# Time scales for large populations birth and death processes - Quasi stationary distributions and resilience

Sylvie Méléard, Ecole Polytechnique, France

Les probabilités de demain, Paris, 2019

With J.R. Chazottes, P. Collet and S. Martinez for the last part.

# Population dynamics

- The dynamics of d interacting species is often modeled by a vector field $B(x) - D(x)$ on $\mathbb{R}_+^d$:

$$\frac{dx}{dt} = B(x) - D(x),$$

  $x_j$ representing the concentration of species $j$, $(1 \leq j \leq d)$.

- $B(x)$ and $D(x)$ have non-negative components.
- $B(0) = D(0) = 0$: absence of spontaneous generation (no immigration).
- Link with individuals: a parameter $K > 0$ (large) - the order of magnitude of the number of individuals compatible with the available resources.
- If $n_j$ is the number of individuals of species $j$, then

$$x_j = \frac{n_j}{K}.$$

# Birth and Death Process

The microscopic time evolution is given by the birth and death process $(N_t^K, t \geq 0)$ on $\mathbb{Z}_+^d$ with transition rates $K\,B(n/K)$ and $K\,D(n/K)$.

$$\mathbb{P}(N_j^K(t{+}dt) = n_j{+}1, N_q^K(t{+}dt) = n_q, \forall q \neq j \mid N(t) = n) = K\,B_j(n/K)\,dt\,;$$

$$\mathbb{P}(N_j^K(t{+}dt) = n_j{-}1, N_q^K(t{+}dt) = n_q, \forall q \neq j \mid N(t) = n) = K\,D_j(n/K)\,dt\,.$$

$$\mathbb{P}(N^K(t+dt) = n \mid N(t) = n) = 1 - K\sum_{j=1}^{d}\big(B_j(n/K) + D_j(n/K)\big)\,dt.$$

- For example if $B_j(x) = \lambda_j x_j$, we have

$$\mathbb{P}(N_j^K(t+dt) = n_j + 1 \mid N(t) = n) = \lambda_j\,n_j\,dt.$$

# Time scale of order one

**Theorem** (Kurtz '71) *For any* $T > 0$*, for any* $x_0 \in \mathbb{R}^d_+$*, if* $\lim_K \frac{N^K(0)}{K} = x_0$*, then for any* $\varepsilon > 0$*,*

$$\lim_K \mathbb{P}\Big(\sup_{t\leq T}\Big|\frac{N^K(t)}{K} - x(t)\Big| > \varepsilon\Big) = 0,$$

*with* $dx/dt = B(x) - D(x)$ *and* $X(0) = x_0$.

• On the finite time interval $[0, T]$, the trajectory of the process $N^K/K$ stays close to the trajectory of the diffusion process $Z$ given by

$$dZ_j(t) = (B_j(Z) - D_j(Z))dt + \frac{\sqrt{B_j(Z) + D_j(Z)}}{\sqrt{K}}dW_t,$$

where $W$ is a Brownian motion.

*This is the standard stochastic fluctuation.*

# Assumptions on $B$ and $D$

- $B(0) = D(0) = 0$: 0 *is an absorbing point.*
- *B and D are smooth.*
- *"Descent from infinity":*

$$\lim_{x \to \infty} \frac{\sup_j B_j(x)}{\inf_j D_j(x)} = 0.$$

- 0 *is a repeller*
- *There exists a unique positive fixed point* $x_*$ *for* $B - D$, *lying in* $Int(\mathbb{R}^d_+)$, *linearly stable and globally attracting.*

**Standard assumptions in ecology**: logistic birth-and-death process

$B(x) = b\,x$ ; $D(x) = x(d + cx)$ and $\frac{dx}{dt} = x(b - d - cx)$.

# What about very large time scale?

Under our assumptions, the process $(N_t^K, t \geq 0)$ attains 0 almost surely in finite time.

Let $T_0 = \inf\{t > 0; N_t^K = 0\}$ be the extinction time.

$$\forall n \in \mathbb{N}^d \setminus \{0\}, \quad \mathbb{P}_n(T_0 < \infty) = 1.$$

- From Kurtz's Theorem, $N_t^K$ should be close to $[x_* K]$ for large $t$.
- Then the limits in $t$ and $K$ cannot be interchanged.

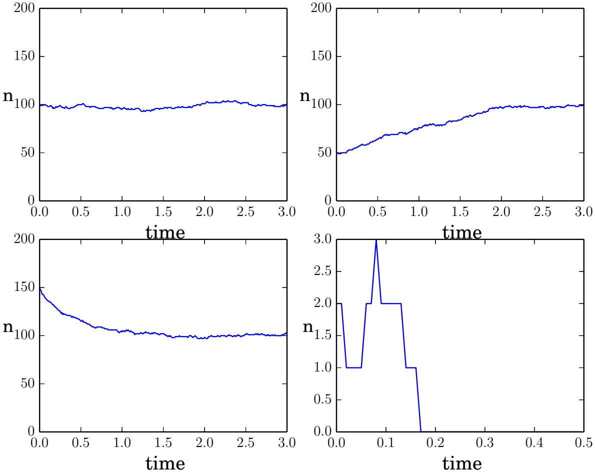
What happens in a larger time scale?

How long does it take for the process to reach 0?

What is the time scale of $T_0$?

# Trajectories of $N_t^K$

$d = 1$, $K = 100$.



# Quasi-stationary distribution

**Theorem** (Van Doorn '91)
*For fixed $K$ , there exists a unique probability measure $\nu^K$ on $\mathbb{N}^d \setminus \{0\}$ such that*

$$\mathbb{P}_{\nu^K}\left(N_t^K \in A \,\middle|\, T_0 > t\right) = \nu^K(A) \qquad \forall t > 0,\ A \subset \mathbb{N}^d \setminus \{0\}.$$

Moreover, for all $n \in \mathbb{N}^d \setminus \{0\}$, we have

$$\lim_{t \to \infty} \mathbb{P}_n\left(N_t^K \in A \,\middle|\, T_0 > t\right) = \nu^K(A)\ .$$

$\nu^K$ is called a quasi-stationary distribution (QSD).

Large literature on the topics, in particular Cattiaux et al. '09, M.-Villemonais '12, Collet-Martinez-San Martin '13, Champagnat-Villemonais '16.

One can show that there exists $\rho_0(K) > 0$, extinction rate from the QSD $\nu^K$, such that for all $t > 0$

$$\mathbb{P}_{\nu^K}\big(T_0 > t\big) = e^{-\rho_0(K)\,t}\;.$$

In particular,

$$\mathbb{E}_{\nu^K}(T_0) = \frac{1}{\rho_0(K)}.$$

Can we obtain the exact dependence of $\rho_0$ as function of $K$, for large $K$?

How do the trajectories behave, for large $K$ ?

Can we see the QSD?

Which information can we deduce from the observation of the process?

• If the time it takes for the process to reach the QSD is significantly less than $1/\rho_0(K)$, we can see the QSD.

• We will prove that there exists another time scale $\frac{1}{\rho_1(K)} \ll \frac{1}{\rho_0(K)}$ which describes the time it takes to reach the QSD.

The problem is generically not self-adjoint (except in dimension 1).

*A necessary and sufficient condition for the existence and uniqueness of a QSD together with the convergence in total variation is proved by N. Champagnat and D. Villemonais, 2016*.

They provide in particular an estimate for the rate of convergence (spectral gap).

$$\sup_{n\in\mathbb{N}^d\backslash\{0\}} \left\|\mathbb{P}_n\left(N_t^K \in \cdot \mid T_0 > t\right) - \nu^K\right\|_{TV} \leq 2\left(1 - b_1\, b_2\right)^{t/t_0} .$$

They require two conditions.

**Condition A1**: There exist two positive numbers $b_1$ and $t_0$ and a probability measure $\theta_K$ on $\mathbb{N}^d\backslash\{0\}$ such that for any subset $A$ of $\mathbb{N}^d\backslash\{0\}$

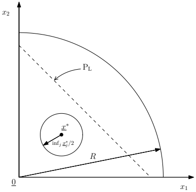
$$\inf_{n\in\mathbb{N}^d\backslash\{0\}} \mathbb{P}_n\left(N_{t_0}^K \in A \,\middle|\, T_0 > t_0\right) \geq b_1\, \theta_K(A) .$$

Note that in general $\theta_K$ is not the QSD.

In our case, we choose the uniform distribution on $\mathcal{B}(Kx_*, \sqrt{K})$.

**Condition A2**: There exists a positive number $b_2$ such that

$$\mathbb{P}_{\theta_K}\big(T_0 > t\big) \geq b_2 \sup_{n \in \mathbb{N}^d \setminus \{0\}} \mathbb{P}_n\big(T_0 > t\big).$$




We have proven that for large $K$ the constants $b_1$ and $b_2$ can be chosen independent of $K$ while

$$t_0 = \mathcal{O}(1)_d \, \log K \, .$$

*The proof relies on descent from infinity, Lyapounov function and lower bounds on transition probabilities (no symmetry, no Harnack inequality available, no Gaussian bound known).*

We obtain that

$$\frac{1}{\rho_0(K)} = e^{\mathcal{O}(1)\,K},$$

with a very precise estimate for $d = 1$.

For the convergence rate, we get for some $a > 0$ independent of $K$, for all $n$,

$$\left\| \mathbb{P}_n\big(N_t^K \in \cdot\,\big) - \nu^K(\cdot) \right\|_{TV} \leq 2\, e^{-a\,t/\log K} + \mathbb{P}_n(T_0 \leq t).$$

We can prove that for some $b > 0$, $c > 0$, $f > 0$ and $D > 0$,

$$\mathbb{P}_n(T_0 \leq t) \leq e^{-b(\|n\|_1 \wedge (c\,K))} + t\,D\,e^{-f\,K}.$$

Therefore

$$\left\| \mathbb{P}_n\big(N_t^K \in \cdot\,\big) - \nu^K(\cdot) \right\|_{TV} \leq 2\, e^{-a\,t/\log K} + e^{-b(\|n\|_1 \wedge (c\,K))} + t\,D\,e^{-f\,K}.$$

This error estimate $2\, e^{-a\,t/\log K} + e^{-b(\|n\|_1 \wedge (c\,K))} + t\,D\,e^{-f\,K}$ reflects what we saw in the simulations.

- If the starting point $n$ is of order one, the error is not small and the population can disappear in a time of order one.
- If the starting point $n$ is of order $K$, the error decreases with time at an exponential rate of order $1/\log K$ and becomes small (for large $K$).
- If $t \approx e^{f\,K}$, the error becomes large again.
- Hence if $\log K \ll t \ll e^{f\,K}$, the distribution of $N^K(t)$ is very near to $\nu^K$ (for a starting point of order $K$).

  Note the huge difference of time scales between $\log K$ (rate of convergence to $\nu^K$), and $e^{f\,K}$ (lower bound on the time scale of extinction), if $K$ is large.

# Key Properties

Let $S_t^K$ be the semigroup of $N^K$. Then there exists $C$ independent of $K$ s.t.

$$\sup_{n\in\mathbb{N}^d} S_1^K(e^{\|\cdot\|})(n) \leq e^{C\,K}.$$

In particular, $S_1^K$ maps polynomially growing functions to bounded functions and is a compact operator in such Banach spaces.

For the QSD, we have

- Exponential moments:

$$\nu^K(e^{\|n\|}) \leq e^{\mathcal{O}(1)\,K}.$$

- $\nu^K(n) = K\,x_* + \mathcal{O}(1)$.
- For $\ell \in \mathbb{N}$, there exist $C_\ell > 0$ and $C' > 0$ such that for all $K \geq 1$,

$$\nu^K(\|n - \nu^K(n)\|^{2\ell}) \leq C_\ell\, K^\ell \quad ; \quad \nu^K(\|n - \nu^K(n)\|^2) \geq C'\, K,$$

- There is a Gaussian approximation of $\nu^K$ near $\nu^K(n)$ with variance of order $K$.

# Properties of the QSD for $d = 1$

For $K$ large enough,

$$\rho_0(K) = \left(a + \mathcal{O}\left(\frac{(\log K)^3}{\sqrt{K}}\right)\right) \sqrt{K}\, e^{-b\,K},$$

$$a = \frac{1}{\sqrt{2}\,\pi} \left(\sqrt{\frac{B'(0)}{D'(0)}} - \sqrt{\frac{D'(0)}{B'(0)}}\right) \sqrt{\frac{D'(x_*)}{D(x_*)} - \frac{B'(x_*)}{B(x_*)}}\; x_*\; B(x_*)\;,$$

$$b = \int_0^{x_*} \frac{B(x)}{D(x)}\, dx\;.$$

We have

$$\sup_{n\in\mathbb{N}^d\setminus\{0\}} \left\| \mathbb{P}_n(N_t^K \in \cdot\,) \;-\; \alpha_n(K)\,\nu^K + (1-\alpha_n(K))\,\delta_0 \right\|_{TV} \leq \mathcal{O}(1)\times$$

$$\left(\sqrt{K}\,\log K\, e^{-c\,K} + (1 - e^{-\rho_0(K)t}) + K e^{-d\,t/4} + K^{3/4} e^{\ell\,K} e^{-\rho_1(K)t}\right)$$

for $c, d, \ell$ positive constants independent of $K$ and

$$\alpha_n(K) = 1 - \left(\frac{D'(0)}{B'(0)}\right)^n + \frac{\mathcal{O}(1)}{K} \quad , \quad \rho_1(K) \geq \frac{\mathcal{O}(1)}{\log K}.$$

# Resilience

Back to the dynamical system. Let

$$M = J(B - D)(x_*).$$

The engineering resilience is defined by

$$\mathcal{R} = - \sup_{z \in Sp(M)} \mathcal{R}e(z) > 0.$$

*Engineering resilience is useful for at least two major purposes:*

1) It gives the exponential rate of relaxation to the equilibrium after a (small) perturbation. Large resilience means more stability.

2) It gives an estimation of the change of the equilibrium after a (small) perturbation of the system.

# How to determine the resilience?

Can one measure the resilience just by observing and recording the time dynamics of the system?

We prove the relation

$$M\Sigma^K + \Sigma^K M^t + 2\mathcal{D}^K = \mathcal{O}(\sqrt{K}),$$

where $\mathcal{D}^K$ is the diagonal matrix with entries

$$\mathcal{D}^K_{ii} = KD_i(x_*) = KB_i(x_*)$$

and $\Sigma^K$ is the covariance matrix

$$\Sigma^K_{i,j} = \int (n_i - \mu^K_i)(n_j - \mu^K_j)\,\nu^K(dn) \quad ; \quad \mu^K = \int n\,\nu^K(dn) = \nu^K(n).$$

Given a trajectory $(N^K(t), t \leq T)$, one can estimate $\Sigma^K$ and $\mathcal{D}^K$ .

$\mathbf{d = 1}$:

$$\mathcal{R} = \frac{\mathcal{D}^K}{\Sigma^K} \qquad \text{up to } \frac{1}{\sqrt{K}}.$$

# Case $\mathbf{d} > 1$

The equation $M\Sigma^K + \Sigma^K M^t + 2\mathcal{D}^K = 0$ has many solutions for $M$ (which generically is not symmetric).

One uses the time correlations. Define for $\tau > 0$

$$\Sigma^K_{i,j}(\tau) = \mathbb{E}_{\nu^K}\left((N^K_i(\tau) - \mu^K_i)(N^K_j(0) - \mu^K_j)\right).$$

Note that $\Sigma^K(0) = \Sigma^K$.

One can prove that

$$e^{\tau M} = \Sigma^K(\tau)\,\Sigma^K(0)^{-1} + \mathcal{O}(1)(1/\sqrt{K}).$$

The matrices $\Sigma^K(\tau)$ and $\Sigma^K$ can be estimated from the data $(N^K(t), 0 \leq t \leq T)$, and choosing for example $\tau = 1$, one can estimate the matrix $M$ and hence the resilience.

# Statistics

One can introduce statistics to estimate the various quantities of interest from the data. For $T > 0$, let

$$S_i^{\mu}(T,K) = \frac{1}{T}\int_0^T N_i^K(s)ds,$$

$$S_{i,j}^{\Sigma}(T,K) = \frac{1}{T}\int_0^T (N_i^K(s) - S_i^{\mu}(T,K))(N_j^K(s) - S_j^{\mu}(T,K))ds,$$

$$S_i^{\mathcal{D}}(T,K) = \frac{1}{T}\#\left\{\text{birth of species i for } t \in [0,T]\right\}$$

$$S_{i,j}^{C}(T,\tau,K) = \frac{1}{T-\tau}\int_0^{T-\tau} (N_i^K(s+\tau) - S_i^{\mu}(T,K))(N_j^K(s) - S_j^{\mu}(T,K))ds$$

# Rates of convergence of the statistics

The errors in the inferences depend on $T$ and on the starting point.

We have estimates for the $L^2$-distance between each of the above statistics and the quantities to infer, starting from an initial condition or in the QSD.

For example there exist $C > 0, a, b, c, d > 0$ such that for all $K > 2$, for all $n$,

$$\mathbb{E}_n\left(\|S^\mu(T,K) - \nu^K(n)\|^2\right) \leq C(K^2 + \|n\|^2)\left(\frac{\log K}{T} + Te^{-bK} + e^{-c(\|n\| \wedge (dK))}\right)$$

$$\mathbb{E}_{\nu^K}\left(\|S^\mu(T,K) - \nu^K(n)\|^2\right) \leq C\left(\frac{K^2 \log K}{T} + K^2(1+T)e^{-aK}\right)$$

The last inequality has an interest only if $K^2 \log K \ll T \ll e^{aK}$.

# Thank you for your attention!