branching diffusions with immigration: time-discrete observation, a reconstruction algorithm for particle trajectories, and estimation of the diffusion coefficient

Reinhard Höpfner, Universität Mainz
ongoing joint work with Matthias Hammer, TU Berlin,
and Tobias Berg, Universität Mainz

DYNSTOCH Rennes 08.–10.06.2016

our aims

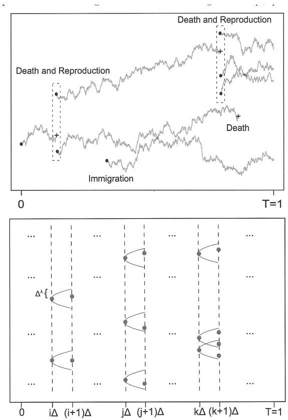
branching diffusion with immigration (BDI), ergodic setting,
time-continuous observation $\leftrightarrow$ time-discrete observation at small step size Δ:

in continuous time:

- independent diffusion paths
- position-dependent branching rate
- random displacement of offspring
- immigration at constant rate

in discrete time:

- pairs of successive configurations
- no information on particle identities
- seemingly 'identifiable' pairs – and others
- $\rightarrow$ reconstruction algorithm ??
- $\rightarrow$ regression schemes for estimation ??



BDI process, ergodicity, invariant measure

single particle space $(E, \mathcal{E})$: $E = \mathbb{R}^d$ with Borel σ-field

configuration space $(S, \mathcal{S})$: $S = \bigcup_{\ell \in \mathbb{N}_0} E^\ell$ with Borel σ-field, $E^0 = \{\delta\}$ void conf.

$x = (x_1, \ldots, x_\ell)$ elements of S, $\ell : S \to \mathbb{N}_0$ length of a configuration,
$x \in S$, $A \in \mathcal{E}$: $x(A) = \sum_{i=1}^{\ell(x)} \epsilon_{x_i}(A)$ number of particles visiting A,
$f : E \to \mathbb{R}$ a function, $x \in S$: $\overline{f}(x) = \sum_{i=1}^{\ell(x)} f(x_i)$ with convention $\overline{f}(\delta) := 0$

particles travel on independent diffusion paths $d\xi_s = b(\xi_s)ds + \sigma(\xi_s)dW_s$,
$a = \sigma\sigma^\top$, assume: drift $\mathcal{C}_b^1(E)$, diffusion coefficient $\mathcal{C}_b^2(E)$

branching at position-dependent rate $\kappa(\cdot) \in \mathcal{C}_b(E)$

when branching happens in position $y \in E$: number k of descendants and locations $y + v_1, \ldots, y + v_k$ for offspring selected by Markov kernel

$$K_1(y, dk)K_2((y, k), dv_1, \ldots, dv_k) = p_k(y) \prod_{j=1}^{k} Q^{\mathrm{r}}(dv_j) ,$$

finite reproduction means: $\rho(\cdot) = \sum_{k \geq 1} k p_k(\cdot)$ in $\mathcal{C}_b(E)$

immigration: PPP on $(0, \infty) \times E$ with intensity $c\,ds\, Q^{\mathrm{i}}(dy)$

with these ingredients:
construct the BDI process as S-valued strong Markov process $(\eta_t)_{t\geq 0}$

- infinite lifetime, no accumulation of jumps in finite time, jumps (branching or immigration) arriving at rate $(c+\overline{\kappa}(\eta_s))ds$
- sequence $T_n \uparrow \infty$ of stopping times such that on $[[T_n, T_{n+1}[[$, ℓ-particle configurations travel on diffusion paths

subprocesses $(\eta^{\mathfrak{r}}_{s+h})_{h\geq 0}$ of all direct descendants of one ancestor at time s are branching diffusions without immigration, occupation time kernel

$$H(y,g) = E_y\left(\int_0^\infty \overline{g}(\eta^{\mathfrak{r}}_h)\, dh\right) \ , \quad y\in E, g\in \mathcal{E}^+$$

for which there exists a jump diffusion $\widetilde{\xi}$ such that

$$H(y,g) = E_y\left(\int_0^\infty g(\widetilde{\xi}_t)\, e^{-\int_0^t [\kappa(1-\rho)](\widetilde{\xi}_s)\, ds}\, dt\right)$$

specified through its generator (writing $\mathcal{L}$ for the generator of ξ)

$$\widetilde{\mathcal{L}}g(y) = \mathcal{L}g(y) + [\kappa\rho](y)\int_E Q^{\mathfrak{r}}(dv)[g(y+v)-g(y)]$$

proposition 1: the BDI process $(\eta_t)_{t\geq 0}$ is positive Harris recurrent if

(∗) $y \to H(y,1)$ is finite and belongs to $L^1(Q^{\mathrm{i}})$

then: void configuration δ is a recurrent atom, invariant measure on S:

$$\mu(F) = E_\delta\left(\int_0^R 1_F(\eta_s)\,ds\right) \quad , \quad F \in \mathcal{S}$$

with R the time of first return to δ, invariant occupation measure on E:

$$\overline{\mu}(A) = E_\delta\left(\int_0^R \eta_s(A)\,ds\right) \quad , \quad A \in \mathcal{E}\ ;$$

by (∗), $\overline{\mu}$ is a finite measure on E given by (up to constant multiples)

$$\overline{\mu} = Q^{\mathrm{i}} H \quad , \quad \overline{\mu}(E) = E_{Q^{\mathrm{i}}}\left(\int_0^\infty e^{-\int_0^t [\kappa(1-\rho)](\widetilde{\xi}_s)\,ds}\,dt\right) < \infty$$

(case $Q^{\mathrm{r}} = \epsilon_0$: cf. Ikeda, Nagasawa and Watanabe 1969, Nagasawa 1977, Löcherbach 2004, H-L 2005, ... our case: jump laws Q^{r} allow for continuous Lebesgue density of μ on S, Hammer 2012)

more on the invariant measure

associate to the $[\kappa(1-\rho)]$-potential kernel $H(y,g)$ expectation semigroup

$$M_t(y,g) = E_y\left(g(\widetilde{\xi}_t)\, e^{-\int_0^t [\kappa(1-\rho)](\widetilde{\xi}_s)\, ds} \right) \;, \quad t \geq 0$$

(Ikeda, Nagasawa and Watanabe 1969), write $|||M_t||| = \sup_{y\in E} M_t(y,E)$.

theorem 1: can construct $\alpha^\top$ in $\mathcal{C}_b(E)$ and a semimartingale of finite jump intensity $\widetilde{\xi}^\top$ with generator $\widetilde{\mathcal{L}}^\top$ and semigroup

$$M_t^\top(y,g) = E_y\left(g(\widetilde{\xi}_t^\top)\, e^{-\int_0^t \alpha^\top(\widetilde{\xi}_s^\top)\, ds} \right) \;, \quad t \geq 0$$

such that duality of Feller semigroups holds:

$$\langle (\widetilde{\mathcal{L}} - [\kappa(1-\rho)])f \,,\, g \rangle = \langle f \,,\, (\widetilde{\mathcal{L}}^\top - \alpha^\top) g \rangle \;, \quad f, g \in \mathcal{C}_c^\infty(E)$$

$$\langle M_t f, g \rangle = \langle f, M_t^\top g \rangle \;, \quad t \geq 0$$

$$|||M_t||| \leq e^{t\, \|[\kappa(1-\rho)]\|_\infty} \quad , \quad |||M_t^\top||| \leq e^{t\, \|\alpha^\top\|_\infty}$$

(proofs using strongly continuous semigroups, approximation results, bounded operators on $L^1(E)$, bounded perturbation – no flows of diffeomorphisms) formally by prop. 1 and thm. 1, when q^{i} is a Lebesgue density for Q^{i}, write

$$\overline{\mu} f = Q^{\mathrm{i}} H\, f = \langle\, q^{\mathrm{i}}, Hf \,\rangle = \langle\, H^{\top} q^{\mathrm{i}}, f \,\rangle \;\leq\; \infty$$

where $H^{\top}(y,g)$ is the resolvent of the semigroup $(M_t^{\top})_{t\geq 0}$, thus

$$y \longrightarrow H^{\top}(y, q^{\mathrm{i}}) = \int_0^{\infty} dt\, M_t^{\top}(y, q^{\mathrm{i}}) = E_y \left(\int_0^{\infty} q^{\mathrm{i}}(\widetilde{\xi}_t^{\top})\, e^{-\int_0^t \alpha^{\top}(\widetilde{\xi}_s^{\top})\, ds}\, dt \right)$$

should be a Lebesgue density for $\overline{\mu}$ on E

theorem 2: assume $(*)$ and

$$(**) \qquad \limsup_{t\to\infty} \frac{1}{t} \log |||M_t||| \;<\; 0 \quad , \quad \limsup_{t\to\infty} \frac{1}{t} \log |||M_t^{\top}||| \;<\; 0$$

a) if $q^{\mathrm{i}} \in \mathcal{C}_0(E)$, then the invariant occupation measure $\overline{\mu}$ on E admits a Lebesgue density $\overline{\gamma}(\cdot) = H^{\top}(\cdot, q^{\mathrm{i}})$ which is $\mathcal{C}_0(E)$

b) for $q \in \mathbb{N}$: if condition (Mq): $y \to \sum_k k^q p_k(y)$ is $\mathcal{C}_b(E)$ holds, then

$$\mu(\ell^q) = \int_S \ell^q(x)\, \mu(dx) = \sum_{\ell} \ell^q \mu(E^{\ell}) \;<\; \infty$$

the reconstruction algorithm (RA)

from now on: for the semigroup $(P_t^\kappa(y,dz))_{t\geq 0}$ on $E=\mathbb{R}^d$ corresponding to one-particle motion ξ killed at rate $\kappa(\cdot)$, we assume heat kernel bounds

$$(HKB) \qquad p_t^\kappa(y,z) \;\leq\; C\,t^{-\frac{d}{2}}\,e^{-\frac{1}{2}\frac{|z-y|^2}{C\,t}}\ , \quad 0<t\leq t_0\ ,\ y,z\in E$$

definition 1: call a two-particle configuration $(x,x')\in E\times E$ ε-wellspread if $\min_{1\leq j\leq d}|x_j-x_j'|\;\geq\;\varepsilon$; call a configuration $x\in S$ ε-wellspread if arbitrary two-particle subconfigurations of x are ε-wellspread

write N_ε for the set of all configurations in S which are not ε-wellspread

theorem 3: under (HKB) and $(M3)$, we have

$$\mu(N_\varepsilon)\;\leq\;O(\varepsilon)\quad \text{as } \varepsilon\downarrow 0.$$

proof: resolvant calculations for multi-particle motions under killing rate $\kappa(\cdot)$ establish $\mu(N_\varepsilon\cap E^\ell)\leq D\,\varepsilon\,\ell^3\mu(E^\ell)$ for $\ell\geq 2$

when we observe discretely in time: recording pairs of successive observations $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ only, any information on individual particle trajectories between times $i\Delta$ and $(i+1)\Delta$ is lost ... $\hookrightarrow$ problem of particle identification !!!

definition 2: call $(x,y) \in S \times S$ (Δ,λ)-identifiable if $(\Delta > 0,\ 0 < \lambda < \frac{1}{2})$

- $\ell(x) = \ell(y) = \ell$ for some $\ell \in \mathbb{N}$
- $x = (x_1, \ldots, x_\ell)$ is $4\Delta^\lambda$-wellspread, $y = (y_1, \ldots, y_\ell)$ is $2\Delta^\lambda$-wellspread
- exists permutation π of the ℓ particles (necessarily unique) such that

$$\max_{1 \le j \le d} |y_{\pi(m),j} - x_{m,j}| < \Delta^\lambda \quad , \quad 1 \le m \le \ell$$

write $\mathtt{ID}(\Delta,\lambda)$ for the subset of (Δ,λ)-identifiable pairs in $S \times S$, call pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ (Δ,λ)-identifiable if $(\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \mathtt{ID}(\Delta,\lambda)$ holds

reconstruction algorithm: for (Δ,λ)-identifiable pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$, writing $x := \eta_{i\Delta}$, $y := \eta_{(i+1)\Delta}$, ℓ and π as in definition 2, we decide to view

$$(RA) \qquad \left\{ \begin{array}{l} y_{\pi(m)} \text{ as the position at time } (i+1)\Delta \text{ of the particle} \\ \text{which was in position } x_m \text{ at time } i\Delta, \text{ for } 1 \le m \le \ell \end{array} \right.$$

(of course, this decision may be wrong ...)

theorem 4: for $(\eta_t)_{t\geq 0}$ stationary, $0<\lambda<\frac{1}{2}$ fixed, as $\Delta\downarrow 0$:

$$Q_\mu\left((\eta_{i\Delta},\eta_{(i+1)\Delta}) \notin \mathtt{ID}(\Delta,\lambda) \text{ or } \text{(RA) decision incorrect} \right) \;\leq\; O(\Delta^\lambda)$$

proof: define 'good = correctly identifiable' path segments

$$\eta_{[i\Delta,(i+1)\Delta]} := (\eta_s)_{i\Delta\leq s\leq(i+1)\Delta}$$

as elements of the cadlag path space $D([i\Delta,(i+1)\Delta],S)$ such that

$$\{\eta_{[i\Delta,(i+1)\Delta]}\text{'good'}\} \;\subset\; \{(\eta_{i\Delta},\eta_{(i+1)\Delta})\in \mathtt{ID}(\Delta,\lambda) \text{ and } \text{(RA) decision is correct}\}$$

we prove under assumptions (HKB) and $(M1)$ that

$$Q_\mu\left(\eta_{[i\Delta,(i+1)\Delta]} \text{ is } (\Delta,\lambda)\text{-good}\right) \;\geq\; 1-O(\Delta^\lambda)$$

as $\Delta\downarrow 0$, with leading contribution on l.h.s. (cf. thm. 3 and def. 2)

$$Q_\mu\left(\eta_{i\Delta} \text{ is } 4\Delta^\lambda\text{-wellspread}\right) = 1-\mu(N_{4\Delta^\lambda}) \;\geq\; 1-O(\Delta^\lambda)$$

which explains the rate in theorem 4; as a by-product:

$$(+) \qquad Q_\mu\left((\eta_{i\Delta},\eta_{(i+1)\Delta})\in \mathtt{ID}(\Delta,\lambda)\,,\; \text{(RA) decision incorrect}\right) \;\leq\; O(\Delta)$$

filling regression schemes for $\sigma^2(\cdot)$ (dim $d = 1$)

$E := \mathbb{R}$, let $\mathtt{CI}(\Delta, \lambda)$ denote the set of '(Δ,λ)-good' path segments:

$$\{\, \eta_{[m\Delta,(m+1)\Delta]} \in \mathtt{CI}(\Delta,\lambda) \,\} \subset \{\, (\eta_{m\Delta}, \eta_{(m+1)\Delta}) \in \mathtt{ID}(\Delta,\lambda) \,\}$$

consider any interval A such that $\inf_{x \in A} \overline{\gamma}(x) > 0$, w.l.o.g. $A := [0,1]$

subdivide A into n cells A_j of equal length, by ergodicity

(since $\lim_{i\to\infty} \frac{1}{i} \sum_{m=1}^{i} 1_{\{\eta_{m\Delta}(A_j)\geq 1\}} 1_{\{\, \eta_{[m\Delta,(m+1)\Delta]} \in \mathtt{CI}(\Delta,\lambda)\,\}}$ exists in $(0,\infty)$)

associate to every cell A_j a stopping time τ_j with finite mean, $1 \leq j \leq n$:

$$\tau_j \;:=\; \min\left\{ m \in \mathbb{N}_0 : \eta_{m\Delta}(A_j) \geq 1 \text{ and } (\eta_{m\Delta}, \eta_{(m+1)\Delta}) \in \mathtt{ID}(\Delta,\lambda) \right\}$$

take some m_j such that particle m_j in configuration $\eta_{\tau_j\Delta}$ visits A_j, then put

$$(RS) \quad \mathcal{X}_j := (\eta_{\tau_j\Delta})_{m_j} \in A_j \;,\; \mathcal{Y}_j := \left(\frac{(\eta_{(\tau_j+1)\Delta})_{\pi(m_j)} - (\eta_{\tau_j\Delta})_{m_j}}{\sqrt{\Delta}} \right)^2 ,\; 1 \leq j \leq n$$

with π prescribed by the reconstruction algorithm (RA) (we are allowed to fill several cells at the same time) theorem 5: assume $d = 1$, (HKB), (M3), ..., then the regression scheme (RS)

$$(\mathcal{X}_j \, , \, \mathcal{Y}_j) \; , \; 1 \le j \le n$$

has the following properties as $n \to \infty$ and $\Delta \downarrow 0$:

a) the $\mathcal{X}_j$, $1 \le j \le n$, are approximately aequidistant

b) on an exceptional event of small probability $\le O(n\Delta)$,
some $\mathcal{Y}_j$'s in (RS) may not correspond to underlying one-particle motions; but (RA) always provides a trivial deterministic bound $|\mathcal{Y}_j| \le \Delta^{2\lambda-1}$, $1 \le j \le n$

c) on an event of large probability $\ge 1 - O(n\Delta)$, the scheme (RS) satisfies

$$\mathcal{Y}_j \;=\; \sigma^2(\mathcal{X}_j)(1+U_j) + R_j \; , \quad 1 \le j \le n$$

where for iid BM's W_j, $1 \le j \le n$, iid pairs (U_j, R_j) and some global cst

$$U_j = 2\int_0^1 W_j dW_j = [W_j(1)^2 - 1] \quad , \quad E(R_j^2) \le C\,\Delta$$

(proof: recall $(\eta_{\tau_j \Delta}, \eta_{(\tau_j+1)\Delta}) \in \mathtt{ID}(\Delta, \lambda)$ for all $1 \le j \le n$ by construction;
c) $\leftrightarrow$ event that (RA) decides correctly, thus all $\mathcal{Y}_j$ are taken out of some true underlying one-particle motion; b) $\leftrightarrow$ event $(+)$ in theorem 4)

estimation of the diffusion coefficient at points $a \in \mathrm{int}(A)$, $A = [0,1]$: for n large and Δ small, consider regression schemes (RS)

unknown $\sigma^2(\cdot)$, assume $\sigma^2 \in \mathcal{H}^\beta$, the Hölder class of smoothness $\beta > 2$

consider as an example kernel estimators (other examples could be local polynomial estimators, cf. Tsybakov 2009 section 1.6)

kernel K of order $\lfloor \beta \rfloor$, bandwidth $h = n^{-\frac{1}{2\beta+1}}$, $K_h = \frac{1}{h} K(\frac{\cdot}{h})$, estimator

$$\widehat{\sigma^2_{n,\Delta}}(a) \ := \ \frac{1}{n} \sum_{j=1}^{n} \mathcal{Y}_j \, K_h(\mathcal{X}_j - a)$$

(since $\mathcal{X}_j$, $1 \le j \le n$, approximately aequidistant: $\frac{1}{n} \sum_{j=1}^{n} K_h(\mathcal{X}_j - a) \approx 1$)

quadratic risk when estimating from (RS)

$$\sup_{\sigma^2 \in \mathcal{H}^\beta} E_{\sigma^2} \left(\left| \widehat{\sigma^2_{n,\Delta}}(a) - \sigma^2(a) \right|^2 \right) \ \le \ O\left(n \, \Delta^{2\lambda} \right) + O\left(n^{-\frac{2\beta}{2\beta+1}} \right)$$

arises as the sum of two terms:

- on the exceptional set of probability $O(n\Delta)$: all $|\mathcal{Y}_j| \le \Delta^{2\lambda - 1}$
- usual nonparametric squared risk on event of probability $1 - O(n\Delta)$

both terms will be of same order

$$O\left(n^{-\frac{2\beta}{2\beta+1}}\right) = O\left(\frac{1}{n\,h_n}\right) = O\left(n\,\Delta^{2\lambda}\right)$$

if we take $\Delta = \Delta_n$ such that $\Delta^{2\lambda} = \frac{1}{n^2 h_n} = n^{-\frac{4\beta+1}{2\beta+1}}$: we thus arrive at

corollary: estimating unknown $\sigma^2 \in \mathcal{H}^\beta$, $\beta > 2$ fixed, from time-discrete observation of the BDI process at step size Δ

$$\Delta = \Delta_n \text{ such that } \Delta^\lambda = \frac{1}{n\sqrt{h_n}} = \frac{1}{n}\, n^{+\frac{1/2}{2\beta+1}}$$

using reconstruction algorithm (RA) to fill a regression scheme (RS) with n cells, we have asymptotically as $n \to \infty$

$$\sup_{\sigma^2 \in \mathcal{H}^\beta} E_{\sigma^2}\left(\left|\widehat{\sigma^2_{n,\Delta}}(a) - \sigma^2(a)\right|^2\right) \leq O\left(n^{-\frac{2\beta}{2\beta+1}}\right)$$

i.e. we attain the nonparametric rate known to be optimal (Tsybakov 2009 sect. 2.5) for squared risk in standard regression schemes for unknown $f \in \mathcal{H}^\beta$

$$(\mathcal{U}_j, \mathcal{V}_j)\ ,\ \mathcal{V}_j = f(\mathcal{U}_j) + \epsilon_j\ ,\ 1 \leq j \leq n$$

with iid errors ϵ_j and equispaced deterministic $\mathcal{U}_j$

–THE END–

references

Berg (2015): Nonparametric estimation of the diffusion coefficient of a branching diffusion with immigration. PhD thesis, Mainz.

Brandt (2005): Partial reconstruction of the trajectories of a discretely observed branching diffusion with immigration ... PhD thesis, Mainz.

Hammer (2012): Ergodicity and regularity of invariant measure for branching Markov processes with immigration. PhD thesis, Mainz.

Höpfner, Löcherbach (2005): Remarks on ergodicity and invariant occupation measure in branching diffusions with immigration. AIHP **41**.

Ikeda, Nagasawa, Watanabe (1969): Branching Markov processes III. J. Math. Kyoto Univ. **9**.

Kessler, Lindner, Sørensen (2012): Statistical methods for stochastic differential equations. CRC Press.

Löcherbach (2004): Smoothness of the intensity measure density for interactiong branching diffusions with immigration. JFA **215**.

Nagasawa (1977): Basic models of branching processes. Bull. Intern. Statist. Instit. **27**.

Tsybakov (2008): Introduction to nonparametric estimation. Springer.