with Lemma 7.3, note that the estimates for  $n = 1$  reduce to

$$E_\mu(\kappa_h^\varepsilon - 1)_+ \ll \begin{cases} \varepsilon^{d-2}, & d \geq 3, \\ |\log \varepsilon|^{-1}, & d = 2. \end{cases}$$

PROOF. On the sets

$$A_{h,\varepsilon} = \bigcap_{j \leq n} \left\{ \sum_{k \in \mathbb{N}} 1\{\eta_h^k B_{x_j}^\varepsilon > 0\} \leq n \right\}, \quad h, \varepsilon > 0,$$

we have

$$\begin{aligned} & \left( \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ \\ & \leq \sum_{k \in \mathbb{N}^n} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} = \prod_{j \leq n} \sum_{k \in \mathbb{N}} 1\{\eta_h^k B_{x_j}^\varepsilon > 0\} \leq n^n. \end{aligned}$$

On  $A_{h,\varepsilon}^c$  we note that  $\prod_{j \leq n} \eta_h^{k_j} B_{x_j}^\varepsilon > 0$  implies  $\eta_h^l B_{x_i}^\varepsilon > 0$  for some  $i \leq n$  and  $l \neq k_1, \dots, k_n$ , and so

$$\begin{aligned} & \left( \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ \\ & \leq \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} \\ (19) \quad & \leq \sum_{i \leq n} \sum_{(k,l) \in \mathbb{N}^{(n+1)}} 1\{\eta_h^l B_{x_i}^\varepsilon > 0\} \prod_{j \leq n} 1\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\} \\ & = \sum_{i \leq n} \int \int \zeta_s^{(n+1)}(du dv) 1\{\eta_h^v B_{x_i}^\varepsilon > 0\} \prod_{j \leq n} 1\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\}, \end{aligned}$$

where  $u = (u_1, \dots, u_n) \in \mathbb{R}^{nd}$  and  $v \in \mathbb{R}^d$ . Finally, writing  $U_{h,\varepsilon}$  for the union in Lemma 7.4 and using Lemma 2.11, we get on  $U_{h,\varepsilon}^c$

$$\begin{aligned} & \left( \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{kj} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ \\ &= \left( \prod_{j \leq n} \sum_{k \in \mathbb{N}} 1\{\eta_h^{kj} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ \\ &\leq \sum_{i \leq n} \left( \sum_{l \in \mathbb{N}} 1\{\eta_h^l B_{x_i}^\varepsilon > 0\} - 1 \right) \prod_{j \leq n} \sum_{k \in \mathbb{N}} 1\{\eta_h^k B_{x_j}^\varepsilon > 0\} \\ &= \sum_{i \leq n} \sum_{(k,l) \in \mathbb{N}^{(n+1)}} 1\{\eta_h^l B_{x_i}^\varepsilon > 0\} \prod_{j \leq n} 1\{\eta_h^{kj} B_{x_j}^\varepsilon > 0\}, \end{aligned}$$

which agrees with the bound in (19). Now let  $q_{\mu,s}^m$  denote the continuous density of  $E_\mu \xi_s^{\otimes m}$  in Theorem 5.3. Since  $\Omega = (U_{h,\varepsilon} \cap A_{h,\varepsilon}) \cup U_{h,\varepsilon}^c \cup A_{h,\varepsilon}^c$ , we may combine the previous estimates and use Lemmas 5.6 and 7.1(i) to get, for  $d \geq 3$ ,

$$\begin{aligned} & E_\mu \left( \sum_{k \in \mathbb{N}^{(n)}} \prod_{j \leq n} 1\{\eta_h^{kj} B_{x_j}^\varepsilon > 0\} - 1 \right)_+ - n^n P_\mu U_{h,\varepsilon} \\ &\leq \sum_{i \leq n} E_\mu \int \int \zeta_s^{(n+1)}(du dv) 1\{\eta_h^v B_{x_i}^\varepsilon > 0\} \prod_{j \leq n} 1\{\eta_h^{uj} B_{x_j}^\varepsilon > 0\} \\ &= \sum_{i \leq n} \int \int E_\mu \xi_s^{\otimes (n+1)}(du dv) P_v\{\eta_h B_{x_i}^\varepsilon > 0\} \prod_{j \leq n} P_{u_j}\{\eta_h B_{x_j}^\varepsilon > 0\} \\ &\leq \varepsilon^{(n+1)(d-2)} \sum_{i \leq n} \int \int q_{\mu,s}^{n+1}(u, v) du dv p_{h_\varepsilon}(x_i - v) \prod_{j \leq n} p_{h_\varepsilon}(x_j - u_j) \\ &= \varepsilon^{(n+1)(d-2)} \sum_{i \leq n} (q_{\mu,s}^{n+1} * p_{h_\varepsilon}^{\otimes (n+1)})(x, x_i) \\ &\leq \varepsilon^{(n+1)(d-2)} h^{1-d/2} \ll \varepsilon^{n(d-2)}. \end{aligned}$$

The term  $n^n P_\mu U_{h,\varepsilon}$  on the left is of the required order by Lemma 7.4. For  $d = 2$ , a similar argument, based on Lemma 7.1(ii), yields the bound  $|\log \varepsilon|^{-n-1} \times |\log h| \ll |\log \varepsilon|^{-n}$ .  $\square$

We may combine the last two lemmas into a useful approximation:

**COROLLARY 7.6.** *Fix any  $t, x$  and  $\mu$ , let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with  $h$ -clusters  $\eta_h^k$  at time  $t$  and let  $\gamma$  be a random element in a space  $T$ . Put  $B = (B_0^1)^n$ .*

Then as  $\varepsilon, h \rightarrow 0$  subject to (18), and for  $d \geq 3$  or  $d = 2$ , respectively,

$$\left\| \mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \gamma) - E_\mu \int \zeta_s^{(n)}(du) 1\{((\eta_h^{u_j} S_{x_j}^\varepsilon)_{j \leq n}, \gamma) \in \cdot\} \right\|_B \ll \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}. \end{cases}$$

PROOF. It is enough to establish the corresponding bounds for

$$E_\mu \left| f((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \gamma) - \int \zeta_s^{(n)}(du) f((\eta_h^{u_j} S_{x_j}^\varepsilon)_{j \leq n}, \gamma) \right|,$$

uniformly for  $H_B$ -measurable functions  $f$  on  $\mathcal{M}_d^n \times T$  with  $0 \leq f \leq 1_{H_B}$ . Writing  $\Delta_h^\varepsilon$  for the absolute value on the left,  $U_{h,\varepsilon}$  for the union in Lemma 7.4 and  $\kappa_h^\varepsilon$  for the sum in Lemma 7.5, we note that

$$\Delta_h^\varepsilon \leq \kappa_h^\varepsilon 1\{\kappa_h^\varepsilon > 1\} + 1\{\kappa_h^\varepsilon \leq 1; U_{h,\varepsilon}\}.$$

Since  $k 1\{k > 1\} \leq 2(k - 1)_+$  for any  $k \in \mathbb{Z}_+$ , we get

$$E_\mu \Delta_h^\varepsilon \leq 2E_\mu(\kappa_h^\varepsilon - 1)_+ + P_\mu U_{h,\varepsilon},$$

which is of the required order by Lemmas 7.4 and 7.5.  $\square$

We proceed to estimate the contribution of  $\xi_t$  to the balls  $B_{x_i}^\varepsilon$  from distantly rooted  $h$ -clusters. Define  $B_x^r = \bigcup_{j \leq n} B_{x_j}^r$  for  $x = (x_j) \in (\mathbb{R}^d)^n$ .

LEMMA 7.7. Fix any  $t, r > 0, x \in (\mathbb{R}^d)^{(n)}$  and  $\mu$ , and put  $B_x = \bigcup_{j \leq n} B_{x_j}^r$ . Then as  $\varepsilon, h \rightarrow 0$ , with  $\varepsilon^2 \leq h$  for  $d \geq 3$  and  $\varepsilon \leq h$  for  $d = 2$ ,

$$E_\mu \int_{B_x^c} \zeta_s^{(n)}(du) \prod_{j \leq n} 1\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\} \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

PROOF. Let  $q_{\mu,s}^n$  denote the jointly continuous density of  $E_\mu \xi_s^{\otimes n}$  from Theorem 5.3. For  $d \geq 3$  we may use the conditional independence and Lemma 7.1(i) to write the ratio between the two sides as

$$\begin{aligned} & \varepsilon^{n(2-d)} \int_{B_x^c} E_\mu \xi_s^{\otimes n}(du) \prod_{j \leq n} P_{u_j} \{\eta_h B_{x_j}^\varepsilon > 0\} \\ & \leq \int_{B_x^c} q_{\mu,s}^n(u) p_{h_\varepsilon}^{\otimes n}(x - u) du \\ & = (q_{\mu,s}^n * p_{h_\varepsilon}^{\otimes n})(x) - \int_{B_0} q_{\mu,s}^n(x - u) p_{h_\varepsilon}^{\otimes n}(u) du, \end{aligned}$$

where  $h_\varepsilon = h + \varepsilon^2$ . Here the first term on the right tends to  $q_{\mu,t}^n(x)$  as  $h_\varepsilon \rightarrow 0$  by Lemma 5.6, and the same limit is obtained for the second term by the joint continuity in Theorem 5.3 and elementary estimates. Hence, the difference tends

to 0. For  $d = 2$ , the same argument yields a similar bound with  $\varepsilon^{n(d-2)}$  replaced by  $|\log(\varepsilon^2/h)|^{-n}$ . Since  $0 < \varepsilon \leq h \rightarrow 0$ , we have  $|\log(\varepsilon^2/h)| \geq |\log \varepsilon|$ , and the assertion follows.  $\square$

Next we estimate the probability that a closed set  $G^c \subset \mathbb{R}^d$  is hit by some  $h$ -cluster in  $\xi_t$ , rooted within a compact subset  $B \subset G$ . For our present purposes, it suffices with a bound that tends to 0 as  $h \rightarrow 0$  faster than any power of  $h$ .

LEMMA 7.8. *Let  $t = s + h$  with  $0 < h \leq s$ , let  $G \subset \mathbb{R}^d$  be open with a compact subset  $B$ , and let  $r$  denote the minimum distance between  $B$  and  $G^c$ . Write  $\zeta_s$  for the ancestral process of  $\xi_t$  at time  $s$ . Then*

$$P_\mu \left\{ \int_B \zeta_s(du) \eta_h^u G^c > 0 \right\} \leq r^d h^{-1-d/2} e^{-r^2/2h} (\mu * \nu_t) B.$$

PROOF. Letting  $h' = h^{1/2} < r/2$ , we get by Theorem 3.3(b) in [3]

$$\begin{aligned} P_\mu \left\{ \int_B \zeta_s(du) \eta_h^u G^c > 0 \right\} &\leq P_\mu \left\{ \int_B \zeta_s(du) \eta_{h'}^u (B'_u)^c > 0 \right\} \\ &= E_\mu P_{(\xi_s B) \delta_0} \{ \xi_h (B'_0)^c > 0 \} \\ &\leq E_\mu \xi_s B r^{-2} (r/h')^{d+2} e^{r^2/2h} \\ &\leq (\mu * \nu_t) B r^d h^{-1-d/2} e^{-r^2/2h}. \end{aligned}$$

Here the first step holds by the definition of  $r$ , the second step holds by conditional independence and shift invariance, and the last step holds since  $E_\mu \xi_s = \mu * \nu_s$  and  $p_s \leq p_t$  when  $s \in [t/2, t]$ .  $\square$

The last two lemmas yield a useful decoupling approximation:

COROLLARY 7.9. *For a DW-process  $\xi$  in  $\mathbb{R}^d$  and times  $t$  and  $s = t - h$ , choose  $\tilde{\xi}_t \perp\!\!\!\perp_{\xi_s} \xi_t$  with  $(\xi_s, \xi_t) \stackrel{d}{=} (\xi_s, \tilde{\xi}_t)$ . Let  $G \subset \mathbb{R}^d$  be open, put  $B = (B_0^1)^n \times G^c$ , and fix any  $x \in G^{(n)}$ . Then as  $\varepsilon, h \rightarrow 0$  subject to (18)*

$$\| \mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t) - \mathcal{L}_\mu((\tilde{\xi}_t S_{x_j}^\varepsilon)_{j \leq n}, \tilde{\xi}_t) \|_B \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

PROOF. Letting  $U_x^r = \bigcup_j B_{x_j}^r$ , fix any  $r > 0$  with  $U_x^{2r} \subset G$ , and write  $\xi_t = \xi'_t + \xi''_t$  and  $\tilde{\xi}_t = \tilde{\xi}'_t + \tilde{\xi}''_t$ , where  $\xi'_t$  is the sum of clusters in  $\xi_t$  rooted in  $U_x^r$ , and similarly for  $\tilde{\xi}'_t$ . Putting  $D_n = \mathbb{R}^{nd} \setminus (\mathbb{R}^d)^{(n)}$  and letting  $\zeta_s$  be the ancestral process

of  $\xi_t$  at time  $s$ , we get

$$\begin{aligned} & \| \mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t) - \mathcal{L}_\mu((\xi'_t S_{x_j}^\varepsilon)_{j \leq n}, \xi''_t) \|_B \\ & \leq P_\mu \left\{ \prod_{j \leq n} \xi_t B_{x_j}^\varepsilon > \prod_{j \leq n} \xi'_t B_{x_j}^\varepsilon \right\} + P_\mu \{ \xi'_t G^c > 0 \} \\ & \leq P_\mu \left\{ \int_{D_n} \zeta_s^{\otimes n}(du) \prod_{j \leq n} \eta_h^{u_j} B_{x_j}^\varepsilon > 0 \right\} \\ & \quad + E_\mu \int_{(B'_x)^c} \zeta_s^{(n)}(du) \prod_{j \leq n} 1\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\} \\ & \quad + P_\mu \left\{ \int_{U'_x} \zeta_s(du) \eta_h^u G^c > 0 \right\} \ll \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}, \end{cases} \end{aligned}$$

by Lemmas 7.4, 7.7 and 7.8. Since the last estimate only depends on the marginal distributions of the pairs  $((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t)$  and  $((\xi'_t S_{x_j}^\varepsilon)_{j \leq n}, \xi''_t)$ , the same bound applies to

$$\| \mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \tilde{\xi}_t) - \mathcal{L}_\mu((\xi'_t S_{x_j}^\varepsilon)_{j \leq n}, \tilde{\xi}''_t) \|_B.$$

It remains to note that  $(\xi'_t, \xi''_t) \stackrel{d}{=} (\tilde{\xi}'_t, \tilde{\xi}''_t)$ .  $\square$

**8. Scaling limits and local approximation.** Here we study the pseudo-random measures  $\tilde{\xi}$  and  $\tilde{\eta}$ , which provide local approximations of the DW-process  $\xi$  in  $\mathbb{R}^d$  and its canonical cluster  $\eta$ . We begin with some scaling properties of  $\xi$  and  $\eta$ . Given a suitable measure-valued process  $\eta$  in  $\mathbb{R}^d$ , we define the *span* of  $\eta$  from 0 as the random variable

$$\rho = \inf\{r > 0; \sup_t \eta_t(B'_0)^c = 0\}.$$

Recall that a DW-cluster has a.s. finite span. For any process  $X$  on  $\mathbb{R}_+$ , we define the process  $\hat{X}_h$  on  $\mathbb{R}_+$  by  $(\hat{X}_h)_t = X_{ht}$ . Let  $\mathcal{L}_x[\eta \|\eta_h]_y$  denote the continuous versions of the Palm distributions described in [2, 4].

LEMMA 8.1. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ , and let  $\rho$  denote the span of  $\eta$  from 0. Then for any  $\mu, x \in \mathbb{R}^d$  and  $r, c > 0$ :*

- (i)  $\mathcal{L}_{\mu S_r}(r^2 \xi_1) = \mathcal{L}_{r^2 \mu}(\xi_{r^2} S_r)$ ;
- (ii)  $\mathcal{L}_{\mu S_r}(r^2 \eta) = r^2 \mathcal{L}_\mu(\hat{\eta}_{r^2} S_r)$ ;
- (iii)  $\mathcal{L}_0[r^2 \eta \|\eta_1]_x = \mathcal{L}_0[\hat{\eta}_{r^2} S_r \|\eta_{r^2}]_x$ ;
- (iv)  $P_0[\rho > c \|\eta_{r^2}]_x \leq P_0[r\rho + |x| > c \|\eta_1]_0$ .

Though (i) and (ii) are probably known, they are included here with short proofs for easy reference.

PROOF. (i) If  $v$  solves the evolution equation for  $\xi$ , then so does  $\tilde{v}(t, x) = r^2 v(r^2 t, rx)$ . Writing  $\tilde{\xi}_t = r^{-2} \xi_{r^2 t} S_r$ ,  $\tilde{\mu} = r^{-2} \mu S_r$ , and  $\tilde{f}(x) = r^2 f(rx)$ , we get

$$E_\mu e^{-\tilde{\xi}_t \tilde{f}} = E_\mu e^{-\xi_{r^2 t} f} = e^{-\mu v_{r^2 t}} = e^{-\tilde{\mu} \tilde{v}_t} = E_{\tilde{\mu}} e^{-\tilde{\xi}_t \tilde{f}},$$

and so  $\mathcal{L}_\mu(\tilde{\xi}) = \mathcal{L}_{\tilde{\mu}}(\tilde{\xi})$ , which is equivalent to (i); cf. [7], page 51.

(ii) Define the cluster kernel  $v$  by  $v_x = \mathcal{L}_x(\eta)$ ,  $x \in \mathbb{R}^d$ , and consider the cluster decomposition  $\xi = \int m \zeta(dm)$ , where  $\zeta$  is a Poisson process with intensity  $\mu v$  when  $\xi_0 = \mu$ . Here

$$r^{-2} \xi_{r^2 t} S_r = \int (r^{-2} m_{r^2 t} S_r) \zeta(dm), \quad r, t > 0.$$

Using (i) and the uniqueness of the Lévy measure, we obtain  $(r^{-2} \mu S_r)v = \mu(v\{r^{-2} \hat{m}_{r^2} S_r \in \cdot\})$ , which is equivalent to

$$r^{-2} \mathcal{L}_{\mu S_r}(\eta) = \mathcal{L}_{r^{-2} \mu S_r}(\eta) = \mathcal{L}_\mu(r^{-2} \hat{\eta}_{r^2} S_r).$$

(iii) By Palm disintegration, we get, from (ii),

$$\begin{aligned} \int E_0 \eta_1(dx) E_0[f(x, r^2 \eta) | \eta_1]_x &= E_0 \int \eta_1(dx) f(x, r^2 \eta) \\ &= r^2 E_0 \int \eta_{r^2}(dx) f(x, \hat{\eta}_{r^2} S_r) \\ &= r^2 \int E_0 \eta_{r^2}(dx) E_0[f(x, \hat{\eta}_{r^2} S_r) | \eta_{r^2}]_x, \end{aligned}$$

and (iii) follows by the continuity of the Palm kernel.

(iv) By (iii) we have  $\mathcal{L}_0[\rho | \eta_{r^2}]_x = \mathcal{L}_0[r \rho | \eta_1]_x$ , and (iv) follows for  $x = 0$ . For general  $x$  it is enough to take  $r = 1$ . Then recall from Corollary 4.1.6 in [4] or Theorem 11.7.1 in [2] that, under  $\mathcal{L}_0[\eta | \eta_1]_x$ , the cluster  $\eta$  is a countable sum of conditionally independent subclusters, rooted along the path of a Brownian bridge on  $[0, 1]$  from 0 to  $x$ . In particular, the evolution of the process after time 1 is the same as under the original distribution  $\mathcal{L}(\eta)$  (cf. Lemma 3.13), hence independent of  $x$ . We may now construct a cluster with distribution  $\mathcal{L}_0[\eta | \eta_1]_x$  from one with distribution  $\mathcal{L}_0[\eta | \eta_1]_0$ , simply by shifting every subcluster born at time  $s \leq 1$  by an amount  $(1 - s)x$ . Since all mass of  $\eta$  is then shifted by at most  $|x|$ , the span from 0 of the entire cluster is increased by at most  $|x|$ , and the assertion follows. □

We may now summarize the basic properties of the pseudo-random measure  $\tilde{\xi}$ , introduced in Section 8 of [19]. Here and below, we write

$$\mathcal{L}_\mu^0(\xi_t) = \mathcal{L}_\mu[\xi_t | \xi_t]_0, \quad \mathcal{L}_\mu^0(\eta_t) = \mathcal{L}_\mu[\eta_t | \eta_t]_0,$$

where  $\mathcal{L}_\mu[\xi_t | \xi_t]_x$  and  $\mathcal{L}_\mu[\eta_t | \eta_t]_x$  denote the continuous versions of the Palm distributions constructed in [2, 4]. Since the pseudo-random measures  $\tilde{\xi}$  and  $\tilde{\eta}$

below are stationary by definition, we may further take  $\mathcal{L}^0(\tilde{\xi})$  and  $\mathcal{L}^0(\tilde{\eta})$  to be the unique invariant versions of the associated Palm distributions  $\mathcal{L}[\theta_{-x}\tilde{\xi} \|\tilde{\xi}\|_x$  and  $\mathcal{L}[\theta_{-x}\tilde{\eta} \|\tilde{\eta}\|_x$ , respectively.

**THEOREM 8.2.** *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with  $d \geq 3$ . Then there exists a pseudo-random measure  $\tilde{\xi}$  on  $\mathbb{R}^d$ , such that:*

(i) *as  $\varepsilon \rightarrow 0$  for fixed  $\mu$  and  $t > 0$*

$$\|\mathcal{L}_\mu(\xi_t) - \mu p_t \mathcal{L}(\tilde{\xi})\|_{B_0^\varepsilon} \ll \varepsilon^{d-2}, \quad \|\mathcal{L}_\mu^0(\xi_t) - \mathcal{L}^0(\tilde{\xi})\|_{B_0^\varepsilon} \rightarrow 0,$$

*and similarly with  $\xi_t$  replaced by  $\eta_t$ ;*

(ii) *for any  $r > 0$ ,*

$$\mathcal{L}(\tilde{\xi} S_r) = r^{d-2} \mathcal{L}(r^2 \tilde{\xi}), \quad \mathcal{L}^0(\tilde{\xi} S_r) = \mathcal{L}^0(r^2 \tilde{\xi});$$

(iii)  *$\tilde{\xi}$  is stationary with  $E\tilde{\xi} = \lambda^{\otimes d}$ ;*

(iv)  *$\mathcal{L}(\tilde{\xi})$  is an invariant measure for  $\xi$ ;*

(v) *as  $r \rightarrow \infty$ , we have in total variation on  $H_B$  for bounded  $B$*

$$r^{d-2} \mathcal{L}_{r^2-d\lambda^{\otimes d}}(\xi_{r^2}) \rightarrow \mathcal{L}(\tilde{\xi}).$$

**PROOF.** (i)–(ii) In Theorems 8.1–2 of [19] we proved the existence of a stationary pseudo-random measure  $\tilde{\xi}$  on  $\mathbb{R}^d$  satisfying (ii), and such that as  $\varepsilon \rightarrow 0$  for bounded  $B \in \mathcal{B}^d$ ,

$$(20) \quad \begin{aligned} \|\varepsilon^{2-d} \mathcal{L}_\mu(\varepsilon^{-2}\xi_t S_\varepsilon) - \mu p_t \mathcal{L}(\tilde{\xi})\|_B &\rightarrow 0, \\ \|\mathcal{L}_\mu^0(\varepsilon^{-2}\xi_t S_\varepsilon) - \mathcal{L}^0(\tilde{\xi})\|_B &\rightarrow 0, \end{aligned}$$

and similarly with  $\xi_t$  replaced by  $\eta_t$ . Under (ii), the latter properties are equivalent to (i).

(iii) In our proof of Theorem 8.2 in [19] (display (20) in [19], page 2210) we showed that for any  $B \in \hat{\mathcal{B}}^d$

$$\|\varepsilon^{2-d} E_\mu[\varepsilon^{-2}\xi_t B_0^\varepsilon; \varepsilon^{-2}\xi_t S_\varepsilon \in \cdot] - \mu p_t E[\tilde{\xi} B_0^1; \tilde{\xi} \in \cdot]\|_B \rightarrow 0.$$

Taking  $B = B_0^1$  and  $\mu = \lambda^{\otimes d}$ , we get, in particular,  $\varepsilon^{-d} E_{\lambda^{\otimes d}} \xi_t S_\varepsilon B_0^1 \rightarrow E\tilde{\xi} B_0^1$ , which extends by stationarity to arbitrary  $B$ . Hence,

$$\lambda^{\otimes d} = \varepsilon^{-d} \lambda^{\otimes d} S_\varepsilon = \varepsilon^{-d} E_{\lambda^{\otimes d}}(\xi_t S_\varepsilon) \rightarrow E\tilde{\xi},$$

and so  $E\tilde{\xi} = \lambda^{\otimes d}$ .

(iv) Let  $(\tilde{\xi}_t)$  denote the DW-process  $\xi$  with initial measure  $\tilde{\xi}$ . Using (ii) and Lemma 8.1(i), we get for any  $r > 0$ ,

$$\begin{aligned} \mathcal{L}(\tilde{\xi}_{r^2} S_r) &= E\mathcal{L}_{\tilde{\xi}}(\xi_{r^2} S_r) = E\mathcal{L}_{r^{-2}\tilde{\xi} S_r}(r^2 \xi_1) \\ &= r^{d-2} E\mathcal{L}_{\tilde{\xi}}(r^2 \xi_1) = r^{d-2} \mathcal{L}(r^2 \tilde{\xi}_1) = \mathcal{L}(\tilde{\xi} S_r), \end{aligned}$$

which implies  $\tilde{\xi}_{r,2} S_r \stackrel{d}{=} \tilde{\xi} S_r$ . Hence,  $\tilde{\xi}_t \stackrel{d}{=} \tilde{\xi}$  for all  $t \geq 0$ .

(v) Using Lemma 8.1(i) and (20) above and noting that  $\lambda^{\otimes d} S_r = r^d \lambda^{\otimes d}$ , we get, as  $r \rightarrow \infty$ ,

$$r^{d-2} \mathcal{L}_{r^{2-d} \lambda^{\otimes d}}(\xi_{r,2}) = r^{d-2} \mathcal{L}_{\lambda^{\otimes d}}(r^2 \xi_1 S_{1/r}) \rightarrow \mathcal{L}(\tilde{\xi}). \quad \square$$

For  $d = 2$ , there is no random measure  $\tilde{\xi}$  with the stated properties. However, a similar role is then played by the *stationary cluster*  $\tilde{\eta}_t$  with pseudo-distribution  $\mathcal{L}(\tilde{\eta}_t) = \mathcal{L}_{\lambda^{\otimes d}}(\eta_t)$ . Writing  $\tilde{\eta} = \tilde{\eta}_1$ , we have the following approximation and scaling properties:

**THEOREM 8.3.** *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ . Then as  $\varepsilon \rightarrow 0$  for fixed  $\mu$  and  $t > 0$ :*

$$(i) \quad \|\mathcal{L}_\mu(\xi_t) - \mu p_t \mathcal{L}(\tilde{\eta})\|_{B_0^\varepsilon} \ll \begin{cases} \varepsilon^{d-2}, & d \geq 3, \\ |\log \varepsilon|^{-1}, & d = 2, \end{cases}$$

$$(ii) \quad \|\mathcal{L}_\mu^0(\xi_t) - \mathcal{L}^0(\tilde{\eta})\|_{B_0^\varepsilon} \rightarrow 0, \quad d \geq 2,$$

and similarly with  $\xi_t$  replaced by  $\eta_t$ . Furthermore,

(iii) for any  $r > 0$ ,

$$\mathcal{L}(\tilde{\eta}_{r,2} S_r) = r^{d-2} \mathcal{L}(r^2 \tilde{\eta}), \quad \mathcal{L}^0(\tilde{\eta}_{r,2} S_r) = \mathcal{L}^0(r^2 \tilde{\eta}),$$

(iv) for  $d \geq 3$  and as  $t \rightarrow \infty$ , in total variation on  $H_B$  for bounded  $B$ ,

$$\mathcal{L}(\tilde{\eta}_t) \rightarrow \mathcal{L}(\tilde{\xi}).$$

Though for  $d \geq 3$  the pseudo-random measures  $\tilde{\xi}$  and  $\tilde{\eta}$  have many similar properties, we note that  $\tilde{\eta}$  has weaker scaling properties. The Palm distributions  $\mathcal{L}_\mu^0(\xi_t)$  and  $\mathcal{L}_\mu^0(\eta_t)$  both exist, since the common intensity measure  $E_\mu \xi_t = E_\mu \eta_t$  is locally finite. Our proof of Theorem 8.3 requires a simple comparison of  $\mathcal{L}_\mu(\xi_t)$  and  $\mathcal{L}_\mu(\eta_t)$ .

**LEMMA 8.4.** *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with canonical cluster  $\eta$ . Then for any  $\mu, B$  and  $t > 0$ :*

$$(i) \quad P_\mu\{\eta_t B > 0\} = -\log(1 - P_\mu\{\xi_t B > 0\});$$

$$(ii) \quad \|\mathcal{L}_\mu(\xi_t) - \mathcal{L}_\mu(\eta_t)\|_B \leq (P_\mu\{\eta_t B > 0\})^2.$$

In particular,  $P_\mu\{\xi_t B > 0\} \sim P_\mu\{\eta_t B > 0\}$  as either side tends to 0.

PROOF. (i) See Lemma 4.1 in [19].

(ii) Using the cluster representation  $\xi_t = \int m \zeta_t(dm)$ , where  $\zeta_t$  is a Poisson process on  $\mathcal{M}_d$  with intensity measure  $\mathcal{L}_\mu(\eta_t)$ , we get

$$1_B \xi_t = \int 1_B m \zeta_t(dm) = \int 1_B m \zeta_t^B(dm),$$

where  $\zeta_t^B$  denotes the restriction of  $\zeta_t$  to the set of measures  $m$  with  $mB > 0$ . For any measurable function  $f \geq 0$  on  $\mathcal{M}_d$  with  $f(0) = 0$ , Lemma 2.10 yields

$$|E_\mu f(1_B \xi_t) - E_\mu f(1_B \eta_t)| \leq \|f\| (P_\mu\{\eta_t B > 0\})^2,$$

and the assertion follows since  $f$  is arbitrary.  $\square$

PROOF OF THEOREM 8.3. Some crucial ideas in the following proof are adapted from the corresponding arguments in Section 8 of [19].

(i) For  $d \geq 3$ , the assertion follows from Theorem 8.2(i), applied to  $\xi_t$  under  $P_\mu$  and to  $\eta_1$  under  $P_{\lambda \otimes d}$ . Now let  $d = 2$ . Fixing  $t > 0$ ,  $\mu \in \hat{\mathcal{M}}_2$ , and  $B \in \hat{\mathcal{B}}^2$ , writing  $\eta_h^i$  for the  $h$ -clusters of  $\xi_t$  with associated point process  $\zeta_s$  of ancestors at time  $s = t - h$  and letting  $\varepsilon, h \rightarrow 0$  with  $|\log h| \ll |\log \varepsilon|$ , we get as in [19] [display (12)], for any  $\mathcal{H}_B$ -measurable function  $f$  with  $0 \leq f \leq 1_{H_B}$ ,

$$\begin{aligned} E_\mu f(\xi_t S_\varepsilon) &= E_\mu f\left(\sum_i \eta_h^i S_\varepsilon\right) \approx E_\mu \sum_i f(\eta_h^i S_\varepsilon) \\ &= E_\mu \int \zeta_s(dx) f(\eta_h^x S_\varepsilon) = \int E_\mu \xi_s(dx) E_x f(\eta_h S_\varepsilon) \\ &= \int \mu(dy) \int p_s(x - y) E_x f(\eta_h S_\varepsilon) dx \\ &\approx \mu p_t \int E_x f(\eta_h S_\varepsilon) dx = \mu p_t E f(h \tilde{\eta} S_{\varepsilon/\sqrt{h}}), \end{aligned}$$

where the third equality holds by the conditional independence of the clusters and the Cox nature of  $\zeta_s$ , and the last equality holds by Lemma 8.1(ii).

As for the first approximation, we get, by Lemma 7.3,

$$\begin{aligned} E_\mu \left| f\left(\sum_i \eta_h^i S_\varepsilon\right) - \sum_i f(\eta_h^i S_\varepsilon) \right| &\leq E_\mu [\kappa_h^{c\varepsilon}, \kappa_h^{c\varepsilon} > 1] \\ &\leq \frac{\log(t/h) \mu p_t + (\mu p_{t(h,\varepsilon)})^2}{|\log \varepsilon|^2} \\ &\leq \frac{|\log h| + 1}{|\log \varepsilon|^2} \ll |\log \varepsilon|^{-1}, \end{aligned}$$

where  $\kappa_h^\varepsilon$  denotes the number of clusters  $\eta_h^i$  hitting  $B_0^\varepsilon$ . For the second approximation, we get, by Lemma 7.1(ii) as  $\varepsilon \leq h \rightarrow 0$ ,

$$\begin{aligned} & \left| \int \mu(dy) \int (p_s(y-x) - p_t(y)) E_x f(\eta_h S_\varepsilon) dx \right| \\ & \leq |\log(\varepsilon^2/h)|^{-1} \int \mu(dy) \int |p_s(y-x) - p_t(y)| p_{h_{ce}}(x) dx \\ & \leq |\log \varepsilon|^{-1} \int \mu(dy) E |p_s(y - \gamma h_{ce}^{1/2}) - p_t(y)|, \end{aligned}$$

where  $\gamma$  denotes a standard normal random vector in  $\mathbb{R}^d$ . Since  $p_s(y - \gamma h_{ce}^{1/2}) \rightarrow p_t(y)$  by the joint continuity of  $p_t(x)$  and

$$E p_s(y - \gamma h_{ce}^{1/2}) = (p_s * p_{h_{ce}})(y) = p_{s+h_{ce}}(y) \rightarrow p_t(y),$$

the last expectation tends to 0 by Lemma 1.32 in [15], and so the integral on the right tends to 0 by dominated convergence.

In summary, noting that both approximations are uniform in  $f$ , we get as  $\varepsilon, h \rightarrow 0$  with  $|\log h| \ll |\log \varepsilon|$

$$(21) \quad \|\mathcal{L}_\mu(\xi_t S_\varepsilon) - \mu p_t \mathcal{L}(h \tilde{\eta} S_{\varepsilon/\sqrt{h}})\|_B \ll |\log \varepsilon|^{-1},$$

which extends to unbounded  $\mu$  by an easy truncation argument. Furthermore, Lemmas 7.2(ii) and 8.4 yield, as  $\varepsilon \rightarrow 0$ ,

$$(22) \quad |E_\mu f(\xi_t S_\varepsilon) - E_\mu f(\eta_t S_\varepsilon)| \leq (P_\mu\{\eta_t B_0^{ce} > 0\})^2 \leq |\log \varepsilon|^{-2}.$$

Hence, (21) remains valid with  $\xi_t$  replaced by  $\eta_t$ . Now (i) follows as we take  $t = 1$  and  $\mu = \lambda^{\otimes 2}$  and combine with (21). The corresponding result for  $\eta_t$  follows by means of (22).

(ii) Once again, the statement for  $d \geq 3$  follows from Theorem 8.2(i). For  $d = 2$ , we see from Lemma 7.2(ii) above and Lemma 3.4 in [19] that as  $\varepsilon \rightarrow 0$ , for fixed  $t > 0$ ,

$$\begin{aligned} P_\mu\{\xi_t B_0^\varepsilon > 0\} & \asymp |\log \varepsilon|^{-1} \mu p_t, \\ E_\mu(\xi_t B_0^\varepsilon)^2 & \asymp \varepsilon^4 |\log \varepsilon| (\lambda^{\otimes 2} B_0^1)^2 \mu p_t \asymp \varepsilon^4 |\log \varepsilon| \mu p_t, \end{aligned}$$

and similarly for  $\eta_t$ . Hence,

$$E_\mu[(\xi_t B_0^\varepsilon)^2 | \xi_t B_0^\varepsilon > 0] = \frac{E_\mu(\xi_t B_0^\varepsilon)^2}{P_\mu\{\xi_t B_0^\varepsilon > 0\}} \asymp \varepsilon^4 |\log \varepsilon|^2,$$

and similarly for  $\eta_t$ . Thus,  $\xi_t B_0^\varepsilon / \varepsilon^2 |\log \varepsilon|$  is uniformly integrable, conditionally on  $\xi_t B_0^\varepsilon > 0$ , and correspondingly for  $\eta_t$  under both  $P_\mu$  and  $P_{\lambda^{\otimes 2}}$ . Noting that, by (i),

$$\|\mathcal{L}_\mu[\xi_t S_\varepsilon | \xi_t B_0^\varepsilon > 0] - \mathcal{L}[\tilde{\eta} S_\varepsilon | \tilde{\eta} B_0^\varepsilon > 0]\|_{B_0^1} \rightarrow 0,$$

we obtain

$$\|E_\mu[\xi_t B_0^\varepsilon; \xi_t S_\varepsilon \in \cdot | \xi_t B_0^\varepsilon > 0] - E[\tilde{\eta} B_0^\varepsilon; \tilde{\eta} S_\varepsilon \in \cdot | \tilde{\eta} B_0^\varepsilon > 0]\|_{B_0^1} \ll \varepsilon^2 |\log \varepsilon|,$$

and so, by (i),

$$\|E_\mu[\xi_t B_0^\varepsilon; \xi_t S_\varepsilon \in \cdot] - \mu p_t E[\tilde{\eta} B_0^\varepsilon; \tilde{\eta} S_\varepsilon \in \cdot]\|_{B_0^1} \ll \varepsilon^2,$$

and similarly for  $\eta_t$ . Next, Lemma 4.1 yields

$$E_\mu \xi_t B_0^\varepsilon = \lambda^{\otimes 2} (\mu * p_t) 1_{B_0^\varepsilon} \asymp \varepsilon^2 \mu p_t.$$

Combining the last two estimates with Lemma 6.6, and using Lemma 3.11 in a version for pseudo-random measures, we obtain the desired convergence.

(iii) Use Lemma 8.1(ii)–(iii).

(iv) From (iii) and Theorem 8.2(i)–(ii), we get, as  $r \rightarrow \infty$ ,

$$\mathcal{L}(\tilde{\eta}_{r,2}) = r^{d-2} \mathcal{L}(r^2 \tilde{\eta} S_{1/r}) \rightarrow \mathcal{L}(\tilde{\xi}). \quad \square$$

Though the scaling properties of  $\tilde{\eta}$  are weaker than those of  $\tilde{\xi}$  when  $d \geq 3$ ,  $\tilde{\eta}$  does satisfy a strong continuity property under scaling, which extends Lemma 5.1 in [19].

**THEOREM 8.5.** *Let  $\tilde{\eta}$  be the stationary cluster of a DW-process in  $\mathbb{R}^2$ , and define a kernel  $\nu$  from  $(0, \infty)$  to  $\mathcal{M}_2$  by*

$$\nu(r) = |\log r| \mathcal{L}(r^{-2} \tilde{\eta} S_r), \quad r > 0.$$

*Then the kernel  $t \mapsto \nu(\exp(-e^t))$  is uniformly continuous on  $[1, \infty)$ , in total variation on  $H_B$  for bounded  $B$ .*

**PROOF.** For any  $\varepsilon, r, h \in (0, 1)$ , let  $\zeta_s$  denote the ancestral process of  $\xi_1$  at time  $s = 1 - h$ , and let  $\eta_h^u$  be the  $h$ -clusters rooted at the associated atoms at  $u$ . Then

$$\begin{aligned} |\log \varepsilon|^{-1} \nu(\varepsilon) &\approx r^{-1} \mathcal{L}_{r\lambda^{\otimes 2}}(\varepsilon^{-2} \xi_1 S_\varepsilon) \approx r^{-1} E_{r\lambda^{\otimes 2}} \int \zeta_s(du) 1\{\varepsilon^{-2} \eta_h^u S_\varepsilon \in \cdot\} \\ &= \int \mathcal{L}_u(\varepsilon^{-2} \eta_h S_\varepsilon) du = \mathcal{L}(\varepsilon^{-2} h \tilde{\eta} S_{\varepsilon/\sqrt{h}}) \\ &\approx |\log \varepsilon|^{-1} \nu(\varepsilon/\sqrt{h}), \end{aligned}$$

with all relations explained and justified below. The first equality holds by the conditional independence of the clusters and the fact that  $E_{r\lambda^{\otimes 2}} \zeta_s = (r/h)\lambda^{\otimes 2}$ . The second equality follows from Lemma 8.1(i) by an elementary substitution.

To estimate the error in first approximation, we see from Lemmas 7.1(ii) and 8.4 that for  $\varepsilon < \frac{1}{2}$ ,

$$\begin{aligned} \|r |\log \varepsilon|^{-1} \nu(\varepsilon) - \mathcal{L}_{r\lambda^{\otimes 2}}(\varepsilon^{-2} \xi_1 S_\varepsilon)\|_B &= \|\mathcal{L}_{r\lambda^{\otimes 2}}(\eta_1 S_\varepsilon) - \mathcal{L}_{r\lambda^{\otimes 2}}(\xi_1 S_\varepsilon)\|_B \\ &\leq (r P\{\tilde{\eta}(\varepsilon B) > 0\})^2 \leq r^2 |\log \varepsilon|^{-2}, \end{aligned}$$

where  $\eta_h^1, \eta_h^2, \dots$  are the  $h$ -clusters of  $\xi$  at time  $t$ . As for the second approximation, we get, by Lemma 7.3 for small enough  $\varepsilon/h$ ,

$$\begin{aligned} & \left\| E_{r\lambda^{\otimes 2}} \sum_k 1\{\eta_h^k S_\varepsilon \in \cdot\} - \mathcal{L}_{r\lambda^{\otimes 2}}(\xi_1 S_\varepsilon) \right\|_B \\ & \leq E_{r\lambda^{\otimes 2}} \left( \sum_k 1_{+}(\eta_h^k(\varepsilon B)) - 1 \right)_{+} \\ & \leq \frac{|\log h| r \lambda^{\otimes 2} p_1 + (r \lambda^{\otimes 2} p_{t(h,\varepsilon)})^2}{|\log \varepsilon|^2} \\ & = r \frac{|\log h| + r}{|\log \varepsilon|^2}. \end{aligned}$$

The third approximation relies on the estimate

$$\|v(\varepsilon/\sqrt{h})\|_B \left| \frac{|\log \varepsilon|}{|\log(\varepsilon/\sqrt{h})|} - 1 \right| \leq \frac{|\log h|}{|\log \varepsilon|},$$

which holds for  $\varepsilon \leq h$  by the boundedness of  $v$ . Combining those estimates and letting  $r \rightarrow 0$  gives

$$(23) \quad \|v(\varepsilon) - v(\varepsilon/\sqrt{h})\|_B \leq \frac{|\log h|}{|\log \varepsilon|}, \quad \varepsilon \ll h < 1.$$

Putting  $\varepsilon = e^{-u}$  and  $\varepsilon/\sqrt{h} = e^{-v}$  and writing  $v_A(x) = v(x, A)$  for measurable sets  $A \subset H_B$ , we get for  $u - v \ll u$  (with  $0/0 = 1$ ),

$$\left| \log \frac{v_A(e^{-u})}{v_A(e^{-v})} \right| \leq \left| \frac{v_A(e^{-u})}{v_A(e^{-v})} - 1 \right| \leq |v_A(e^{-u}) - v_A(e^{-v})| \leq \frac{u - v}{u} \leq \left| \log \frac{u}{v} \right|,$$

and so, for  $u = e^s$  and  $v = e^t$  (with  $\infty - \infty = 0$ ),

$$|\log v_A(\exp(-e^t)) - \log v_A(\exp(-e^s))| \leq |t - s|,$$

which extends immediately to arbitrary  $s, t \geq 1$ . Since  $v$  is bounded on  $H_B$  by Lemma 7.2, the function  $v_A(\exp(-e^t))$  is again uniformly continuous on  $[1, \infty)$ , and the assertion follows since all estimates are uniform in  $A$ .  $\square$

The exact scaling properties of  $\tilde{\eta}$  in Theorem 8.3(iii) may be supplemented by the following asymptotic age invariance, which may be compared with the exact age invariance of  $\tilde{\xi}$  in Theorem 8.2(iv).

**COROLLARY 8.6.** *Let  $\varepsilon \rightarrow 0$  and  $h > 0$  with  $\varepsilon^2 \ll h \ll \varepsilon^{-2}$  for  $d \geq 3$  and  $|\log \varepsilon| \gg |\log h|$  for  $d = 2$ . Then, as  $\varepsilon \rightarrow 0$ ,*

$$\|\mathcal{L}(\tilde{\eta}_h) - \mathcal{L}(\tilde{\eta}_1)\|_{B_0^\varepsilon} \ll \begin{cases} \varepsilon^{d-2}, & d \geq 3, \\ |\log \varepsilon|^{-1}, & d = 2. \end{cases}$$

PROOF. Fix any  $B \in \hat{\mathcal{B}}^d$ . For  $d \geq 3$ , we get by Theorems 8.2 and 8.3

$$\begin{aligned} & \|\mathcal{L}(\tilde{\eta}S_\varepsilon) - r^{d-2}\mathcal{L}(r^2\tilde{\eta}S_{\varepsilon/r})\|_B \\ & \leq \|\mathcal{L}(\tilde{\eta}S_\varepsilon) - \mathcal{L}(\tilde{\xi}S_\varepsilon)\|_B + r^{d-2}\|\mathcal{L}(\tilde{\eta}S_{\varepsilon/r}) - \mathcal{L}(\tilde{\xi}S_{\varepsilon/r})\|_B \\ & \ll \varepsilon^{d-2} + r^{d-2}(\varepsilon/r)^{d-2} \leq \varepsilon^{d-2}. \end{aligned}$$

When  $d = 2$ , we may use (23) instead to get

$$\begin{aligned} & |\log \varepsilon| \|\mathcal{L}(\tilde{\eta}S_\varepsilon) - \mathcal{L}(r^2\tilde{\eta}S_{\varepsilon/r})\|_B \\ & \leq \|v(\varepsilon) - v(\varepsilon/r)\|_B + \left| \frac{|\log \varepsilon|}{|\log(\varepsilon/r)|} - 1 \right| \|v(\varepsilon/r)\|_B \\ & \leq \frac{|\log r|}{|\log \varepsilon|} + \frac{|\log r|}{|\log \varepsilon| - |\log r|} \rightarrow 0. \end{aligned}$$

It remains to note that  $r^{d-2}\mathcal{L}(r^2\tilde{\eta}S_{1/r}) = \mathcal{L}(\tilde{\eta}_{r^2})$  by Theorem 8.3(iii).  $\square$

**9. Local conditioning and global approximation.** Here we state and prove our main approximation theorem, which contains multivariate versions of the local approximations in Section 8 and shows how the multivariate Palm distributions of a DW-process can be approximated by elementary conditional distributions.

Given a DW-process  $\xi$  in  $\mathbb{R}^d$ , let  $\tilde{\xi}$  and  $\tilde{\eta}$  be the associated pseudo-random measures from Theorems 8.2 and 8.3. Let  $q_{\mu,t}$  denote the continuous versions of the moment densities of  $E_\mu \xi_t^{\otimes n}$  from Theorem 5.3, and write  $\mathcal{L}_\mu[\xi_t \|\xi_t^{\otimes n}]_x$  for the regular, multivariate Palm distributions considered in Theorem 6.3. Define  $c_d$  and  $m_\varepsilon$  as in Lemma 7.2. Write  $f \sim g$  for  $f/g \rightarrow 1$ ,  $f \approx g$  for  $f - g \rightarrow 0$ , and  $f \ll g$  for  $f/g \rightarrow 0$ . The notation  $\|\cdot\|_B$  with associated terminology is explained in Section 1 above.

**THEOREM 9.1.** *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$  with  $d \geq 2$ , and let  $\varepsilon \rightarrow 0$  for fixed  $\mu, t > 0$  and open  $G \subset \mathbb{R}^d$ . Then:*

(i) *for any  $x \in (\mathbb{R}^d)^{(n)}$ ,*

$$P_\mu\{\xi_t^n B_x^\varepsilon > 0\} \sim q_{\mu,t}(x) \begin{cases} c_d^n \varepsilon^{n(d-2)}, & d \geq 3, \\ m_\varepsilon^n |\log \varepsilon|^{-n}, & d = 2; \end{cases}$$

(ii) *for any  $x \in G^{(n)}$ , in total variation on  $(B_0^1)^n \times G^c$ ,*

$$\mathcal{L}_\mu[(\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t \|\xi_t^{\otimes n} B_x^\varepsilon > 0] \approx \mathcal{L}^{\otimes n}[\tilde{\eta}S_\varepsilon | \tilde{\eta}B_0^\varepsilon > 0] \otimes \mathcal{L}_\mu[\xi_t \|\xi_t^{\otimes n}]_x;$$

(iii) *for  $d \geq 3$  we have, in the same sense,*

$$\mathcal{L}_\mu[(\varepsilon^{-2}\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t \|\xi_t^{\otimes n} B_x^\varepsilon > 0] \rightarrow \mathcal{L}^{\otimes n}[\tilde{\xi} | \tilde{\xi}B_0^1 > 0] \otimes \mathcal{L}_\mu[\xi_t \|\xi_t^{\otimes n}]_x.$$

Here (i) extends some asymptotic results for  $n = 1$  from [3, 19, 27]. Parts (ii) and (iii) show that, asymptotically as  $\varepsilon \rightarrow 0$ , the contributions of  $\xi_t$  to the sets  $B_{x_1}^\varepsilon, \dots, B_{x_n}^\varepsilon$  and  $G^c$  are conditionally independent. They further imply the multivariate Palm approximation

$$\mathcal{L}_\mu[1_{G^c} \xi_t | \xi_t^{\otimes n} B_x^\varepsilon > 0] \rightarrow \mathcal{L}_\mu[1_{G^c} \xi_t | \xi_t^{\otimes n}]_x, \quad x \in G^{(n)},$$

and they contain the asymptotic equivalence or convergence on  $B_0^1$  for any  $x \in (\mathbb{R}^d)^{(n)}$ ,

$$\mathcal{L}_\mu[\xi_t S_{x_j}^\varepsilon | \xi_t^{\otimes n} B_x^\varepsilon > 0] \begin{cases} \approx \mathcal{L}[\tilde{\eta} S_\varepsilon | \tilde{\eta} B_0^\varepsilon > 0], & d \geq 2, \\ \rightarrow \mathcal{L}[\tilde{\xi} | \tilde{\xi} B_0^1 > 0], & d \geq 3, \end{cases}$$

extending the versions for  $n = 1$  implicit in Theorems 8.2 and 8.3. Analogous results for simple point processes and regenerative sets appear in [16, 22].

Given (i), assertions (ii) and (iii) are essentially equivalent to the following estimate, which we prove first. Here and below,  $q_{\mu,t}(x) = q_{\mu,t}^x$ .

LEMMA 9.2. *Let  $\xi$  be a DW-process in  $\mathbb{R}^d$ , fix any  $\mu, t > 0$  and open  $G \subset \mathbb{R}^d$ , and put  $B = (B_0^1)^n \times G^c$ . Then, as  $\varepsilon \rightarrow 0$  for fixed  $x \in G^{(n)}$ ,*

$$\|\mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t) - q_{\mu,t}^x \mathcal{L}^{\otimes n}(\tilde{\eta} S_\varepsilon) \otimes \mathcal{L}_\mu[\xi_t | \xi_t^{\otimes n}]_x\|_B \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

PROOF. We may regard  $\xi_t$  as a sum of conditionally independent clusters  $\eta_h^u$  of age  $h \in (0, t)$ , rooted at the points  $u$  of the ancestral process  $\zeta_s$  at time  $s = t - h$ . Choose the random measure  $\xi'_t$  to satisfy

$$\xi'_t \perp\!\!\!\perp_{\xi_s} (\xi_t, \zeta_s, (\eta_h^u)), \quad (\xi_s, \xi_t) \stackrel{d}{=} (\xi_s, \xi'_t).$$

Our argument can be summarized as follows:

$$\begin{aligned} \mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t) &\approx E_\mu \int \zeta_s^{(n)}(du) 1_{(\cdot)}((\eta_h^{u_j} S_{x_j}^\varepsilon)_{j \leq n}, \xi'_t) \\ &= \int E_\mu \xi_s^{\otimes n}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes \mathcal{L}_\mu[\xi_t | \xi_s^{\otimes n}]_u \\ (24) \quad &\approx (E_\mu \xi_s^{\otimes n} * p_h^{\otimes n})_x \mathcal{L}^{\otimes n}(\tilde{\eta}_h S_\varepsilon) \otimes \mathcal{L}_\mu[\xi_t | \xi_t^{\otimes n}]_x \\ &\approx q_{\mu,t}^x \mathcal{L}^{\otimes n}(\tilde{\eta} S_\varepsilon) \otimes \mathcal{L}_\mu[\xi_t | \xi_t^{\otimes n}]_x, \end{aligned}$$

where  $h$  and  $\varepsilon$  are related as in (8.18), and the approximations hold in the sense of total variation on  $H_B$  of the order  $\varepsilon^{n(d-2)}$  or  $|\log \varepsilon|^{-n}$ , respectively. Detailed justifications are given below.

The first relation in (24) is immediate from Corollaries 7.6 and 7.9. To justify the second relation, we provide some intermediate steps:

$$\begin{aligned}
 & E_\mu \int \zeta_s^{(n)}(du) 1_{(\cdot)}((\eta_h^{u_j} S_{x_j}^\varepsilon)_{j \leq n}, \xi_t') \\
 &= E_\mu \int \zeta_s^{(n)}(du) \mathcal{L}_\mu[(\eta_h^{u_j} S_{x_j}^\varepsilon)_{j \leq n}, \xi_t' | \xi_s, \zeta_s] \\
 &= h^n E_\mu \int \zeta_s^{(n)}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes \mathcal{L}_\mu[\xi_t' | \xi_s] \\
 &= E_\mu \int \xi_s^{(n)}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes \mathcal{L}_\mu[\xi_t | \xi_s] \\
 &= E_\mu \int \xi_s^{(n)}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes 1_{(\cdot)}(\xi_t) \\
 &= \int E_\mu \xi_s^{(n)}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes \mathcal{L}_\mu[\xi_t \| \xi_s^{\otimes n}]_u.
 \end{aligned}$$

Here the first and fourth equalities hold by disintegration and Fubini's theorem. The second relation holds by the conditional independence of the  $h$ -clusters and the process  $\xi_t'$ , along with the normalization of  $\mathcal{L}(\eta)$ . The third relation holds by the choice of  $\xi_t'$  and the moment relation  $E_\mu[\zeta_s^{(n)} | \xi_s] = h^{-n} \xi_s^{\otimes n}$  from [22]. The fifth relation holds by Palm disintegration.

To justify the third relation in (24), we first consider a change in the last factor. By Lemmas 3.4 and 7.1,

$$\begin{aligned}
 & \left\| \int E_\mu \xi_s^{\otimes n}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \otimes (\mathcal{L}_\mu[\xi_t \| \xi_s^{\otimes n}]_u - \mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x) \right\|_B \\
 & \leq \int E_\mu \xi_s^{\otimes n}(du) p_{h_\varepsilon}^{\otimes n}(x - u) \| \mathcal{L}_\mu[\xi_t \| \xi_s^{\otimes n}]_u - \mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x \|_{G^c} \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}, \end{cases}
 \end{aligned}$$

with  $h_\varepsilon$  defined as in Lemma 7.1 with  $t$  replaced by  $h$ . Choosing  $r > 0$  with  $B_x^{2r} \subset G^{(n)}$ , we may estimate the integral on the right by

$$\begin{aligned}
 & (E_\mu \xi_s^{\otimes n} * p_{h_\varepsilon}^{\otimes n}(x)) \sup_{u \in B_x^r} \| \mathcal{L}_\mu[\xi_t \| \xi_s^{\otimes n}]_u - \mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x \|_{G^c} \\
 & + \int_{(B_x^r)^c} E_\mu \xi_s^{\otimes n}(du) p_{h_\varepsilon}^{\otimes n}(x - u).
 \end{aligned}$$

Here the first term tends to 0 by Lemmas 5.6 and 6.5, whereas the second term tends to 0 as in the proof of Lemma 7.7. Hence, in the second line of (24), we may replace  $\mathcal{L}_\mu[\xi_t \| \xi_s^{\otimes n}]_u$  by  $\mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x$ .

By a similar argument based on Lemma 3.4 and Corollary 8.6, we may next replace  $\mathcal{L}(\tilde{\eta}S_\varepsilon)$  in the last line by  $\mathcal{L}(\tilde{\eta}_h S_\varepsilon)$ . It is then enough to prove that

$$\int E_\mu \xi_s^{\otimes n}(du) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_{x_j}^\varepsilon) \approx q_t(x) \mathcal{L}^{\otimes n}(\tilde{\eta}_h S_\varepsilon),$$

where  $q_t$  denotes the continuous density of  $E_\mu \xi_t^{\otimes n}$  in Theorem 5.3. Here the total variation distance may be expressed in terms of densities as

$$\begin{aligned} & \left\| \int (q_s(x-u) - q_t(x)) \bigotimes_{j \leq n} \mathcal{L}_{u_j}(\eta_h S_\varepsilon) du \right\|_B \\ & \leq \int |q_s(x-u) - q_t(x)| p_{h_\varepsilon}^{\otimes n}(u) du \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}. \end{cases} \end{aligned}$$

Letting  $\gamma$  be a standard normal random vector in  $\mathbb{R}^{nd}$ , we may write the integral on the right as  $E|q_s(x - \gamma h_\varepsilon^{1/2}) - q_t(x)|$ . Here  $q_s(x - \gamma h_\varepsilon^{1/2}) \rightarrow q_t(x)$  a.s. by the joint continuity of  $q_s(u)$ , and Lemma 5.6 yields

$$E q_s(x - \gamma h_\varepsilon^{1/2}) = (q_s * p_{h_\varepsilon}^{\otimes n})(x) \rightarrow q_t(x).$$

The former convergence then extends to  $L^1$  by Lemma 1.32 in [15], and the required approximation follows.  $\square$

PROOF OF THEOREM 9.1. (i) For any  $x \in (\mathbb{R}^d)^{(n)}$ , Lemma 9.2 yields

$$|P_\mu\{\xi_t^{\otimes n} B_x^\varepsilon > 0\} - q_{\mu,t}^x(P\{\tilde{\eta} B_0^\varepsilon > 0\})^n| \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

It remains to note that, by Lemma 7.2,

$$P\{\tilde{\eta} B_0^\varepsilon > 0\} \sim \begin{cases} c_d \varepsilon^{d-2}, & d \geq 3, \\ m(\varepsilon) |\log \varepsilon|^{-1}, & d = 2. \end{cases}$$

(ii) Assuming  $x \in G^{(n)}$  and using (i) and Lemma 9.2, we get, in total variation on  $(B_0^1)^n \times G^c$ ,

$$\begin{aligned} \mathcal{L}_\mu[(\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t | \xi_t^{\otimes n} B_x^\varepsilon > 0] &= \frac{\mathcal{L}_\mu((\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t)}{P_\mu\{\xi_t^{\otimes n} B_x^\varepsilon > 0\}} \\ &\approx \frac{q_{\mu,t}^x \mathcal{L}^{\otimes n}(\tilde{\eta} S_\varepsilon) \otimes \mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x}{q_{\mu,t}^x (P\{\tilde{\eta} B_0^\varepsilon > 0\})^n} \\ &= \mathcal{L}^{\otimes n}[\tilde{\eta} S_\varepsilon | \tilde{\eta} B_0^\varepsilon > 0] \otimes \mathcal{L}_\mu[\xi_t \| \xi_t^{\otimes n}]_x. \end{aligned}$$

(iii) When  $d \geq 3$ , Theorem 8.2(i) yields

$$\varepsilon^{2-d} \mathcal{L}(\varepsilon^{-2} \tilde{\eta} S_\varepsilon) \rightarrow \mathcal{L}(\tilde{\xi}),$$

in total variation on  $B_0^1$ . Hence,

$$\mathcal{L}[\varepsilon^{-2}\tilde{\eta}S_\varepsilon|\tilde{\eta}B_0^\varepsilon > 0] \rightarrow \mathcal{L}[\tilde{\xi}|\tilde{\xi}B_0^1 > 0],$$

and the assertion follows by means of (ii).  $\square$

**Acknowledgment.** My sincere thanks to the referees for their careful reading and many helpful remarks.

## REFERENCES

- [1] DALEY, D. J. and VERE-JONES, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure*, 2nd ed. Springer, New York. [MR2371524](#)
- [2] DAWSON, D. A. (1993). Measure-valued Markov processes. In *École D'Été de Probabilités de Saint-Flour XXI—1991. Lecture Notes in Math.* **1541** 1–260. Springer, Berlin. [MR1242575](#)
- [3] 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. [MR1012498](#)
- [4] DAWSON, D. A. and PERKINS, E. A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93** (454) iv+179. [MR1079034](#)
- [5] DYNKIN, E. B. (1991). Branching particle systems and superprocesses. *Ann. Probab.* **19** 1157–1194. [MR1112411](#)
- [6] DYNKIN, E. B. (1994). *An Introduction to Branching Measure-valued Processes. CRM Monograph Series* **6**. Amer. Math. Soc., Providence, RI. [MR1280712](#)
- [7] ETHERIDGE, A. M. (2000). *An Introduction to Superprocesses. University Lecture Series* **20**. Amer. Math. Soc., Providence, RI. [MR1779100](#)
- [8] GOROSTIZA, L. G., ROELLY-COPPOLETTA, S. and WAKOLBINGER, A. (1990). Sur la persistance du processus de Dawson–Watanabe stable. L'interversion de la limite en temps et de la renormalisation. In *Séminaire de Probabilités, XXIV, 1988/89. Lecture Notes in Math.* **1426** 275–281. Springer, Berlin. [MR1071545](#)
- [9] GOROSTIZA, L. G. and WAKOLBINGER, A. (1991). Persistence criteria for a class of critical branching particle systems in continuous time. *Ann. Probab.* **19** 266–288. [MR1085336](#)
- [10] JAGERS, P. (1973). On Palm probabilities. *Z. Wahrsch. Verw. Gebiete* **26** 17–32. [MR0339330](#)
- [11] KALLENBERG, O. (1977). Stability of critical cluster fields. *Math. Nachr.* **77** 7–43. [MR0443078](#)
- [12] KALLENBERG, O. (1986). *Random Measures*, 4th ed. Akademie-Verlag, Berlin. [MR0854102](#)
- [13] KALLENBERG, O. (1999). Palm measure duality and conditioning in regenerative sets. *Ann. Probab.* **27** 945–969. [MR1698987](#)
- [14] KALLENBERG, O. (2001). Local hitting and conditioning in symmetric interval partitions. *Stochastic Process. Appl.* **94** 241–270. [MR1840832](#)
- [15] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. [MR1876169](#)
- [16] KALLENBERG, O. (2003). Palm distributions and local approximation of regenerative processes. *Probab. Theory Related Fields* **125** 1–41. [MR1952454](#)
- [17] KALLENBERG, O. (2005). *Probabilistic Symmetries and Invariance Principles*. Springer, New York. [MR2161313](#)
- [18] KALLENBERG, O. (2007). Some problems of local hitting, scaling, and conditioning. *Acta Appl. Math.* **96** 271–282. [MR2327541](#)
- [19] KALLENBERG, O. (2008). Some local approximations of Dawson–Watanabe superprocesses. *Ann. Probab.* **36** 2176–2214. [MR2478680](#)

- [20] KALLENBERG, O. (2009). Some local approximation properties of simple point processes. *Probab. Theory Related Fields* **143** 73–96. [MR2449123](#)
- [21] KALLENBERG, O. (2010). Commutativity properties of conditional distributions and Palm measures. *Commun. Stoch. Anal.* **4** 21–34. [MR2643111](#)
- [22] KALLENBERG, O. (2011). Iterated Palm conditioning and some Slivnyak-type theorems for Cox and cluster processes. *J. Theoret. Probab.* **24** 875–893. [MR2822485](#)
- [23] KUMMER, G. and MATTHES, K. (1970). Verallgemeinerung eines Satzes von Sliwnjak. II–III. *Rev. Roumaine Math. Pures Appl.* **15** 845–870, 1631–1642. [MR0270416](#)
- [24] LE GALL, J.-F. (1986). Sur la saucisse de Wiener et les points multiples du mouvement brownien. *Ann. Probab.* **14** 1219–1244. [MR0866344](#)
- [25] LE GALL, J.-F. (1986). Une approche élémentaire des théorèmes de décomposition de Williams. In *Séminaire de Probabilités XX. Lecture Notes in Math.* **1204**, 447–464. Springer, Berlin.
- [26] LE GALL, J.-F. (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.* **19** 1399–1439. [MR1127710](#)
- [27] LE GALL, J.-F. (1994). A lemma on super-Brownian motion with some applications. In *The Dynkin Festschrift. Progress in Probability* **34** 237–251. Birkhäuser, Boston, MA. [MR1311723](#)
- [28] LE GALL, J.-F. (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Birkhäuser, Basel. [MR1714707](#)
- [29] LIEMANT, A., MATTHES, K. and WAKOLBINGER, A. (1988). *Equilibrium Distributions of Branching Processes. Mathematical Research* **42**. Akademie-Verlag, Berlin. [MR0969126](#)
- [30] MATTHES, K., KERSTAN, J. and MECKE, J. (1978). *Infinitely Divisible Point Processes*. Wiley, Chichester. [MR0517931](#)
- [31] MECKE, J. (1967). Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. *Z. Wahrsch. Verw. Gebiete* **9** 36–58. [MR0228027](#)
- [32] PERKINS, E. (2002). Dawson–Watanabe superprocesses and measure-valued diffusions. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1999). Lecture Notes in Math.* **1781** 125–324. Springer, Berlin. [MR1915445](#)
- [33] SALISBURY, T. S. and VERZANI, J. (1999). On the conditioned exit measures of super Brownian motion. *Probab. Theory Related Fields* **115** 237–285. [MR1720367](#)
- [34] SALISBURY, T. S. and VERZANI, J. (2000). Nondegenerate conditionings of the exit measures of super Brownian motion. *Stochastic Process. Appl.* **87** 25–52. [MR1751163](#)
- [35] SLIVNYAK, I. M. (1962). Some properties of stationary flows of homogeneous random events. *Theory Probab. Appl.* **7** 336–341; **9** 168.
- [36] WILLIAMS, D. (1974). Path decomposition and continuity of local time for one-dimensional diffusions. I. *Proc. London Math. Soc.* (3) **28** 738–768. [MR0350881](#)
- [37] ZÄHLE, U. (1988). The fractal character of localizable measure-valued processes. II. Localizable processes and backward trees. *Math. Nachr.* **137** 35–48. [MR0968985](#)

DEPT. OF MATHEMATICS AND STATISTICS  
AUBURN UNIVERSITY  
221 PARKER HALL, AUBURN  
ALABAMA 36849  
USA  
E-MAIL: kalleoh@auburn.edu