Convergence exponentielle uniforme vers la distribution quasi-stationnaire en dynamique des populations

Nicolas Champagnat, Denis Villemonais

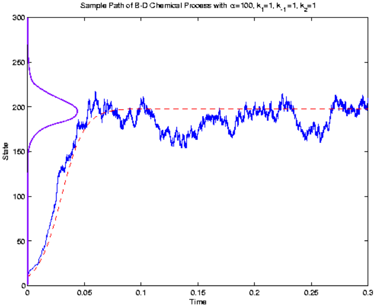
Institut ÉLIE CARTAN Nancy

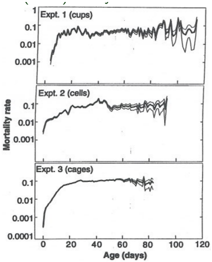
Inria informatiques mathématiques

Colloque Franco-Maghrébin d'Analyse Stochastique, Nice, 25/11/2015

Population dynamics with extinction

- Stationary behavior of population dynamics: usually extinction
- However, we only observe populations which are not extinct, and most often, they exhibit a stationary behavior
- Quasi-stationary distributions are stationary distributions conditionally on non-extinction
- Quasi-stationary populations exhibit a mortality plateau
- Questions: Speed of convergence of conditional distributions? Dependence w.r.t. initial distribution?





Carey et al., 1992

Notations

We consider a general stochastic population dynamics $(X_t, t \geq 0)$ with values in $E \cup \{\partial\}$ s.t.

- E is arbitrary (measurable space), e.g. $E = \mathbb{N}$, $E = (0, \infty)$
- $\partial \notin E$ is a cemetary point (corresponding to extinction or absorption)
- X is a Markov process
- $\mathbb{P}_x$ denotes its law given $X_0 = x \in E \cup \{\partial\}$, and for μ probability measure on E, $\mathbb{P}_\mu = \int_E \mathbb{P}_x \, d\mu(x)$
- $\tau_\partial := \inf\{t \geq 0 : X_t = \partial\}$

Assumptions:

- $\forall t \geq \tau_\partial$, $X_t = \partial$ (absorbing point)
- $\forall x \in E$, $\mathbb{P}_x(\tau_\partial < \infty) = 1$ (almost sure absorption)
- $\forall x \in E, \forall t \geq 0$, $\mathbb{P}_x(t < \tau_\partial) > 0$ (possible survival up to any time)

Quasi-stationary distributions

Définition

- *Quasi-stationary distribution (QSD):* $\alpha \in \mathcal{P}(E)$ *s.t.* $\forall t \geq 0$, $\mathbb{P}_\alpha(X_t \in \cdot \mid t < \tau_\partial) = \alpha$
- *Quasi-limiting distribution (QLD):* $\alpha \in \mathcal{P}(E)$ *s.t.* $\exists \mu \in \mathcal{P}(E)$, $\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) \to \alpha$ *(for some topology)*
- *Yaglom limit:* $\alpha \in \mathcal{P}(E)$ *s.t.* $\forall x \in E$, $\mathbb{P}_x(X_t \in \cdot \mid t < \tau_\partial) \to \alpha$
- *Universal QLD:* $\alpha \in \mathcal{P}(E)$ *s.t.* $\forall \mu \in E$, $\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) \to \alpha$

Proposition

- α *universal QLD* $\Rightarrow \alpha$ *Yaglom limit* $\Rightarrow \alpha$ *QLD* $\Leftrightarrow \alpha$ *QSD*
- *If* α *is a QSD, then* $\exists \lambda_0 > 0$ *s.t.* $\mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda_0 t}$
- $L^*\alpha = -\lambda_0 \alpha$, *where* L *is the infinitesimal generator of* X

Observation of a QSD and domain of attraction

- Even in cases where all states of the population communicate, there is not necessarily uniqueness of a QSD
  $\rightsquigarrow$ for example, subcritical branching processes have infinitely many QSDs.
- Then, which QSD is observed in a population? After what time?
- To each QSD α is associated a basin of attraction, which is the set of initial distributions for which α is the QLD.
- Usually, the basin of attraction of a QSD is determined by the tail of the initial distribution (e.g. for branching processes, finiteness of moments)
- In practice, when the initial distribution is only estimated from observations, one can't determine to which basin of attraction it belongs
- In addition, the speed of convergence of conditional distributions strongly depends on the basin of attraction.

Aim of the talk

- Important question in practice: is there a unique QSD which attracts all the initial distributions, i.e. a universal QLD, and is the speed of convergence of conditional distributions uniform?
- First goal: necessary and sufficient conditions for the existence (and uniqueness) of a universal QSD with uniform exponential rate of convergence.
- Second goal: apply these criteria to various classes of population dynamics and types of extinction.

- The mathematical study of QSD makes usually use of spectral tools (Perron-Frobenius, Krein-Rutman, Sturm-Liouville, L^2 theory for self-adjoint operators...)
- This talk: probabilistic criteria (coupling estimates...) avoiding the use of spectral theory (e.g. no need to assume that the processes are self-adjoint)
- Our criteria are inspired from known criteria in the irreducible case (ergodicity coefficient of Dobrushin, Doeblin condition, cf. Meyn and Tweedie, 1993)

The criterion

Assumption (A)

There exists a probability measure ν on E s.t.

(A1) $\exists t_0, c_1 > 0$ *s.t.* $\forall x \in E$,

$$\mathbb{P}_x(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu(\cdot)$$

(A2) $\exists c_2 > 0$ *s.t.* $\forall x \in E$ *and* $\forall t \geq 0$,

$$\mathbb{P}_\nu(t < \tau_\partial) \geq c_2 \mathbb{P}_x(t < \tau_\partial)$$

Interpretation of Conditions (A)

(A1) Coupling of conditional laws independently of the initial condition.
Typically, $\nu = c\mathbb{1}_K$ where K is a compact set.
Example: diffusion on $E = \mathbb{R}_+^*$ and $\partial = 0$, $K = \{a \leq x \leq b\}$.

- if $x > b$, we want $\mathbb{P}_x(\tau_K < t_0) \geq c$ (here conditioning plays no role)
- if $x < a$, we want $\mathbb{P}_x(\tau_K < t_0 \mid t_0 < \tau_\partial) \geq c$.

$\Longrightarrow$ the process must

- go down from infinity ($+\infty$ is an entrance boundary)
- go away from 0 conditionally on non-absorption when it starts close to 0
- couple inside K

(A2) Harnack inequality: the populations survives nowhere much better than starting from ν.
True for example if $\inf_{y\in K} \mathbb{P}_y(t < \tau_\partial) \geq c \sup_{x\in E} \mathbb{P}_x(t < \tau_\partial)$

Ergodicity coefficient of Dobrushin

In the case where $(X_t, t \geq 0)$ is irreducible on E, if there exist $t_0, c_1 > 0$ and $\nu \in \mathcal{P}(E)$ s.t. $\forall x \in E$, $\mathbb{P}_x(X_{t_0} \in \cdot) \geq c_1 \nu$, then

- $\forall x, y \in E$,

$$\|\mathbb{P}_x(X_{t_0} \in \cdot) - \mathbb{P}_y(X_{t_0} \in \cdot)\|_{TV} \leq 2(1-c_1) = (1-c_1)\|\delta_x - \delta_y\|_{TV} < 2$$

- $\forall \mu_1, \mu_2 \in \mathcal{P}(E)$ (the set of probability measures on E),

$$\|\mathbb{P}_{\mu_1}(X_{t_0} \in \cdot) - \mathbb{P}_{\mu_2}(X_{t_0} \in \cdot)\|_{TV} \leq (1 - c_1)\|\mu_1 - \mu_2\|_{TV},$$

- and by the Markov property, $\forall x, y \in E$,

$$\|\mathbb{P}_x(X_{kt_0} \in \cdot) - \mathbb{P}_y(X_{kt_0} \in \cdot)\|_{TV}$$
$$\leq (1-c_1)\|\mathbb{P}_x(X_{(k-1)t_0} \in \cdot) - \mathbb{P}_y(X_{(k-1)t_0} \in \cdot)\|_{VT} \leq \ldots \leq 2(1-c_1)^k$$

- $\rightsquigarrow$ exponential convergence in total variation to a unique invariant measure.

Necessary and sufficient condition for the existence of a universal QLD [C-V, PTRF, 2015]

Theorem

Condition (A) is equivalent to the existence of a universal QLD α *s.t.* $\forall \mu \in \mathcal{P}(E)$

$$\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \alpha\|_{VT} \leq Ce^{-\gamma t}$$

for two constants $C, \gamma > 0$ *independent of* μ.
In addition, under Condition (A), we can take $C = 2/(1 - c_1 c_2)$ *and* $\gamma = -\log(1 - c_1 c_2)/t_0$.

Some comments

- We obtain a necessary and sufficient condition
- The convergence is uniform w.r.t. the initial distribution
- A bound on the exponential rate of convergence can be computed from the constants c_1, c_2, but this is non-optimal
- Difficulty: how to check (A1) and (A2) in practice?

Idea of the proof of sufficiency ($t_0 = 1$)

Step 1: $\mathbb{P}_x(X_1 \in \cdot \mid t < \tau_\partial) \geq c_1 c_2 \nu_t$ where $\nu_t \in \mathcal{P}(E)$.

$$\begin{aligned}
\mathbb{P}_x(X_1 \in A \mid t < \tau_\partial) &= \frac{\mathbb{P}_x(1 < \tau_\partial)}{\mathbb{P}_x(t < \tau_\partial)} \mathbb{E}_x[\mathbb{1}_A(X_1)\mathbb{P}_{X_1}(t-1 < \tau_\partial) \mid 1 < \tau_\partial] \\
&\geq c_1 \frac{\mathbb{P}_x(1 < \tau_\partial)}{\mathbb{P}_x(t < \tau_\partial)} \nu[\mathbb{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial)] \\
&\geq c_1 \frac{\nu[\mathbb{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial)]}{\sup_{y \in E} \mathbb{P}_y(t-1 < \tau_\partial)} \geq c_1 c_2 \nu_t(A)
\end{aligned}$$

where $\nu_t(A) = \dfrac{\nu[\mathbb{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial)]}{\mathbb{P}_\nu(t-1 < \tau_\partial)}$.

Idea of the proof (continued)

Step 2: define
$R^T_{s,t}f(x) = \mathbb{E}_x(f(X_{t-s}) \mid T-s < \tau_\partial) = \mathbb{E}(f(X_t) \mid X_s = x, T < \tau_\partial)$.
Then, for all $u \leq s \leq t \leq T$,

$$R^T_{u,s} R^T_{s,t} f = R^T_{u,t} f.$$

Step 3: Then, step 1 says that $\delta_x R^T_{s,s+1} - c_1 c_2 \nu_{T-s} \geq 0$, with mass $\leq 1 - c_1 c_2$, and so

$$\|\delta_x R^T_{s,s+1} - \delta_y R^T_{s,s+1}\|_{VT} \leq 2(1 - c_1 c_2) = (1 - c_1 c_2)\|\delta_x - \delta_y\|_{VT}.$$

The end of the proof is similar to the classical coupling coefficient argument of Dobrushin.

Asymptotic behavior of survival probabilities and eigenfunction

We can also deduce from Condition (A) the spectral properties of the QSD.

Proposition

Under Condition (A), there exists $\eta : E \cup \{\partial\} \to \mathbb{R}_+$ bounded such that $\eta(\partial) = 0$ and $\eta > 0$ on E s.t.

$$\eta(x) = \lim_{t\to\infty} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\alpha(t < \tau_\partial)} = \lim_{t\to\infty} e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial)$$

where the convergence is uniform w.r.t. x.
In addition $\eta \in \mathcal{D}(L)$ where L is the infinitesimal generator of P_t, and $L\eta = -\lambda_0 \eta$.

This result explains the phenomenon of mortality plateau.

Exponential ergodicity of the Q-process

Theorem

Assume (A), then

(i) $\exists(\mathbb{Q}_x)_{x\in E}$ *proba s.t. for all* $s>0$ *and* $A\in\mathcal{F}_s$,

$$\lim_{t\to\infty}\mathbb{P}_x(A\mid t<\tau_\partial)=\mathbb{Q}_x(A),$$

and X_t *is an homogeneous Markov process under* $(\mathbb{Q}_x)_{x\in E}$

(ii) *Under* $(\mathbb{Q}_x)_{x\in E}$, X_t *has semi-group* $\tilde{P}_t$ *defined by*

$$\tilde{P}_t\varphi(x)=\frac{e^{\lambda_0 t}}{\eta(x)}P_t(\eta\varphi)(x)$$

(iii) $\beta(dx)=\dfrac{\eta(x)\alpha(dx)}{\int\eta\,d\alpha}$ *is the unique invariant distribution of* X *under* $(\mathbb{Q}_x)$ *and* $\forall\mu\in\mathcal{P}(E)$, $\|\mathbb{Q}_\mu(X_t\in\cdot)-\beta\|_{VT}\leq Ce^{-\gamma t}$

Applications to 1D birth and death processes

Let X_t be a BDP on $\mathbb{Z}_+$ with birth rate λ_n and death rate μ_n in state n.

- $\partial = 0$ is absorbing if $\lambda_0 = \mu_0 = 0$; we assume $\lambda_n > 0$ and $\mu_n > 0$ for all $n \geq 1$.
- (A1) is true for all measure ν with finite support, for example $\nu = \delta_1$, if the BDP comes down from infinity, i.e. if $\inf_n \mathbb{P}_n(T_k < t) > 0$ for some $k \geq 1$ and $t > 0$, which is equivalent to

$$S = \sum_{n \geq 0} \frac{1}{\mu_{n+1}} \left(1 + \frac{\lambda_n}{\mu_n} + \ldots + \frac{\lambda_n \ldots \lambda_1}{\mu_n \ldots \mu_1} \right) < \infty$$

- The difficulty is to prove (A2)

Proposition (Martinez, San Martin, Villemonais, 2012)

Assume $S < \infty$. Then (A2) is satisfied, and there exists a unique universal QLD, with exponential convergence in total variation.

Proof of (A2)

- We have $\sup_x \mathbb{E}_x(\tau_\partial) = \sum_{k\geq 1} \mathbb{E}_k(T_{k-1}) = S < \infty$, where T_k is the first hitting time of k.
- It is standard to deduce (cf. e.g. Cattiaux et al., 2009) that $\forall \lambda' > 0$, $\exists k \geq 1$ s.t. $\sup_{x\geq k} \mathbb{E}_x(e^{\lambda' T_k}) < \infty$.
- De plus, $\mathbb{P}_1(t < \tau_\partial) \geq \mathbb{P}_1(X_t = 1) \geq e^{-(\lambda_1+\mu_1)t}$. Then we can choose $\lambda' > \lambda_1 + \mu_1$ in the last statement.
- On a aussi $\mathbb{P}_1(t \leq \tau_\partial) \geq \mathbb{P}_1(t+1 < \tau_\partial) \geq \mathbb{P}_1(X_1 = k)\mathbb{P}_k(t < \tau_\partial) \geq C\mathbb{P}_k(t < \tau_\partial)$.

Proof of (A2) (continued)

Therefore,

$$\begin{aligned}
\mathbb{P}_\infty(t<\tau_\partial) &\leq \mathbb{P}_\infty(t<T_k)+\mathbb{P}_\infty(T_k<t<\tau_\partial)\\
&\leq Ce^{-\lambda' t}+\int_0^t \mathbb{P}_k(t-s<\tau_\partial)d\mathbb{P}_\infty(T_k>s)\\
&\leq Ce^{-\lambda' t}+C\int_0^t \mathbb{P}_1(t-s<\tau_\partial)d\mathbb{P}_\infty(T_k>s)\\
&\leq Ce^{-\lambda' t}+C\int_0^t \mathbb{P}_1(t-s<\tau_\partial)d\mathbb{P}_\infty(T_k>s)\\
&\leq Ce^{-\lambda' t}+C\mathbb{P}_1(t<\tau_\partial)\int_0^t \frac{d\mathbb{P}_\infty(T_k>s)}{\mathbb{P}_1(s<\tau_\partial)}\\
&\leq Ce^{-\lambda' t}+C\mathbb{P}_1(t<\tau_\partial)\int_0^t e^{(\lambda_1+\mu_1)s}d\mathbb{P}_\infty(T_k>s).
\end{aligned}$$

Hence,

$$\mathbb{P}_\infty(t<\tau_\partial)\leq C\mathbb{P}_1(t<\tau_\partial).$$

Comments and extensions

- Most of the classical biological models with density-dependence satisfy $S < \infty$. E.g. the logistic BDP, Gompertz BDP... Subcritical branching processes do not satisfy $S < \infty$.
- For 1D BDP, the full picture is known, since when $S = \infty$, it is known that there is no uniqueness of the QSDs (van Doorn, 1991).
- The method is easy to extend to many cases where the absorption rate is bounded. For example, multi-dimensional BDP aborbed at $(0, \dots, 0)$, BDP processes with catastrophes at bounded rate...
- Cases with unbounded absorption rate are harder to study, because in this case, it is usually not sufficient that the process comes down from infinity to check (A1) $\rightsquigarrow$ next parts.

Multi-dimensional BDP absorbed at the boundary [C-V, preprint, 2015]

We consider a BDP in $\mathbb{Z}_+^r$, absorbed at $\partial = \mathbb{Z}_+^r \setminus (\mathbb{N})^r$, with transition rates from $n = (n_1, \ldots, n_r)$ to

$$\begin{cases} n + e_j & \text{with rate } n_j b_j(n), \\ n - e_j & \text{with rate } n_j \left(d_j(n) + \sum_{k=1}^r c_{jk}(n) n_k\right), \end{cases}$$

where $b_j(n), d_j(n), c_{jk}(n) \geq 0$.

We assume that $b_j(n) \leq \bar{b}$ and $d_j(n) \leq \bar{d}$ and that $c_{jk}(n) \geq \underline{c} > 0$.

Exemple : neutral logistic BDP: jumps from

$$\begin{aligned} &n \text{ to } n + e_i \quad \text{with rate} \quad b n_i \\ &n \text{ to } n - e_i \quad \text{with rate} \quad n_i \left(d + c \sum_{j=1}^{d} n_j\right). \end{aligned}$$

Assumption

We are able to prove that (A1) and (A2) are satisfied when, for $|n|$ large enough,

$$c_{ii}(n) \geq C_r \left(\sum_{1 \leq j \neq k \leq r} c_{jk}(n) + \frac{1}{|n|} \sum_{j=1}^{r} c_{jj}(n) \right),$$

for some explicit constant C_r depending only on r.

This means that intra-specific competition is stronger than the inter-specific competition.

Theorem

Under the previous conditions, Condition (A) is satisfied, and there exists a unique universal QLD, with exponential convergence in total variation.
In the special case of the neutral logistic BDP, Condition (A) is satisfied if $r \leq 3$.

Comments

- The quasi-stationary behavior of multi-dimensional BDP process absorbed at the boundary have been the object of very few works, except in the case of multi-type Galton-Watson processes. The case of diffusions on $\mathbb{R}_+^r$ was more studied (cf. Cattiaux, Méléard, 2010).
- Our results are certainly not optimal. In particular, we expect that the neutral logistic BDP satisfies (A) in any dimension, but our method of proof does not work in general.
- The existence of a Yaglom limit (instead of a universal QLD) with exponential speed of convergence (but non-uniform w.r.t. the initial distribution) can be proved in general using another approach based on Lyapunov functions [C-V, in preparation].

Method of proof

Proof of (A1): Find Lyapunov functions for the conditional distributions: check that

$$\mu_t(V) - V(n) \leq \int_0^t \left[\mu_s(LV) - \mu_s(V)\mu_s(L\mathbb{1}_{\mathbb{N}^r})\right] ds,$$

and check that

$$V_\varepsilon(n) = \sum_{j=1}^{|n|} \frac{1}{j^{1+\varepsilon}},$$

satisfies for all probability measure μ

$$\mu(LV_\varepsilon) - \mu(V_\varepsilon)\mu(L\mathbb{1}_{\mathbb{N}^r}) \leq C - \frac{1}{C} \sum_{n\in\mathbb{N}^r} |n|^{\gamma-\gamma\beta_2-\varepsilon}\mu(n).$$

Proof of (A2): similar as in the 1D case.

Application to one-dimensional diffusions [C-V, preprint, 2015]

We consider a diffusion $(X_t, t \geq 0)$ sur $[0, \infty)$ absorbed at 0 on natural scale, i.e. (typically)

$$dX_t = \sigma(X_t)dB_t, \quad \forall t < \tau_\partial,$$

(e.g. Fisher diffusion, or Wright-Fisher on $[0, 1]$). We assume that 0 is accessible (i.e. either exit or regular).

Theorem

Condition (A) is equivalent to the condition

(B) *The diffusion $(X_t, t \geq 0)$ comes down from infinity (i.e. $+\infty$ is an entrance boundary) and there exist $t, A > 0$ s.t.*

$$\mathbb{P}_x(t < \tau_\partial) \leq Ax, \quad \forall x > 0.$$

We also prove that nearly all diffusions coming down from infinity and a.s. absorbed at 0 satisfy (A) (the only case we do not cover are those of strong oscillations of σ close to 0).

A glimpse of the proof

Again, we focus on (A) $\Longrightarrow$ (B), and only explain the

New difficulty: prove that conditionally on extinction, the diffusion started close to 0 goes away from 0: $\exists \varepsilon, c > 0$ s.t.

$$\mathbb{P}_x(X_{t_1} \geq \varepsilon \mid t_1 < \tau_\partial) \geq c, \quad \forall x > 0.$$

Key argument: Since X is a local martingale,

$$\begin{aligned} x = \mathbb{E}_x(X_{t_1 \wedge T_1}) &= \mathbb{P}_x(t_1 < \tau_\partial)\mathbb{E}_x(X_{t_1 \wedge T_1} \mid t_1 < \tau_\partial) + \mathbb{P}_x(T_1 < \tau_\partial \leq t_1) \\ &\leq \mathbb{P}_x(t_1 < \tau_\partial)\mathbb{E}_x(X_{t_1 \wedge T_1} \mid t_1 < \tau_\partial) + \mathbb{P}_x(T_1 < \tau_\partial)\mathbb{P}_1(\tau_\partial \leq t_1) \\ &\leq \mathbb{P}_x(t_1 < \tau_\partial)\mathbb{E}_x(X_{t_1 \wedge T_1} \mid t_1 < \tau_\partial) + x\mathbb{P}_1(\tau_\partial \leq t_1), \end{aligned}$$

and so

$$\mathbb{E}_x(X_{t_1 \wedge T_1} \mid t_1 < \tau_\partial) \geq c > 0.$$

This implies that, conditionally on non-extinction, X hits some $\varepsilon > 0$ before t_1 with a positive probability.

Comments

- The quasi-stationary behavior of 1D diffusion has been extensively studied in the literature (Cattiaux et al., 2009, Kolb and Steinsaltz, 2012, Littin, 2012, Miura, 2014). We do not improve much the known conditions for existence and uniqueness of the QSD, but we improve the convergence, since we prove that it is uniform and exponential w.r.t. the TV norm.
- Again, it is possible to extend our method to models with killing, or to cases with jump.

Multi-dimensional diffusions [C, Coulibaly-Pasquier, V, in prep.]

We can also consider solutions to an SDE in a bounded domain $D \subset \mathbb{R}^d$ absorbed at the boundary of D, or diffusions on a compact Riemannian manifold (cf. Cattiaux and Méléard, 2010 and Knobloch and Partzsch, 2010)

Method for (A1): the process goes away from the boundary when it is not-absorbed: comparison/coupling of the distance to the boundary to use the results of dimension 1.

Methods for (A2):

- estimates of the gradient of the semi-group [Wang, 2004] : $\|\nabla P_t f\|_\infty \leq c_t \|P_{t-1} f\|_\infty$
- two-sided estimates on the density of the absorbed process [Lierl, Saloff-Coste, 2014]

Concluding remarks

- We obtained explicit conditions ensuring the existence of a universal QLD, which is the only case where the QSD can be expected to be observed regardless of the (unknown) initial distribution.
- We showed how these conditions can be checked in classical models of population dynamics.
- The method we use is flexible enough to cover easily basic extensions of the model, including processes with killing, jumps...

Open question/work in progress:

- Improve conditions for multidimensional BDP absorbed at the boundary (require new methods)
- "Simple" criteria expressed in terms of Lyapunov functions (like those used in [Collet et al., 2011])
- Criteria for the convergence of conditional distributions non-uniform w.r.t. the initial distribution.