The Annals of Statistics
1996, Vol. 24, No. 3, 1161–1184

ASYMPTOTIC LIKELIHOOD ESTIMATION FROM BIRTH AND DEATH ON A FLOW[1]

BY MICHAEL J. PHELAN

University of Pennsylvania

Birth and death on a flow refers to a particle system on a Brownian flow. Particles are born in a point process and move on the flow subject to position-dependent killing. They die eventually and leave the flow. The particle process is a measure-valued Markov process tracking these motions. Its law depends on the distribution of births, the coefficients of the flow and the rate of killing. We treat asymptotic likelihood estimation of these parameters from chronicles of the particle process as observed over a long period of time.

1. Introduction. In their 1992 model of mass transport, Çinlar and Kao imagined particles on a turbulent fluid flow. With a Brownian flow describing the fluid-flow map of this fluid over its domain, a Poisson process regulates the birth of particles that live and die there. The particles move with the motion of the flow. They dissipate in response to position-dependent killing or decay. Eventually, they die and leave the flow.

Çinlar and Kao (1992a, b) studied the configuration of live particles on the flow in terms of a measure-valued Markov process. The law of this particle process depends on the distribution of births, the coefficients of the flow and the rate of killing.

Phelan (1996) developed likelihood methods for parametric estimation of the system parameters from chronicles of the particle process as observed over a fixed period of time. In this work, we treat asymptotic, maximum-likelihood estimation from chronicles of the particle process as observed over a long period of time. As in our earlier work, the martingale dynamics over the particle process play a central role.

Our results recall in particular the work of Feigin (1976), Basawa, Feigen and Heyde (1976), Godambe and Heyde (1987), Barndorff-Nielsen and Sørensen (1994) and Sørensen (1989, 1990) on likelihood methods for stochastic processes. The addition here lies in treating a spatial stochastic process.

2. Birth and death on a Brownian flow. We recall Phelan's (1996) statistical model of birth and death on a Brownian flow. Of course, we take our description of the process itself from that of Çinlar and Kao (1992). The

Received October 1993; revised October 1994.
[1]Supported by NSF Grant DMS-92-07094 and ONR Grant N00014-93-0216.
AMS 1991 *subject classifications*. Primary 62M05; secondary 62M30.
Key words and phrases. Brownian flow, Poisson process, maximum likelihood, asymptotic estimation.

probability space $(\Omega, \mathscr{H}, \mathbb{P})$ supports all of the random variables that appear in the following text. The set E denotes Euclidean space of dimension d.

A particle system. The particle system is a countable system of particles in motion on a Brownian flow $F = (F_{st})$, $0 \le s \le t \le \infty$, on E. The life process L is an independent Poisson random measure on $\mathbb{R}_+ \times E \times \mathbb{R}_+$ whose atoms identify particles that enter, live and die on the flow. That is, let (S, X, U) denote an atom of L, identifying a generic particle. Our interpretation is that the particle enters the flow at random time S and random position X. The trajectory $t \to F_{St}X$ describes its motion on the flow. This motion stops when the particle dies and leaves the flow at time T,

$$T = \inf\left\{t > S \colon \int_S^t dr\, k(F_{Sr}X) > U\right\},$$

where U is the particle's intrinsic lifetime and k is the position-dependent rate of killing.

The particle process $t \to M_t$ is a Markov process on the space of counting measures on E. It tracks the position of the living particles on the flow. According to Çinlar and Kao's (1992b) master formula, the particle process satisfies the equation

$$M_t f = \int L(ds, dx, du) f(F_{st}x) 1_{[0,t]}(s) 1_{[0,u)}\left(\int_s^t dr\, k(F_{sr}x)\right) \tag{1}$$

for every function f in the space $\mathbb{C}_K$ of continuous functions having compact support in E. Since all particles move on the same flow, contemporaneous particles move dependently.

We suppose that F is a Brownian flow of homeomorphisms on E having infinitesimal mean b and infinitesimal covariance c. The mean b is a mapping from E to E. We suppose that there is a mapping γ from E to the space of $d \times m$ matrices such that c is the mapping $(x, y) \to \gamma(x)\gamma^{\mathsf{T}}(y)$, where T denotes transpose. The coefficients b and γ satisfy the global Lipschitz condition

$$|b(x) - b(y)| + \|\gamma(x) - \gamma(y)\| \le K|x - y| \tag{2}$$

and the linear growth condition

$$|b(x)| + \|\gamma(x)\| \le K(1 + |x|) \tag{3}$$

for some constant K. We refer to Kunita (1990) for more on flows.

The life process L is a Poisson random measure on $\mathbb{R}_+ \times E \times \mathbb{R}_+$. Its mean measure λ satisfies the equation

$$\lambda(ds, dx, du) = \delta_0(\{s\})\mu_0(dx)\,du\,e^{-u} + ds\,\pi(dx)\,du\,e^{-u} \tag{4}$$

for every $s \ge 0$, $x \in E$, $u \ge 0$, where δ_0 is the Dirac measure at zero and μ_0 and π are finite measures on E. Thus, the placement of the initial particles on the flow has distribution μ_0 on E. Thereafter, particles enter the flow at rate $\pi(E)$ while having distribution π on E. The intrinsic lifetime of all particles is exponential with mean 1. We refer to Karr (1986) for more on random measures.

For later reference, we notice that the one-point motions are diffusions having Markov generator A that satisfies the equation

$$Af(x) = \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} c_{ij}(x, x) \frac{\partial^2 f}{\partial x_i \, \partial x_j}(x) \tag{5}$$

for every function f having two continuous derivatives.

A statistical model. A statistical model here refers to a sample space, an indexed family of probability laws for the particle system and an observational scheme.

The particle process takes values in the space $(M_b, \mathscr{M}_b)$ of bounded counting measures on E with its Borel sigma algebra relative to the vague topology. Its sample paths belong to the Skorokhod space $\mathbb{D}(M_b)$ of right-continuous functions from $\mathbb{R}_+$ to M_b having vague limits from the left. The coordinate mappings on this space generate a filtration $\mathbf{G} = (\mathscr{G}_t)$, $t \geq 0$, and the filtered space $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G})$. This space is the canonical setting for the particle process and our sample space.

We specify a family of laws for the particle process as indexed by a finite-dimensional parameter on the system parameters themselves. The parameter set Θ is an open, bounded subset of Euclidean space of dimension p. As a convenience, we assume it contains the origin as an interior point. The zero theta is the parameter of reference.

For θ in Θ, we introduce the system parameter $(b^\theta, c, \pi^\theta, k^\theta)$ and the probability space $(\Omega, \mathscr{H}, \mathbb{P}^\theta)$ for the life-flow pair (L, F); the covariance c is intentionally free of theta. The life process L is a Poisson random measure having mean measure λ^θ satisfying (4) with π^θ as the distribution of births. The flow F is an independent Brownian flow having infinitesimal mean b^θ and infinitesimal covariance c.

The life-flow pair and the killing function k^θ induce the particle process $t \to M_t^\theta$ as at (1). The particle process has sample paths in the Skorokhod space $\mathbb{D}(M_b)$, so we may speak of its law there. The law of the particle process refers to the probability P^θ, namely, $P^\theta = \mathbb{P}^\theta \circ (M^\theta)^{-1}$, on the filtered space $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G})$.

For a strictly positive time T, the observational scheme is to observe the mapping $t \to M_t$ on $[0, T]$, yielding a chronicle of the particle process or the sigma algebra $\mathscr{G}_T$ as data. We suppose that we know the law of these data only up to theta and consider the problem of estimating theta from $\mathscr{G}_T$. Our approach is maximum-likelihood estimation from the score function as derived in Phelan (1996). We study this problem in the limit as T approaches infinity.

3. Likelihood estimation. Phelan (1996) considered the problem of maximum-likelihood estimation from birth and death on a Brownian flow. We recall his representation of the likelihood and the score function here.

Likelihood. We may view the counting measures on E as a subset of the space of distributions on E. In doing so, we may view the particle process as a semimartingale on the space of distributions. This view is important because it supports a stochastic calculus over the particle process, a special case of the calculus in Itô (1984). Indeed, each of the likelihood, the score and the information matrix have representations in terms of stochastic integrals with respect to the martingale parts of the particle process.

To introduce this, let $\mathscr{D}$ denote the space of infinitely differentiable functions having compact support on E. The dual $\mathscr{D}'$ of $\mathscr{D}$ is the space of distributions on E. We identify a Radon measure μ, for example, with a distribution that takes functions in $\mathscr{D}$ to their integral with respect to μ. As an operator on $\mathscr{D}'$, let A^{0*} denote the adjoint of the operator A^0 of (5) with $b = b^0$. For T in $\mathscr{D}'$, the distribution $A^{0*}T$ is the distribution satisfying the equation $(A^{0*}T)\phi = T(A^0\phi)$ for ϕ in $\mathscr{D}$.

On the stochastic basis $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^0)$, the particle process is by construction right continuous with left limits in the topology of vague convergence. So for $t > 0$, let M_{t-} denote the vague limit of M_s as s increases to t and let ΔM_t denote the signed-Dirac measure $M_t - M_{t-}$ on E. The set $\{(s, \Delta M_s)\colon s > 0,\ M_s \neq M_{s-}\}$ then induces a multivariate point process μ on $\mathbb{R}_+ \times M_{sd}$, where M_{sd} denotes the space of signed-Dirac measures on E. Its compensator ν^0 is a predictable random measure on $\mathbb{R}_+ \times M_{sd}$ satisfying the equation

$$\begin{aligned} &\nu^0(dt, d\eta) \\ &\quad = dt\left[\pi^0 \circ \delta^{-1}(d\eta)1_{\{+1\}}(\eta 1) + \left(k^0 M_{t-}\right) \circ \delta^{-1}(-d\eta)1_{\{-1\}}(\eta 1)\right] \end{aligned} \tag{6}$$

for every signed-Dirac measure η in M_{sd}, where δ is the mapping from E to the space of Dirac measures on E.

Therefore as a process on $\mathscr{D}'$, the particle process satisfies the stochastic integral equation

$$\begin{aligned} M_t &= M_0 + B_t^0 + X_t^c + \int_0^t \int_{M_{sd}} (\mu - \nu^0)(ds, d\eta)\eta, \\ B_t^0 &= \int_0^t ds\left[\pi^0 + k^0 M_{s-} + A^{0*} M_{s-}\right], \qquad t \geq 0, \end{aligned} \tag{7}$$

implicitly defining the continuous-martingale part $t \to X_t^c$ and the discrete-martingale part on the compensated point process $\mu - \nu^0$. The continuous-martingale part $t \to X_t^c$ associates with a characteristic q satisfying the equation

$$q(M_{s-}; \phi, \varphi) = \int_{E \times E} M_{s-}(dx) M_{s-}(dy) \phi(x) c(x, y) \varphi(y) \tag{8}$$

for every ϕ and φ in $\mathscr{D}$.

To specify the likelihood, we need to integrate with respect to X^c and $\mu - \nu^0$. So, we need the following regularity conditions.

CONDITION A. The birth measures π^{θ} and π^{0} are finite measures on $(E, \mathscr{E})$ and π^{θ} admits a density $x \to p^{\theta}(x)$ with respect to π^{0} for every θ in Θ.

CONDITION B. The drift coefficient b^{θ} satisfies the global Lipschitz condition of (2) and the linear growth condition of (3) for every θ in Θ. The coefficient b^{0} is zero; otherwise replace b^{θ} with $b^{\theta} - b^{0}$ in the sequel.

CONDITION C. The square root of c, namely, γ, satisfies the global Lipschitz condition of (2) and the linear growth condition of (3). In addition, we suppose that $x \to \gamma(x)$ is of full column rank on E. The rank condition ensures the existence of a generalized inverse at (10).

CONDITION D. The killing function k^{θ} is a Borel function such that $b \geq k^{\theta} \geq 0$ on E for some constant b, and k^{θ} and k^{0} vanish on the same set for every $\theta \in \Theta$.

This ends the conditions for now. In these conditions and in our treatment below, we treat the zero theta as a reference parameter. To us this seems simpler than carrying a θ_0 throughout the discussion.

We introduce the coefficients of the prelikelihood in Phelan (1996). For θ in Θ and referring to Conditions A and D, let $\eta \to y^{\theta}(\eta)$ denote the mapping on M_{sd} satisfying the equation

$$y^{\theta}(\eta) = p^{\theta} \circ \delta^{-1}(\eta) 1_{\{+1\}}(\eta 1) + \frac{k^{\theta} \circ \delta^{-1}(-\eta)}{k^{0} \circ \delta^{-1}(-\eta)} 1_{\{-1\}}(\eta 1), \tag{9}$$

where δ^{-1} maps Dirac measures on E to their atoms and $0/0$ is by convention zero. Referring to Condition C, let $(t, x) \to f^{\theta}(M_{t-}; x)$ denote the predictable mapping satisfying the equation

$$f^{\theta}(M_{t-}; x) = \gamma(x) \left[\int_E M_{t-}(du) \gamma^{\mathsf{T}}(u) \gamma(u) \right]^{-2} \int_E M_{t-}(dz) \gamma^{\mathsf{T}}(z) b^{\theta}(z) \tag{10}$$

if $M_{t-}(E) > 0$ and $f^{\theta} = 0$ otherwise. In this case, let h_t^{θ} denote the mapping $x \to f^{\theta}(M_{t-}; x)$ on E for every t in $\mathbb{R}_+$. As a process of functionals on distributions, identify the coefficient $t \to H_t^{\theta}$ with the predictable process $t \to h_t^{\theta}$ and the (random) field $\eta \to J^{\theta}(\eta)$ with the mapping $\eta \to y^{\theta}(\eta) - 1$. As a result, we introduce the process $t \to \xi_t^{\theta}$:

$$\xi_t^{\theta} = \int_0^t ds\, q\big(M_{s-}; h_s^{\theta}, h_s^{\theta}\big) + \int_0^t \int_{M_{sd}} \nu^{0}(ds, d\eta) \big(y^{\theta}(\eta) - 1\big)^2. \tag{11}$$

For θ in Θ, one says that P^{θ} is locally absolutely continuous with respect to P^{0} on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G})$ whenever the restriction of P^{θ} to $\mathscr{G}_t$ is absolutely continuous with respect to that restriction of P^{0} for every t in $\mathbb{R}_+$. In this case, there is a density process $t \to Z_t^{\theta}$, for example, for local change of measure from P^{0} to P^{θ}. The random variable Z_T^{θ} is then the likelihood of P^{θ}

relative to P^0 on the data $\mathscr{G}_T$. The mapping $\theta \to Z_T^\theta$ is the likelihood function over our statistical model.

The next proposition summarizes some results from Phelan (1996) on likelihoods for birth and death on a Brownian flow. The centered dot $(\cdot)$ notation denotes stochastic integration for measure-valued processes as defined in Itô (1984); the centered asterisk $(*)$ stands for multivariate point processes as defined in Jacod and Shiryaev (1987).

12. PROPOSITION. *We suppose that we have the regularity conditions A through D. For θ in Θ, we suppose that the process $t \to \xi_t^\theta$ is integrable with respect to both P^θ and P^0. The probability measure P^θ is then locally, absolutely continuous with respect to P^0 on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G})$. In addition, there is a locally square-integrable martingale $t \to Y_t^\theta$,*

$$Y_t^\theta = H^\theta \cdot X_t^c + J^\theta * (\mu - \nu^0)_t,$$

such that its Doléans–Dade exponential $t \to \mathscr{E}(Y^\theta)_t$,

$$\mathscr{E}(Y^\theta)_t = \exp\left(Y_t^\theta - \tfrac{1}{2}\int_0^t ds\, q\left(M_s; h_s^\theta, h_s^\theta\right)\right) \prod_{s \le t} \left(1 + \Delta Y_s^\theta\right)\exp\left(-\Delta Y_s^\theta\right),$$

is a local martingale on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^0)$ such that $\mathscr{E}(Y^\theta)_0 = 1$. The density process $t \to Z_t^\theta$ between P^θ and P^0 is then the process $t \to \mathscr{E}(Y^\theta)_t$ and the likelihood over our statistical model is the mapping $\theta \to \mathscr{E}(Y^\theta)_T$.

We refer to Phelan (1996) for regularity conditions and demonstration of this result.

Score. The score function is the derivative of the natural logarithm of the likelihood with respect to theta. The log-likelihood process $t \to L_t^\theta$ refers naturally to the process $t \to \ln Z_t^\theta$ for every theta. If $\theta \to L_t^\theta$ is differentiable for every t in $\mathbb{R}_+$ and D_θ denotes differentiation with respect to theta, then the score process refers to the process $t \to D_\theta L_t^\theta$ for every theta. The score function is then the mapping $\theta \to D_\theta L_T^\theta$.

We begin with a regularity condition on the system parameters that allow such differentiation. In particular, let g denote a generic function from $\Theta \times E$ into $\mathbb{R}$ or into E. Let D_θ denote the first-order differential operator on functions on Θ. For example, if $\theta \to g(\theta, x)$ is differentiable with respect to theta for every x in E, then $D_\theta g(\theta)$ denotes the mapping $x \to D_\theta g(\theta, x)$ on E.

For real-valued functions, our convention is that the operator D_θ induces the gradient vector as arranged in a row vector. For E-valued functions, our convention is that D_θ induces a matrix of order given by the coordinate dimension of E by that dimension of Θ, giving a rowwise arrangement of gradients.

CONDITION E. In the sense defined in the preceding text, the mappings $\theta \to b^\theta$, $\theta \to p^\theta$, $\theta \to \ln p^\theta$, $\theta \to r^\theta$, where $r^\theta = k^\theta / k^0$, and $\theta \to \ln r^\theta$ are twice continuously differentiable for every theta.

The stochastic basis is $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^0)$. For θ in Θ, Proposition 12 implies that there is a log-likelihood process $t \to L_t^\theta$ satisfying the equation

$$L_t^\theta = H^\theta \cdot X_t^c - \tfrac{1}{2}\int_0^t ds\, q\big(M_s; h_s^\theta, h_s^\theta\big) + \ln y^\theta * (\mu - \nu^0)_t + (\ln y^\theta - J^\theta) * \nu_t^0,$$

after modest reexpression of $\ln Z_t^\theta$.

Condition E implies that the coefficients in the log-likelihood are differentiable with respect to theta. In particular, for θ in Θ, there is a mapping $\eta \to D_\theta \ln y^\theta(\eta)$ on M_{sd} satisfying the equation

$$\begin{aligned} D_\theta \ln y^\theta(\eta) = {} & \big(D_\theta \ln p^\theta\big) \circ \delta^{-1}(\eta) 1_{\{+1\}}(\eta 1) \\ & + \big(D_\theta \ln r^\theta\big) \circ \delta^{-1}(-\eta) 1_{\{-1\}}(\eta 1), \end{aligned}$$

where r^θ is again the function k^θ / k^0. Similarly, there is a mapping $\eta \to D_\theta y^\theta(\eta)$ and so a mapping $\eta \to D_\theta J^\theta(\eta)$. Next, let $(t, x) \to f'^\theta(M_{t-}; x)$ denote the predictable mapping satisfying the equation

$$f'^\theta(M_{t-}; x) = \gamma(x)\left[\int_E M_{t-}(du)\gamma^{\mathsf{T}}(u)\gamma(u)\right]^{-2} \int_E M_{t-}(dz)\gamma^{\mathsf{T}}(z) D_\theta b^\theta(z)$$

if $M_{t-}(E) > 0$ and $f'^\theta = 0$ otherwise. In this case, let $D_\theta h_t^\theta$ denote the mapping $x \to f'^\theta(M_{t-}; x)$ on E for every t in $\mathbb{R}_+$. As a process of functionals on distributions, we then identify the process $t \to D_\theta H_t^\theta$ with the predictable process $t \to D_\theta h_t^\theta$ for the purpose of representing a stochastic integral in subsequent text.

13. PROPOSITION. *We suppose that we have the regularity conditions A through E. We suppose the X^c-integrability of $t \to D_\theta H_t^\theta$, the $(\mu - \nu^0)$-integrability of $D_\theta \ln y^\theta$ and the ν^0-integrability of $D_\theta \ln y^\theta - D_\theta^\theta J^\theta$ for every theta. On the stochastic basis $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^0)$, there is then a well-defined semimartingale $t \to S_t^\theta$,*

$$\begin{aligned} S_t^\theta = {} & \big(D_\theta H^\theta\big) \cdot X_t^c + \big(D_\theta \ln y^\theta\big) * (\mu - \nu^0)_t \\ & + \big(D_\theta \ln y^\theta - D_\theta J^\theta\big) * \nu_t^0 - \int_0^t ds\, q\big(M_{s-}; h_s^\theta, D_\theta h_s^\theta\big) \end{aligned}$$

for every theta. Next, we suppose there is an increasing sequence $n \to T_n$ of stopping times and a sequence $n \to c_n$ of constants such that T_n increases almost surely to infinity with n and such that we have the inequality:

$$E^0 \sup_{t < T_n} \int_\Theta d\theta \left\| \frac{1}{u}\big(L_t^{\theta+u} - L_t^\theta - u S_t^\theta\big)\right\|^2 \le c_n u^2$$

for every n. For every θ in Θ, the process $t \to S_t^\theta$ is then a representation of the score process $t \to D_\theta L_t^\theta$ of interest here.

The score process then has an integral representation with respect to the particle process obtained by formally interchanging differentiation and stochastic integration. We refer to Phelan (1996) for regularity conditions and demonstration of this result.

4. Observed information. We derive a type of Fisherian information that is integral to our asymptotic treatment of maximum-likelihood estimation. Known as the observed information, it requires two derivations of the log-likelihood function. In this section, we specify sufficient conditions in terms of the system parameters for the existence of the information process as well as a representation of it in terms of stochastic integrals over the particle process.

Regularity conditions. The regularity conditions here are differentiability conditions on the system parameters and controls on those derivatives. As in our preparation for Condition E in the second half of Section 3, let g denote a generic function from $\Theta \times E$ into $\mathbb{R}$ or into E. Let D_θ denote the first-order differential operator on functions on Θ, supposing the same conventions as before.

We also introduce the second-order $D^2_{\theta^2}$ and the third-order $D^3_{\theta^3}$ differential operators. In particular, let g denote a generic function from $\Theta \times E$ into $\mathbb{R}$ or into E. If $\theta \to g(\theta, x)$ is thrice differentiable with respect to theta for every x in E, then $D^2_{\theta^2} g(\theta)$ denotes the mapping $x \to D^2_{\theta^2} g(\theta, x)$ on E. Also, $D^3_{\theta^3} g(\theta)$ denotes the mapping $x \to D^3_{\theta^3} g(\theta, x)$ on E.

For real-valued functions, our convention is that the operator D_θ induces the gradient vector as arranged in a row vector. The operator $D^2_{\theta^2}$ then induces the usual matrix of mixed-partial derivatives. The operator $D^3_{\theta^3}$ induces a three-dimensional $p \times p \times p$ matrix of mixed-partial derivatives.

For E-valued functions, our convention is that D_θ induces a matrix of order given by the coordinate dimension of E by that dimension of Θ, giving a row-wise arrangement of gradients. The operator $D^2_{\theta^2}$ then induces a blocked matrix of mixed-partial derivatives obtained by differentiating the rows of gradients as induced by D_θ. Since E has d dimensions and Θ has p, the operator $D^2_{\theta^2}$ induces d blocks $[D^2_{\theta^2} g(\theta)]_1, \dots, [D^2_{\theta^2} g(\theta)]_d$ of $p \times p$ matrices. The third-order operator induces d blocks $[D^3_{\theta^3} g(\theta)]_1, \dots, [D^3_{\theta^3} g(\theta)]_d$ of $p \times p \times p$ matrices.

There are three regularity conditions of interest here. The first specifies excessive measures for the transition semigroup of the one-, two- and three-point motions on the flow. The second imposes a differentiability condition on the system parameters, the last integrability conditions on pairs of controlling functions.

CONDITION F. For $n = 1, 2, 3$, let $t \to T_t^{(n)}$ denote the transition semigroup for the n-point motions on the flow F on $(\Omega, \mathscr{H}, \mathbb{P}^0)$. The measures μ_0 and π^0 are excessive for $t \to T_t^{(1)}$, the product measure of $\mu_0 + \pi^0$ with itself being excessive for $t \to T_t^{(2)}$ and the triple product of $\mu_0 + \pi^0$ with itself being excessive for $t \to T_t^{(3)}$.

CONDITION G. In the sense defined above, the mappings $\theta \to b^{\theta}$, $\theta \to p^{\theta}$, $\theta \to \ln p^{\theta}$, $\theta \to r^{\theta}$, where $r^{\theta} = k^{\theta}/k^{0}$, and $\theta \to \ln r^{\theta}$ are thrice continuously differentiable for every theta. There exist a constant K' and a pair of controlling functions b'' and b''' from E into the respective space of block matrices $E \otimes (\mathbb{R}^{p} \otimes \mathbb{R}^{p})$ and the block matrices $E \otimes (\mathbb{R}^{p} \otimes \mathbb{R}^{p} \otimes \mathbb{R}^{p})$ such that

$$\left\| \sum_{k,l=1}^{n} \sum_{i,j=1}^{d} \left[D^2_{\theta^2} b^{\theta}(x_k) \right]_i c_{ij}(x_k, x_l) \left[D^2_{\theta^2} b^{\theta}(x_l) \right]_j \right\| \\ \le K' \left\| \sum_{k,l=1}^{n} \sum_{i,j=1}^{d} \left[b''(x_k) \right]_i c_{ij}(x_k, x_l) \left[b''(x_l) \right]_j \right\|$$

for every θ in Θ and such that

$$\left|\left|\left| \sum_{k,l=1}^{n} \sum_{i,j=1}^{d} \left[D^3_{\theta_k^3} b^{\theta_k}(x_k) \right]_i c_{ij}(x_k, x_l) \left[D^3_{\theta_l^3} b^{\theta_l}(x_l) \right]_j \right|\right|\right| \\ \le K' \left|\left|\left| \sum_{k,l=1}^{n} \sum_{i,j=1}^{d} \left[b'''(x_k) \right]_i c_{ij}(x_k, x_l) \left[b'''(x_l) \right]_j \right|\right|\right|$$

for every $x_1, \dots, x_n$ in E and for every $\theta_1, \dots, \theta_n$ in Θ, where $|||\cdot|||$ is a suitable norm on three-dimensional matrices. There is a pair of controlling functions p'' and p''' such that

$$\left\| D^2_{\theta^2} p^{\theta}(x) \right\| \le K' \left\| p''(x) \right\| \quad \text{and} \quad \left|\left|\left| D^3_{\theta^3} p^{\theta}(x) \right|\right|\right| \le K' \left|\left|\left| p'''(x) \right|\right|\right|$$

for every x in E and θ in Θ. This last inequality prevails for the derivatives of the remaining parameters, but with corresponding controlling functions (lp'', lp'''), (r'', r''') and (lr'', lr'''). Here the symbol l invokes memory for the control functions on the logarithms of the corresponding parameters, so that lr''' controls $D^3_{\theta^3} \ln r^{\theta}$ for every theta and so on. This condition is essentially a pointwise-Lipschitz condition on two derivatives of the system parameters.

CONDITION H. The pointwise norm of the mappings p'', lp'', p''' and lp''' are square integrable with respect to π^0. The pointwise norm of the mappings r'', lr'', r''' and lr''' are square integrable with respect to $k^0\mu_0$ and $k^0\pi^0$. For z in E, let $\rho(z)$ denote the minimum eigenvalue of the nonsingular matrix $\gamma^{\mathsf{T}}(z)\gamma(z)$. If f denotes the mapping $(x, y, z) \to \sum_{i,j}^{d}[b''(x)]_i c_{ij}(x, y)[b''(y)]_j \rho^{-2}(z)$, then the mapping $x \to \|f(x, x, x)\|$ is integrable with respect to μ_0 and π^0, the mappings $(x, y) \to \|f(x, x, y)\|$, $(x, y) \to \|f(x, y, x)\|$ and $(x, y) \to \|f(x, y, y)\|$ are integrable with respect to $(\mu_0 + \pi^0) \times (\mu_0 + \pi^0)$. The mapping $(x, y, z) \to \|f(x, y, z)\|$ is integrable with respect to the threefold product measure of $\mu_0 + \pi^0$ with itself. Similarly, if f denotes the mapping

$$(x, y, z) \to \sum_{i,j=1}^{d} [b'''(x)]_i c_{ij}(x, y)[b'''(y)]_j \rho^{-2}(z),$$

then the mapping $x \to |||f(x, x, x)|||$ is integrable with respect to μ_0 and π^0, the mappings $(x, y) \to |||f(x, x, y)|||$, $(x, y) \to |||f(x, y, x)|||$ and $(x, y) \to |||f(x, y, y)|||$ are integrable with respect to $(\mu_0 + \pi^0) \times (\mu_0 + \pi^0)$. The mapping $(x, y, z) \to |||f(x, y, z)|||$ is integrable with respect to the threefold product measure of $\mu_0 + \pi^0$ with itself.

This completes our additional conditions on the system parameters. We show next that they are sufficient for the existence of observed information and its explicit representation as a stochastic integral on the particle process.

Information. The information process here refers to minus the second derivative of the log-likelihood process. In particular, if $\theta \to L_t^\theta$ is twice continuously differentiable for every t in $\mathbb{R}_+$, then the information process refers to the process $t \to -D_{\theta^2}^2 L_t^\theta$ for every theta. The observed information over our statistical model is then the mapping $\theta \to -D_{\theta^2}^2 L_T^\theta$.

Under the regularity conditions above, the next proposition shows that there exists an information process, which is calculated explicitly in the proof. The proof relies on arguments in Phelan (1996). We also appeal to the work of Métivier (1982) on interchanging differentiation and stochastic integration for semimartingales.

14. PROPOSITION. *We suppose that we have conditions A through H. For θ in Θ, there is an information process $t \to -D_{\theta^2}^2 L_t^\theta$ having an integral representation with respect to the particle process.*

PROOF. The stochastic basis is $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^0)$. For θ in Θ, Proposition 13 implies that the score process $t \to S_t^\theta$,

$$\begin{aligned} S_t^\theta = {} & \left(D_\theta H^\theta\right) \cdot X_t^c + \left(D_\theta \ln y^\theta\right) * (\mu - \nu^0)_t \\ & + \left(D_\theta \ln y^\theta - D_\theta J^\theta\right) * \nu_t^0 - \int_0^t ds\, q\left(M_{s-}; h_s^\theta, D_\theta h_s^\theta\right), \end{aligned}$$

is a well-defined semimartingale. To get the information process, we work on differentiating it.

Condition G implies that the coefficients in the score are differentiable with respect to theta. In particular, for θ in Θ, there is a mapping $\eta \to D_{\theta^2}^2 \ln y^\theta(\eta)$ on M_{sd} satisfying the equation

$$\begin{aligned} D_{\theta^2}^2 \ln y^\theta(\eta) = {} & \left(D_{\theta^2}^2 \ln p^\theta\right) \circ \delta^{-1}(\eta) 1_{\{+1\}}(\eta 1) \\ & + \left(D_{\theta^2}^2 \ln r^\theta\right) \circ \delta^{-1}(-\eta) 1_{\{-1\}}(\eta 1), \end{aligned}$$

where r^θ is again the function k^θ/k^0. Similarly, there is a mapping $\eta \to D^2_{\theta^2} y^\theta(\eta)$ and so a mapping $\eta \to D^2_{\theta^2} J^\theta(\eta)$. Next, let $(t, x) \to f''^\theta(M_{t-}; x)$ denote the predictable mapping satisfying the equation

$$f''^\theta(M_{t-}; x) = \sum_{i=1}^d \left[\gamma(x) \left[\int_E M_{t-}(du) \gamma^\mathsf{T}(u) \gamma(u) \right]^{-2} \int_E M_{t-}(dz) \gamma^\mathsf{T}(z) \right]_i \left[D^2_{\theta^2} b^\theta(z) \right]_i$$

if $M_{t-}(E) > 0$ and $f''^\theta = 0$ otherwise. In this case, let $D^2_{\theta^2} h^\theta_t$ denote the mapping $x \to f''^\theta(M_{t-}; x)$ on E for every t in $\mathbb{R}_+$. As a process of functionals on distributions, we then identify the process $t \to D^2_{\theta^2} H^\theta_t$ with the predictable process $t \to D^2_{\theta^2} h^\theta_t$ as a candidate for stochastic integration relative to the continuous-martingale part of the particle process.

Our aim is to propose a formula for the information process and then verify it. To do so, we first require the X^c-integrability of $t \to D^2_{\theta^2} H^\theta_t$, the $(\mu - \nu^0)$-integrability of $D^2_{\theta^2} \ln y^\theta - D^2_{\theta^2} J^\theta$ and two other such integrability requirements. Since each of the differentiated coefficients yields a matrix of order $p \times p$, we are referring to elementwise stochastic integration. We divide our argument in two: one part on integrability; the other on information itself.

(i) *Integrability*. For θ in Θ, (6) and Condition G imply that $\|D^2_{\theta^2} \ln y^\theta - D^2_{\theta^2} J^\theta\| * \nu^0_t$ satisfies the inequality

$$\|D^2_\theta \ln y^\theta - D^2_{\theta^2} J^\theta\|^2 * \nu^0_t \le K' t \pi^0\left(\|lp''\|^2 + \|p''\|^2\right) + K' \int_0^t ds\, M_s\left(k^0\left(\|lr''\|^2 + \|r''\|^2\right)\right)$$

for every t in $\mathbb{R}_+$. Therefore, with reference to details in Phelan (1996), Conditions F and H imply the P^0-integrability of $\|D^2_{\theta^2} \ln y^\theta - D^2_{\theta^2} J^\theta\| * \nu^0_t$ and so the almost sure ν^0-integrability of $D^2_{\theta^2} \ln y^\theta - D^2_{\theta^2} J^\theta$ on $[0, t] \times M_{sd}$ for every t in $\mathbb{R}_+$.

Next, we verify the X^c-integrability of $t \to D^2_{\theta^2} H^\theta_t$ and the $(\mu - \nu^0)$-integrability of $D^2_{\theta^2} \ln y^\theta$. To this end, we introduce the following notation: $M^{\otimes 2}_{s-}$ denotes the product measure of M_{s-} with itself and

$$q\left(M_{s-}\ D^2_{\theta^2} h^\theta_s, D^2_{\theta^2} h^\theta_s\right) \equiv \int_{E \times E} M^{\otimes 2}_{s-}(dx, dy) \sum_{i,j=1}^d \left[D^2_{\theta^2} h^\theta_s(x) \right]_i c_{ij}(x, y) \left[D^2_{\theta^2} h^\theta_s(y) \right]_j,$$

recalling that $D^2_{\theta^2} h^\theta_s(x)$ is a d-block matrix of $p \times p$ matrices. For the desired integrability, it suffices to show that the process $t \to \xi^\theta_t$,

$$\xi^\theta_t = \int_0^t ds \left\| q\left(M_{s-}; D^2_{\theta^2} h^\theta_s, D^2_{\theta^2} h^\theta_s\right) \right\| + \int_{[0,t] \times M_{sd}} \nu^0(ds, d\eta) \left\| D^2_{\theta^2} \ln y^\theta(\eta) \right\|^2,$$

is integrable with respect to P^0 for every t in $\mathbb{R}_+$.

First, we simplify and bound the integrand in the first term of ξ^θ_t. In particular, the calculation of Proposition 24 in Phelan (1996), the theory of

Moore–Penrose pseudoinverses and the spectral theory of symmetric matrices yields the inequality

$$\begin{aligned}\|q(M_{s-}; D^2_{\theta^2} h^\theta_3, D^2_{\theta^2} h^\theta_s)\| \\ \leq \left\| \int M^{\otimes 3}_{s-}(dx, dy, dz) \sum_{i,j=1}^{d} \left[D^2_{\theta^2} b^\theta(x)\right]_i c_{ij}(x, y)\left[D^2_{\theta^2} b^\theta(y)\right]_j \rho^{-2}(z)\right\|\end{aligned}$$

for the first term of ξ^θ_t, where $M^{\otimes 3}_{s-}$ denotes the threefold product measure of M_{s-} with itself and where ρ appears at Condition H. Now Condition G implies the inequality

$$\begin{aligned}E^0 \xi^\theta_t \leq K' \int_0^t ds\, E^0 \left\| \int M^{\otimes 3}_{s-}(dx, dy, dz) \sum_{i,j=1}^{d} [b''(x)]_i c_{ij}(x, y)[b''(y)]_j \rho^{-2}(z)\right\| \\ + K' t \pi^0 \|lp''\|^2 + K' \int_0^t ds\, E^0 M_s(k^0 \|lr''\|^2)\end{aligned}$$

for every t in $\mathbb{R}_+$. Finally, with reference to the same calculation of Proposition 24 in Phelan (1996), we apply Conditions F and H to the right-hand side of this last inequality to yield the desired result for ξ^θ_t for every θ in Θ and every t in $\mathbb{R}_+$.

To complete this part of our proof, we validate two other terms for integrability. To this end and for later use, we impose the following notation as natural to our problem:

$$\begin{aligned}q(M_{s-}; h^\theta_s, D^2_{\theta^2} h^\theta_s) \\ \equiv \int_{E\times E} M_{s-}(dx) M_{s-}(dy) \sum_{i,j=1}^{d} \left[h^\theta_s(x)\right]_i c_{ij}(x, y)\left[D^2_{\theta^2} h^\theta_s(y)\right]_j\end{aligned}$$

and

$$q(M_{s-}; D_\theta h^\theta_s, D_\theta h^\theta_s) \equiv \int_{E\times E} M_{s-}(dx) M_{s-}(dy) \left(D_\theta h^\theta_s(x)\right)^{\mathsf{T}} c(x, y) D_\theta h^\theta_s(y),$$

where in the second equivalence we note that $D_\theta h^\theta_s$ is a $d \times p$ matrix (random) function on E. These equivalences properly yield random matrices of order $p \times p$, yielding the mixed-partial derivatives on Θ. With these definitions, the integrability result above and that Θ is a bounded set imply the inequalities

$$E^0 \int_0^t ds\, q(M_{s-}; h^\theta_s, D^2_{\theta^2} h^\theta_s) < \infty \quad \text{and} \quad E^0 \int_0^t ds\, q(M_{s-}; D_\theta h^\theta_s, D^\theta h^\theta_s) < \infty$$

for every θ in Θ and t in $\mathbb{R}_+$.

(ii) *Information*. The integrability conditions above now imply that the process $t \to I^\theta_t$,

$$\begin{aligned}I^\theta_t = -\left(D^2_{\theta^2} H^\theta\right) \cdot X^c_t - \left(D^2_{\theta^2} \ln y^\theta\right) * (\mu - \nu^0)_t - \left(D^2_{\theta^2} \ln y^\theta - D^2_{\theta^2} J^\theta\right) * \nu^0_t \\ + \int_0^t ds\, q(M_{s-}; D_\theta h^\theta_s, D_\theta h^\theta_s) + \int_0^t ds\, q(M_{s-}; h^\theta_s, D^2_{\theta^2} h^\theta_s),\end{aligned}$$

is a well-defined semimartingale for every theta. We show that this process is the information process of interest here.

For each t in $\mathbb{R}_+$, the I_t^θ come of formally interchanging differentiation with respect to theta and stochastic integration in the formula for $-S_t^\theta$. To show that this gives the information, Proposition 13 implies that it suffices to show that $\theta \to -S_t^\theta$ is almost surely continuously differentiable, having derivative I_t^θ. Our method of proof follows that of Section 34 in Métivier (1982).

In particular, we suppose that the dimension p of Θ is 1. For θ in Θ and u in $\mathbb{R}$, the problem here is to show that there is an increasing sequence $n \to T_n$ of stopping times and a sequence $n \to c_n$ of constants such that T_n increases almost surely to infinity with n and such that we have the inequality

$$E^0 \sup_{t<T_n} \int_\Theta d\theta \left\| \frac{1}{u}\left(S_t^{\theta+u} - S_t^\theta + uI_t^\theta\right) \right\|^2 \le c_n u^2$$

for every n. Since Condition G allows us to substitute third-order Taylor expansions for second-order ones where necessary, the calculation of Proposition 25 in Phelan (1996) readily completes the proof. □

COROLLARY. *For θ in Θ, there exists an information process $t \to I_t^\theta$ satisfying the equation*

$$I_t^\theta = -\left(D_{\theta^2}^2 H^\theta\right) \cdot X_t^c - \left(D_{\theta^2}^2 \ln y^\theta\right) * (\mu - \nu^0)_t - \left(D_{\theta^2}^2 \ln y^\theta - D_{\theta^2}^2 J^\theta\right) * \nu_t^0$$
$$+ \int_0^t ds\, q\left(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\right) + \int_0^t ds\, q\left(M_{s-}; h_s^\theta, D_{\theta^2}^2 h_s^\theta\right).$$

The observed information over our statistical model is $\theta \to I_T^\theta$.

In anticipation of the next section, we close this section with a reexpression of the score and of the information. For θ in Θ, we suppose that the stochastic basis is now $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^\theta)$. In this case, the multivariate point process μ has compensator ν^θ obtained by substituting π^θ and k^θ into (6). The process $t \to X_t^c$ has drift so that the process $t \to X_t^{c\theta}$,

$$X_t^{c\theta}\phi = X_t^c\phi - \int_0^t ds \int_E M_s(dx)\, b^{\theta\top}(x) \nabla\phi(x), \qquad \phi \in \mathscr{D}, \tag{15}$$

now defines the continuous-martingale part of the particle process. This change of basis is locally an absolutely continuous change of measure for the particle process. The next proposition lists its effects on the martingale dynamics of the score and of information.

16. PROPOSITION. *For θ in Θ, the score process $t \to S_t^\theta$ is a locally square-integrable local martingale on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^\theta)$ satisfying the stochastic integral equation*

$$S_t^\theta = D_\theta H^\theta \cdot X_t^{c\theta} + D_\theta \ln y^\theta * (\mu - \nu^\theta)_t, \tag{17}$$

following some careful reexpression of the equation in Proposition 13. *Its quadratic characteristic* $t \to \langle S^\theta \rangle_t$ *satisfies the equation*

$$\langle S^\theta \rangle_t = \int_0^t ds\, q\big(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\big) + \int_0^t \int_{M_{sd}} \nu^\theta(ds, d\eta) D_\theta^\top \ln y^\theta(\eta) D_\theta \ln y^\theta(\eta)$$

for every t *in* $\mathbb{R}_+$. *Its quadratic variation* $t \to [S^\theta]_t$ *satisfies the equation*

$$[S^\theta]_t = \int_0^t ds\, q\big(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\big) + \big(D_\theta^\top \ln y^\theta D_\theta \ln y^\theta\big) * \mu_t$$

for every t *in* $\mathbb{R}_+$. *In addition, the information process* $t \to I_t^\theta$,

$$\begin{aligned} I_t^\theta = {} & -\big(D_{\theta^2}^2 H^\theta\big) \cdot X_t^{\theta c} - \big(\big(D_{\theta^2}^2 y^\theta\big)/y^\theta\big) * (\mu - \nu^\theta)_t \\ & + \big(D_\theta^\top \ln y^\theta D_\theta \ln y^\theta\big) * \mu_t + \int_0^t ds\, q\big(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\big), \end{aligned}$$

is a well defined semimartingale such that the process $t \to [S^\theta]_t - I_t^\theta$ *is itself a locally square-integrable local martingale.*

This proposition is a consequence of Proposition 12 and Girsanov's theorem for semimartingales [Jacod and Shiryaev (1987), Theorem III.3.24]. A proof involves a modest amount of stochastic calculus. The proposition introduces the three informational quantities of interest.

In the nomenclature of stochastic inference in Barndorff-Nielsen and Sørensen (1994), the quadratic variation is the incremental observed information, the quadratic characteristic of the incremental expected information and the information process of the observed information. We may simply call them the nonanticipative, predictable and observed information. That the score is locally square integrable and that the difference $[S^\theta] - I^\theta$ is a local martingale may be used to argue that they have the same expectation—Fisher's information.

5. Asymptotic estimation. We devote this section to maximum-likelihood estimation from birth and death on a flow. The aim is for a solution to the estimation equation

$$S_T^\theta = 0$$

for some theta in the parameter set. The problem here is to study the existence and distribution theory of a solution to the estimation equation as the length T of the chronicle increases to infinity.

The resolution of this problem is largely driven off of the asymptotic properties of information. So we treat asymptotic information first, but refer to Phelan (1995) for a gentler treatment of asymptotic information in the context of an equilibrium particle process. We then devote the latter part of this section to the existence, consistency and limiting law of a maximum-likelihood estimator of the system parameter theta.

Asymptotic information. We begin with three more regularity conditions. For differential operators, the notation and conventions here are those of the previous section. Similarly, for any Radon measure ζ and positive integer k, we let $\zeta^{\otimes k}$ denote the k-fold product measure of ζ with itself. As in Section 4, the quantity $\rho(z)$ denotes the value of the smallest eigenvalue of the symmetric nonsingular matrix $\gamma^{\mathsf{T}}(z)\gamma(z)$ for every z in E. Here the quantity $\sigma(z)$ denotes the value of the largest eigenvalue of $\gamma^{\mathsf{T}}(z)\gamma(z)$ for every z in E.

CONDITION I. For each θ in Θ, there exist the quasiinformational quantities $I(p^\theta)$ and $I(k^\theta)$ satisfying the equation

$$I(p^\theta) = \int_E \pi^\theta(dx) D_\theta^{\mathsf{T}} p^\theta(x) D_\theta p^\theta(x) \big(p^\theta(x)\big)^{-2}$$

and, letting g^θ denote the mapping $x \to D_\theta^{\mathsf{T}} k^\theta(x) D_\theta k^\theta(x)/k^\theta(x)$,

$$I(k^\theta) = \int_E \pi^\theta(dx)\big(g^\theta(x)\big)^2 + \int_{E\times E} \pi^\theta(dx)\pi^\theta(dy) g^\theta(x) g^\theta(y).$$

Also, let $f^{*\theta}$ denote the mapping $(x, y, z) \to D_\theta^{\mathsf{T}} b^\theta(x) c(x, y) D_\theta b^\theta(y)\rho^{-2}(z)$ on $E \times E \times E$. Let $f^{*\theta} \otimes f^{*\theta}$ denote the tensor product $(x, y, z, x', y', z') \to f^{*\theta}(x, y, z) f^{*\theta}(x', y', z')$ of $f^{*\theta}$ with itself on $E^3 \times E^3$. Label the coordinates x, y, z, x', y', z' with the labels $1, 2, \dots, 6$, respectively. For each $j = 1, 2, \dots, 6$, let V_j^6 denote a generic partition of the set $\{1, 2, \dots, 6\}$ into j nonempty subset(s). An example of V_2^6 is $\{\{1, 2\}, \{3, 4, 5, 6\}\}$. Then let $\delta_{V_j^6}$ denote the mapping on functions on E^6 into functions on E^j obtained by enforcing equality among the arguments of the original function among the coordinates in the subsets of V_j^6. Using the example above, $\delta_{V_2^6} f^{*\theta}$ is the mapping $(x, y) \to f^{*\theta}(x, x, y, y, y, y)$ on E^2. Using this notation, there exists a quasiinformational quantity $I^*(b^\theta)$ satisfying the equation

$$I^*(b^\theta) = \sum_{j=1}^{6} \sum_{V_j^6} \pi^{\theta\otimes j}\Big(\delta_{V_j^6}\big(f^{*\theta} \otimes f^{*\theta}\big)\Big).$$

Next, for any positive integer m, we introduce the partitions V_j^m of the labels $\{1, \dots, m\}$ into j nonempty subsets and the operators $\delta_{V_j^m}$ on functions on E^m into functions on E^j. In this case, let f_*^θ denote the mapping $(x, y) \to D_\theta^{\mathsf{T}} b^\theta(x) c(x, y) D_\theta b^\theta(y)$ on E^2. For each $m = 0, 1, 2, \dots,$ let $(\sigma^2)^{\otimes m}$ denote the m-fold tensor product of σ^2 with itself; of course $(\sigma^2)^{\otimes 0} \equiv 1$. Using this notation, there exists a positive-definite matrix $B^\theta(1/b)$ satisfying the equation

$$B^{\theta}\left(\frac{1}{b}\right) = \sum_{m=0}^{2n+1} (-1)^m \sum_{j=1}^{2m+2} \left(\frac{1}{b}\right)^{2+j} \times \sum_{V_j^{2m+2}} \pi^{\otimes j} \delta_{V_j^{2m+2}}\left(f_*^{\theta} \otimes (\sigma^2 \otimes \sigma^2)^{\otimes m}\right) + I(p^{\theta}) + \frac{1}{b} I(k^{\theta})$$

for some, possibly infinite, n, where b is the bound of the next Condition J.

CONDITION J. There exist constants b and a such that $b \geq k^{\theta} \geq a > 0$ for every θ in Θ.

CONDITION K. For each θ in Θ and on the probability space $(\Omega, \mathscr{H}, \mathbb{P}^{\theta})$, the Brownian flow $F = (F_{st})$, $0 \leq s \leq t \leq \infty$, preserves π^{θ} on E. For every bounded measurable function having compact support in E, we then have the equation

$$(\pi^{\theta} F_{0t}^{-1}) g = \int_E \pi^{\theta}(dx) g^{\theta}(F_{0t} x) = \int_E \pi^{\theta}(dx) g^{\theta}(x) = \pi^{\theta} g$$

almost surely for every t in $\mathbb{R}_+$.

This completes the conditions. Condition J supplants Condition D, requiring additionally a lower bound on the killing functions. Kunita (1990) treats measure-preserving Brownian flows in his Section 4.3. Because of his Theorem 4.3.2, Condition K supplants Condition F. Now for every θ in Θ, on the probability space $(\Omega, \mathscr{H}, \mathbb{P}^{\theta})$, the k-fold product measure $\pi^{\theta \otimes k}$ is an invariant measure of the k-point motion on the flow.

For each θ in Θ, we recall the gathering of informational statistics at Proposition 16. On the stochastic basis $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^{\theta})$, the score is a locally square-integrable martingale. Its quadratic variation $t \to [S^{\theta}]_t$ is the incremental observed information; its quadratic characteristic $t \to \langle S^{\theta} \rangle_t$ is the incremental expected information. The process $t \to [S^{\theta}]_t - I_t^{\theta}$ is a zero-mean martingale. For each t in $\mathbb{R}_+$, the quantity $E^{\theta} S_t^{\theta \top} S_t^{\theta}$ is the expected information. This next proposition shows that the temporal averages of these statistics converge to a positive-definite matrix.

18. PROPOSITION. *We suppose that we have the regularity conditions A through K. For each θ in Θ, there is a positive-definite matrix Σ^{θ} of rank p such that we have the limits*

$$\lim_{t \to \infty} t^{-1} \langle S^{\theta} \rangle_t = \lim_{t \to \infty} t^{-1} [S^{\theta}]_t = \lim_{t \to \infty} t^{-1} I_t^{\theta} = \lim_{t \to \infty} t^{-1} E^{\theta} S_t^{\theta \top} S_t^{\theta} = \Sigma^{\theta}$$

almost surely on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^{\theta})$.

PROOF. We fix θ in Θ. The stochastic basis is $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^\theta)$. The atoms of M_0 constitute the initial particles on the flow. By virtue of the lower bound in Condition J, their contribution to the asymptotic calculation below vanishes almost surely in finite time. We suppose therefore that $M_0 = 0$, so that the renewal argument below is conditional on this initial condition.

By virtue of Proposition 16, the score process $t \to S_t^\theta$ is a locally square-integrable local martingale. By virtue of Theorem 23.4(3°) in Métivier (1982), we have the equalities

$$E^\theta S_T^{\theta\mathsf{T}} S_T^\theta = E^\theta \langle S^\theta \rangle_T = E^\theta [S^\theta]_T = E^\theta I_T^\theta$$

for every stopping time T. This consequence of the martingale property sharply reduces our computational burden below.

In particular, that π^θ is a finite measure and the lower bound in Condition J imply that there is a sequence $k \to T_k$ of regeneration times for the particle process; compare with the argument in the proof of Lemma 4.10 in Çinlar and Kao (1992b) and Definition 9.2.18 in Çinlar (1975). The distribution of T_1 is not arithmetic and $E^\theta T_1$ is finite. If, for example, R is the stopping time $\inf\{t\colon M_t(E) > 0\}$, then T_1 is the stopping time $\inf\{t > R\colon M_t(E) = 0\}$. We put $T = T_1$ for future purposes.

In light of the martingale property, we claim that it suffices to show that $E^\theta \langle S^\theta \rangle_T$ exists and is positive definite. The desired limits are then a consequence of the regenerative property in Çinlar [(1975), 2.18b of Definition 9.2.18] and the strong law of large numbers for regenerative phenomena.

Let $t \to \zeta_t^\theta$ denote the $p \times p$ matrix-valued process satisfying the equation

$$\zeta_t^\theta = q\big(M_{t-}; D_\theta h_t^\theta, D_\theta h_t^\theta\big) + I(p^\theta) + \int_E \big(k^\theta M_{t-}\big)(dx)\, D_\theta^\mathsf{T} k^\theta(x) D_\theta k^\theta(x) \big(k^\theta(x)\big)^{-2},$$

where $q(M_{t-}; D_\theta h_t^\theta, D_\theta h_t^\theta)$ appears first near the close of part (i) in the proof of Proposition 14 and $I(p^\theta)$ appears in Condition I. Naturally, $\langle S^\theta \rangle_t$ satisfies the equation

$$\langle S^\theta \rangle_t = \int_0^t ds\, \zeta_s^\theta$$

for every t in $\mathbb{R}_+$.

We fix the vector v in $\mathbb{R}^p$ of unit length. For the stopping time T and the quadratic form $v^\mathsf{T} \langle S^\theta \rangle_T v$, we have the equation

$$\begin{aligned} &E^\theta v^\mathsf{T} \langle S^\theta \rangle_T v \\ &\quad = \int_0^\infty ds\, E^\theta v^\mathsf{T} \zeta_s^\theta v 1_{\{T \ge s\}} \\ &\quad = \int_0^\infty ds\, E^\theta v^\mathsf{T} q\big(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\big) v 1_{\{T \ge s\}} + v^\mathsf{T} I(p^\theta) v E^\theta T \\ &\qquad + \int_0^\infty ds\, E^\theta v^\mathsf{T} \int_E \big(k^\theta M_{s-}\big)(dx)\, D_\theta^\mathsf{T} k^\theta(x) D_\theta k^\theta(x) \big(k^\theta(x)\big)^{-2} v 1_{\{T \ge s\}}. \end{aligned}$$

To show that this expectation is finite, therefore, it suffices to do so for the first and third term on the right-hand side of the second equality.

Thus, we exhibit two finite constants B and K, for example, such that we have the inequalities

$$E^{\theta}\left[v^{\mathsf{T}}q\left(M_{s-};D_{\theta}h_{s}^{\theta},D_{\theta}h_{s}^{\theta}\right)v\right]^{2}\leq B$$

and

$$E^{\theta}\left[v^{\mathsf{T}}\int\left(k^{\theta}M_{s-}\right)(dx)D_{\theta}^{\mathsf{T}}k^{\theta}(x)D_{\theta}k^{\theta}(x)\left(k^{\theta}(x)\right)^{-2}v\right]^{2}\leq K$$

for every s in $\mathbb{R}_{+}$. We then show that the mapping $s\to P^{\theta}(T\geq s)^{1/2}$ is integrable on $\mathbb{R}_{+}$. That the expectations above are finite as desired is then a consequence of Hölder's inequality.

In demonstrating the bounds above, we move freely between the canonical setting $(\mathbb{D}(M_{b}),\mathscr{G},P^{\theta})$ with its expectation E^{θ} and the probability space $(\Omega,\mathscr{H},\mathbb{P}^{\theta})$ of Section 2 with its expectation $\mathbb{E}^{\theta}$. In each space, we maintain the notation $t\to M_{t}$ for the particle process.

Let $\mathbb{P}_{F}^{\theta}$ denote conditional probability given the sigma algebra as generated by the flow. For each s in $\mathbb{R}_{+}$, Proposition 2.13 in Çinlar and Kao (1992b) implies that the law of M_{s} under $\mathbb{P}_{F}^{\theta}$ is that of a Poisson random measure having mean μ_{s} satisfying the equation

$$\mu_{s}f=\mathbb{E}_{F}^{\theta}M_{s}f=\int_{0}^{s}dr\int_{E}\pi^{\theta}(dx)f(F_{rs}x)\exp\left(-\int_{r}^{s}dw\,k^{\theta}(F_{rw}x)\right)$$

for every continuous function f having compact support in E, recalling that $M_{0}=0$ here. Subsequently we will repeatedly use this fact and essential results on moment measures of Poisson processes.

First, let g^{θ} denote the mapping on E as defined in Condition I. Let g_{v}^{θ} denote the mapping $x\to v^{\mathsf{T}}g^{\theta}(x)v$ and let $g_{v}^{\theta}\otimes g_{v}^{\theta}$ denote the tensor product $(x,y)\to g_{v}^{\theta}(x)g_{v}^{\theta}(y)$ on $E\times E$. Proposition 2.13 in Çinlar and Kao (1992b), Example 1.15 in Chapter 1 of Karr (1986) and Conditions J and K imply the following inequalities:

$$\begin{aligned}
&E^{\theta}\left[v^{\mathsf{T}}\int\left(k^{\theta}M_{s}\right)(dx)D_{\theta}^{\mathsf{T}}k^{\theta}(x)D_{\theta}k^{\theta}(x)\left(k^{\theta}(x)\right)^{-2}v\right]^{2}\\
&\quad=E^{\theta}M_{s}^{\otimes 2}g_{v}^{\theta}\otimes g_{v}^{\theta}\\
&\quad=\mathbb{E}^{\theta}\int_{0}^{s}dr\int\pi^{\theta}(dx)\left(g_{v}^{\theta}(F_{rs}x)\right)^{2}\exp\left(-\int_{r}^{s}dw\,k^{\theta}(F_{rw}x)\right)\\
&\qquad+\mathbb{E}^{\theta}\int_{0}^{s}dr\int_{0}^{s}du\int\pi^{\theta\otimes 2}(dx,dy)g_{v}^{\theta}\otimes g_{v}^{\theta}(F_{rs}x,F_{us}y)\\
&\qquad\qquad\times\exp\left(-\int_{r}^{s}dw\,k^{\theta}(F_{rw}x)-\int_{u}^{s}dz\,k^{\theta}(F_{uz}y)\right)\\
&\quad\leq a^{-1}v^{\mathsf{T}}I(k^{\theta})v+a^{-2}v^{\mathsf{T}}I^{2}(k^{\theta})v
\end{aligned}$$

for every s in $\mathbb{R}_+$, where $I(k^\theta)$ appears in Condition I. This exhibits one of the desired bounds.

Next, let $f^{*\theta}$ denote the mapping on $E \times E \times E$ as defined in Condition I. Let $f_v^{*\theta}$ denote the mapping $x \to v^{\mathsf{T}} f^{*\theta}(x, y, z)v$ and let $f_v^{*\theta} \otimes f_v^{*\theta}$ denote the tensor product on $E^3 \times E^3$. By virtue of the argument in the proof of Proposition 24, part 3, in Phelan (1996), we have the following inequality:

$$E^\theta\left[v^{\mathsf{T}} q\left(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\right)v\right]^2 \le E^\theta\left[M_{s-}^{\otimes 3} f_v^{*\theta}\right]^2 = E^\theta M_{s-}^{\otimes 6} f_v^{*\theta} \otimes f_v^{*\theta}.$$

By virtue of Proposition 2.13 in Çinlar and Kao (1992b) and Exercise 5.4.5 and Example 7.4(a) in Daley and Vere-Jones (1988), Conditions I, J and K imply the inequality

$$E^\theta\left[v^{\mathsf{T}} q\left(M_{s-}; D_\theta h_s^\theta, D_\theta h_s^\theta\right)v\right]^2 \le \sum_{j=1}^{6} a^{-j} \sum_{V_j} \pi^{\theta \otimes j}\left(\delta_{V_j}\left(f_v^{*\theta} \otimes f_v^{*\theta}\right)\right)$$

for every s in $\mathbb{R}_+$, using the notation of Condition I. This exhibits the second desired bound.

Therefore, it remains to show that $s \to P^\theta(T \ge s)^{1/2}$ is integrable on $\mathbb{R}_+$. We again appeal to the proof of Lemma 4.10 in Çinlar and Kao (1992b). In particular, on the stochastic basis $(\Omega, \mathscr{H}, \mathbf{H}, \mathbb{P}^\theta)$, let $t \to \overline{M}_t(\omega)$ denote the particle process of (1), but with k replaced by the constant a for every ω in Ω. Recalling that $\overline{M}_0 = M_0 = 0$, we introduce, as above, the stopping time $\overline{R} = \inf\{t\colon \overline{M}_t(E) > 0\}$ and the stopping time $\overline{T} = \inf\{t > \overline{R}\colon \overline{M}_t(E) = 0\}$. Since $\overline{T}$ exceeds or equals T almost surely on $(\Omega, \mathscr{H}, \mathbb{P}^\theta)$, it suffices to show that $s \to \mathbb{P}^\theta(\overline{T} \ge s)^{1/2}$ is integrable on $\mathbb{R}_+$.

We show that $\mathbb{P}^\theta(\overline{T} \ge s)$ decays exponentially fast for all sufficiently large s. That is, the process $t \to \overline{M}_t(E)$ is an $M/M/\infty$ queue having arrival rate $c = \pi^\theta(E)$ and expected service time $1/a$. The time $\overline{T}$ equals the sum of an exponential random variable with mean $1/c$ and the length of the first busy period of the queue. In the theory of coverage processes, the length of a busy period in an $M/M/\infty$ queue is the length of a clump in a Boolean model on the line with parameters c and $1/a$. We rescale this model to one with Poisson intensity cs and mean segment length $1/as$ and consider the probability of complete coverage of the unit interval $[0, 1]$. By virtue of Theorem 2.5 in Hall (1988), taking $t = 1$, $\lambda = cs$ and $\alpha = 1/as$ in his (2.21), we have the heavy-traffic limit $\mathbb{P}^\theta(\overline{T} \ge s) \approx \exp(-sc\exp(-c/a))$ for all sufficiently large s, implying the desired integrability.

We next show that the quantity $E^\theta v^{\mathsf{T}} \langle S^\theta \rangle_T v$ is strictly positive. Let $t \to g(M_t)$ denote the mapping $t \to v^{\mathsf{T}} \zeta_t^\theta$. The regenerative property implies that its expectation satisfies the renewal equation

$$E^\theta g(M_t) = E^\theta g(M_t) 1_{\{T \ge t\}} + \int_0^t P^\theta(T \in ds) E^\theta g(M_{t-s}).$$

By virtue of the demonstration above and Proposition 9.2.16(c) in Çinlar (1975), the leading term on the right-hand side above is bounded by a directly Riemann-integrable function. Therefore, by virtue of the quasi-left continuity of the particle process, Çinlar's Proposition 4.2.16(b) implies that the leading term itself is directly Riemann integrable. Since the distribution of T is not arithmetic, the key renewal theorem therefore implies the limit

$$\lim_{t\to\infty} E^{\theta} g(M_t) = \frac{1}{E^{\theta}T} E^{\theta} v^{\mathsf{T}} \langle S^{\theta}\rangle_T v.$$

However, for the leading term in $g(M_t) = v^{\mathsf{T}}\zeta_t^{\theta} v$, we have the inequality

$$v^{\mathsf{T}} q\big(M_{t-}; D_{\theta} h_t^{\theta}, D_{\theta} h_t^{\theta}\big) v \ge \frac{1}{1 + \big[M_{t-}\sigma^2\big]^2} v^{\mathsf{T}} \bigg[\int_{E\times E} M_{t-}^{\otimes 2}(dx, dy) f_*^{\theta}(x, y)\bigg] v,$$

where f_*^{θ} appears in Condition I. We therefore approximate $1/(1 + [M_{t-}\sigma^2]^2)$ from below with $2n + 1$ terms in its Taylor series and find that Proposition 2.13 in Çinlar and Kao (1992b), manipulations on the moment measures of the Poisson process and Conditions I, J and K imply the inequality

$$\lim_{t\to\infty} E^{\theta} g(M_t) \ge v^{\mathsf{T}} B^{\theta}\left(\frac{1}{b}\right) v > 0$$

for every v in $\mathbb{R}^p$ of unit length. Thus, $E^{\theta}\langle S^{\theta}\rangle_T$ is positive definite. This completes the proof. □

As seen in the proof, the work of Çinlar and Kao (1992b) inspired the renewal argument above. We notice that the regenerative property makes the limiting information deterministic, putting our case of stochastic inference among the ergodic ones. This fact simplifies the asymptotic distribution theory below.

Distribution theory. We show here that there exists asymptotically a consistent maximum-likelihood estimator. Its limiting law is Gaussian. We draw our treatment from the program for such inference in Barndorff-Nielsen and Sørensen (1994). Because of Proposition 18, we do not invoke the full generality of their program, but use its simpler form for ergodic inference.

For each θ in Θ and t in $\mathbb{R}_+$, let $D^{\theta}(t)$ denote the diagonal matrix containing the diagonal elements from the information matrix $E^{\theta}\langle S^{\theta}\rangle_t$. The random variable Z is Gaussian on $\mathbb{R}^p$ having zero mean and identity covariance matrix.

19. Proposition. *We suppose conditions A through K. For θ in Θ, the stochastic basis is $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^{\theta})$. For sufficiently large t, there exists a solution $\hat{\theta}_t$ to the estimation equation $S_t^{\vartheta} = 0$, $\vartheta \in \Theta$, such that we have the convergence*

$$\hat{\theta}_t \to \theta$$

in probability as t increases to infinity. Finally, we have the weak limit

$$\langle S^{\theta}\rangle_t^{1/2}\big(\hat{\theta}_t - \theta\big) \to Z$$

as t increases to infinity.

PROOF. We approach this problem from the general theory of filtered statistical models in Barndorff-Nielsen and Sørensen (1994), thereby extending their class of examples to a spatial stochastic process.

It suffices to verify conditions 2.1, 2.2, 2.3, 3.1, A.2 and A.3 in Barndorff-Nielsen and Sørensen (1994). Since Proposition 18 implies that the diagonal elements of $D^\theta(t)$ approach infinity with t at the same rate, the desired result then follows from Theorem A.1 and (3.44) in Barndorff-Nielsen and Sørensen (1994).

Fix θ in Θ and the stochastic basis $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^\theta)$. By virtue of Proposition 16, the score process $t \to S_t^\theta$ is a locally square-integrable local martingale, the information process $t \to I_t^\theta$ is a semimartingale of finite expectation and $t \to I_t^\theta - [S^\theta]_t$ is a local martingale. We thus have Conditions 2.1, 2.2 and 2.3 in Barndorff-Nielsen and Sørensen (1994). Moreover, by virtue of Proposition 18, we have a, c and d of Condition 3.1.

We turn to the infinitesimality condition 3.1.b in Barndorff-Nielsen and Sørensen (1994). In our case, however, we verify a different but still sufficient condition for infinitesimality. The goal is for a central limit theorem for the normalized score. For each t, the natural choice here is to embed $s \to S_s^\theta$, $0 \le s \le t$, into a square-integrable martingale $X^t = (X_u^t)$, $0 \le u \le 1$, such that $X_u^t = S_{ut}^\theta$ for every u so that X^t inherits its characteristics from the score process in a natural way. For example, let ν^{X^t} denote the compensator on the discontinuities of X^t. In light of our Proposition 18, to get a central limit theorem for the normalized score, we verify the Lindeberg–Feller condition

$$t^{-1}|x|^2 1_{\{|x| > \sqrt{t}\,\varepsilon\}} * \nu_1^{X^t} = t^{-1}|D_\theta \ln y^\theta|^2 1_{\{|D_\theta \ln y^\theta| > \sqrt{t}\,\varepsilon\}} * \nu_t^\theta \to_{P^\theta} 0$$

as t increases to infinity for every $\varepsilon > 0$ [cf. Jacod and Shiryaev (1987), VIII.3.23, page 435]. On the other hand, Proposition 18 implies

$$t^{-1}[X^t, X^t]_1 = t^{-1}[S^\theta]_t \to_{P^\theta} \Sigma^\theta \quad \text{and} \quad t^{-1}\langle X^t, X^t\rangle_1 = t^{-1}\langle S^\theta\rangle_t \to_{P^\theta} \Sigma^\theta,$$

namely, (ii) and (iii) of VIII.3.24 in Jacod and Shiryaev (1987) with $t = 1$. As in the proof of VIII.3.22 in Jacod and Shiryaev (1987), this implies the desired Lindeberg–Feller condition above.

We finally verify conditions A.2 and A.3 in Barndorff-Nielsen and Sørensen (1994). Proposition 18 and the regenerative property imply that there exists a positive-definite matrix B^θ such that $[D^\theta(t)]^{-1/2}\langle S^\theta\rangle_t[D^\theta(t)]^{-1/2} \to B^\theta$ in probability as t increases to infinity. So, we introduce the martingale $t \to G_t^\theta$,

$$G_t^\theta = [S^\theta]_t - I_t^\theta,$$

on $(\mathbb{D}(M_b), \mathscr{G}, \mathbf{G}, P^\theta)$. For strictly positive α, we introduce the set

$$B(\alpha, t) = \left\{\theta' : \left\| B^{\theta/2} D^\theta(t)^{1/2}(\theta' - \theta)\right\| \le \alpha\right\}$$

in Θ. The problem here is to show that two quantities

$$\sup_{\theta' \in B(\alpha, t)} \left\| D^\theta(t)^{-1/2} G_t^{\theta'} D^\theta(t)^{-1/2}\right\|$$

and

$$\sup_{\theta' \in B(\alpha, t)} \left\| D^{\theta}(t)^{-1/2} [S^{\theta'}]_t D^{\theta}(t)^{-1/2} - B^{\theta} \right\|$$

converge in probability to zero as t increases to infinity. Since Proposition 18 implies that we have the limits

$$\left\| D^{\theta}(t)^{-1/2} G_t^{\theta} D^{\theta}(t)^{-1/2} \right\| \to_{P^{\theta}} 0$$

and

$$\left\| D^{\theta}(t)^{-1/2} [S^{\theta}]_t D^{\theta}(t)^{-1/2} - B^{\theta} \right\| \to_{P^{\theta}} 0$$

as t increases to infinity, it suffices to estimate the distances between G_t^{θ} and $G_t^{\theta'}$ and between $[S^{\theta}]_t$ and $[S^{\theta'}]_t$ on $B(\alpha, t)$.

First, the quadratic variation $t \to [S^{\theta}]_t$ satisfies the equation

$$[S^{\theta}]_t = \int_0^t ds\, q\big(M_{s-}; D_{\theta} h_s^{\theta}, D_{\theta} h_s^{\theta}\big) + \int_0^t \int_{M_{sd}} \mu(ds, d\eta) D_{\theta}^{\top} \ln y^{\theta}(\eta) D_{\theta} \ln y^{\theta}(\eta)$$

using our convention for $q(M_{s-}; D_{\theta} h_s^{\theta}, D_{\theta} h_s^{\theta})$ from part (i) of the proof of Proposition 14. For any θ' in $B(\alpha, t)$, we have the equality

$$\begin{aligned} &q\big(M_{s-}; D_{\theta} h_s^{\theta}, D_{\theta} h_s^{\theta}\big) - q\big(M_{s-}; D_{\theta'} h_s^{\theta'}, D_{\theta'} h_s^{\theta'}\big) \\ &\quad = q\big(M_{s-}; D_{\theta} h_s^{\theta} + D_{\theta'} h_s^{\theta'}, D_{\theta} h_s^{\theta} - D_{\theta'} h_s^{\theta'}\big). \end{aligned}$$

Simplifying the right-hand side here, we find that Condition G and that the diameter $\|\Theta\|$ of Θ is finite imply the inequality

$$\begin{aligned} &\left\| q\big(M_{s-}; D_{\theta} h_s^{\theta} + D_{\theta'} h_s^{\theta'}, D_{\theta} h_s^{\theta} - D_{\theta'} h_s^{\theta'}\big) \right\| \\ &\quad \le |\theta - \theta'| K' \|\Theta\| \\ &\qquad \times \left\| \int M_{s-}^{\otimes 3}(dx, dy, dz) \sum_{i, j=1}^{d} [b''(x)]_i c_{ij}(x, y) [b''(y)]_j \rho^{-2}(z) \right\|. \end{aligned}$$

More directly, Condition G implies that $\vartheta \to D_{\vartheta} \ln y^{\vartheta}(\eta)$ satisfies a Lipschitz condition for every η in M_{sd}. Therefore, we have the inequality

$$\begin{aligned} &\left\| [S^{\theta}]_t - [S^{\theta'}]_t \right\| \\ &\quad \le |\theta - \theta'| K' \|\Theta\| \\ &\qquad \times \Bigg[\int_0^t ds \left\| \int M_{s-}^{\otimes 3}(dx, dy, dz) \sum_{i, j=1}^{d} [b''(x)]_i c_{ij}(x, y) [b''(y)]_j \rho^{-2}(z) \right\| \\ &\qquad\quad + \int_0^t \int_{M_{sd}} \mu(ds, d\eta) \big(\| p'' \circ \delta^{-1}(\eta) \| 1_{\{+1\}}(\eta(E)) \\ &\qquad\qquad + \| lr'' \circ \delta^{-1}(-\eta) \| 1_{\{-1\}}(\eta(E)) \big) \Bigg] \end{aligned}$$

for every θ' in $B(\alpha, t)$ and t in $\mathbb{R}_+$. Since the diameter of $B(\alpha, t)$ decreases to zero as t increases to infinity, this last inequality, Proposition 18, its regenerative argument and the integrability Condition H imply that the second quantity

$$\sup_{\theta' \in B(\alpha, t)} \left\| D^{\theta}(t)^{-1/2} [S^{\theta'}]_t D^{\theta}(t)^{-1/2} - B^{\theta} \right\|,$$

converges in probability as desired to zero as t increases to infinity.

Last, we have the inequality

$$\|G_t^{\theta} - G_t^{\theta'}\| \le \left\| [S^{\theta}]_t - [S^{\theta'}]_t \right\| + \|I_t^{\theta} - I_t^{\theta'}\|$$

for every θ' in $B(\alpha, t)$. By virtue of the argument above, therefore, it suffices to get an analogous Lipschitz-type estimate for the second term here. However, an analogous argument, this time using the thrice differentiability of the system parameters and the controlling functions b''', p''' and lr''', readily yields the desired estimate. In addition, again by virtue of Proposition 18, its regenerative argument and the integrability Condition H, the quantity of interest,

$$\sup_{\theta' \in B(\alpha, t)} \left\| D^{\theta}(t)^{-1/2} G_t^{\theta'} D^{\theta}(t)^{-1/2} \right\|,$$

converges in probability as desired to zero as t increases to infinity. This completes our proof. □

We chose $\langle S^{\theta} \rangle_t$ as normalizer in the pivot $\langle S^{\theta} \rangle_t^{1/2} (\hat{\theta}_t - \theta)$. Because of Proposition 18, there are several equivalent choices here. In a given parametrization, one of course chooses the simplest. Finally, here is a corollary that provides for the construction of confidence sets for theta.

COROLLARY. *Let χ denote a random variable with the chi-square distribution on p degrees of freedom. For sufficiently large t, there exists almost surely a maximum-likelihood estimator $\hat{\theta}_t$. In this case, let $\tilde{\theta}_t$ be a convex combination of θ and $\hat{\theta}_t$. We then have the weak limit*

$$\left(\hat{\theta}_t - \theta\right)^{\mathsf{T}} I_t^{\tilde{\theta}_t} \left(\hat{\theta}_t - \theta\right) \to \chi$$

as t increases to infinity. Proposition 19 *implies that the limit attains at $\tilde{\theta}_t = \hat{\theta}_t$.*

Acknowledgments. I gratefully acknowledge the constructive criticism of an anonymous referee. I especially thank Erhan Çinlar for introducing his work to me over hours of dark coffee at the Rotunda of Princeton University. Finally, many thanks to Dr. Hu.

REFERENCES

BARNDORFF-NIELSEN, O. E. and SØRENSEN, M. (1994). A review of some aspects of asymptotic likelihood theory for stochastic processes. *Internat. Statist. Rev.* **62** 133–178.

BASAWA, I. V., FEIGIN, P. D. and HEYDE, C. C. (1976). Asymptotic properties of maximum likelihood estimators for stochastic processes. *Sankhyā Ser. A* **38** 259–270.

ÇINLAR, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.
ÇINLAR, E. and KAO, J. S. (1992a). Particle systems on flows. In *Applied Stochastic Models and Data Analysis*. Wiley, New York.
ÇINLAR, E. and KAO, J. S. (1992b). Birth and death on a flow. In *Diffusion Processes and Related Problems in Analysis*. *Z. Stochastic Flows* (M. A. Pinsky and V. Wihstutz, eds.) 121–137. Birkhäuser, Boston.
DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York.
FEIGIN, P. D. (1976). Maximum likelihood estimation for continuous-time stochastic processes. *Adv. in Appl. Probab.* **8** 712–736.
GODAMBE, V. P. and HEYDE, C. C. (1987). Quasi likelihood and optimal estimation. *Internat. Statist. Rev.* **55** 231–244.
HALL, P. (1988). *Introduction to the Theory of Coverage Processes*. Wiley, New York.
ITÔ, K. (1984). *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*. SIAM, Philadelphia.
JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.
KARR, A. F. (1986). *Point Processes and Their Statistical Inference*. Dekker, New York.
KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations*. Cambridge Univ. Press.
MÉTIVIER, M. (1982). *Semimartingales: A Course on Stochastic Processes*. de Gruyter, New York.
PHELAN, M. J. (1995). Asymptotic information for parametric estimation from an equilibrium particle process. *Statist. Probab. Lett.* **24** 193–198.
PHELAN, M. J. (1996). Transient tracers and birth and death on flows: parametric estimation of rates of drift, injection, and decay. *J. Statist. Plann. Inference* **51** 19–40.
SØRENSEN, M. (1989). A note on the existence of a consistent maximum likelihood estimator for diffusions with jumps. In *Markov Processes and Control Theory* **54** 229–234. Akademie-Verlag, Berlin.
SØRENSEN, M. (1990). Likelihood methods for diffusions with jumps. In *Statistical Inference in Stochastic Processes* (N. U. Prabhu and I. V. Basawa, eds.) 67–106. Dekker, New York.

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104-6302
E-MAIL: phelan@stat.wharton.upenn.edu