Uniform convergence to the Q-process

Nicolas Champagnat, Denis Villemonais

Uniform convergence to the Q-process

Nicolas Champagnat[1,2,3], Denis Villemonais[1,2,3]

April 7, 2017

Abstract

The first aim of the present note is to quantify the speed of convergence of a conditioned process toward its Q-process under suitable assumptions on the quasi-stationary distribution of the process. Conversely, we prove that, if a conditioned process converges uniformly to a conservative Markov process which is itself ergodic, then it admits a unique quasi-stationary distribution and converges toward it exponentially fast, uniformly in its initial distribution. As an application, we provide a conditional ergodic theorem.

Keywords: quasi-stationary distribution; Q-process; uniform exponential mixing property; conditional ergodic theorem

2010 Mathematics Subject Classification. 60J25; 37A25; 60B10.

1 Introduction

Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E\cup\{\partial\}})$ be a time homogeneous Markov process with state space $E\cup\{\partial\}$, where E is a measurable space. We assume that $\partial\notin E$ is an absorbing state for the process, which means that $X_s=\partial$ implies $X_t=\partial$ for all $t\geq s$, $\mathbb{P}_x$-almost surely for all $x\in E$. In particular,

$$\tau_\partial := \inf\{t\geq 0, X_t=\partial\}$$

is a stopping time. We also assume that $\mathbb{P}_x(\tau_\partial<\infty)=1$ and $\mathbb{P}_x(t<\tau_\partial)>0$ for all $t\geq 0$ and $\forall x\in E$.

[1]Université de Lorraine, IECL, UMR 7502, Campus Scientifique, B.P. 70239, Vandœuvre-lès-Nancy Cedex, F-54506, France
[2]CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54506, France
[3]Inria, TOSCA team, Villers-lès-Nancy, F-54600, France.
E-mail: Nicolas.Champagnat@inria.fr, Denis.Villemonais@univ-lorraine.fr

A probability measure α on E is called a *quasi-stationary distribution* if

$$\mathbb{P}_\alpha(X_t \in \cdot \mid t < \tau_\partial) = \alpha, \quad \forall t \geq 0.$$

We refer the reader to [7, 9, 4] and references therein for extensive developments and several references on the subject. It is well known that a probability measure α is a quasi-stationary distribution if and only if there exists a probability measure μ on E such that

$$\lim_{t\to+\infty} \mathbb{P}_\mu(X_t \in A \mid t < \tau_\partial) = \alpha(A) \tag{1.1}$$

for all measurable subsets A of E.

In [2], we provided a necessary and sufficient condition on X for the existence of a probability measure α on E and constants $C, \gamma > 0$ such that

$$\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \alpha\|_{TV} \leq C e^{-\gamma t}, \quad \forall \mu \in \mathcal{P}(E), \quad t \geq 0, \tag{1.2}$$

where $\|\cdot\|_{TV}$ is the total variation norm and $\mathcal{P}(E)$ is the set of probability measures on E. This immediately implies that α is the unique quasi-stationary distribution of X and that (1.1) holds for any initial probability measure μ.

The necessary and sufficient condition for (1.2) is given by the existence of a probability measure ν on E and of constants $t_0, c_1, c_2 > 0$ such that

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

and

$$\mathbb{P}_\nu(t < \tau_\partial) \geq c_2 \mathbb{P}_x(t < \tau_\partial), \quad \forall t \geq 0,\ x \in E.$$

The first condition implies that, in cases of unbounded state space E (like $\mathbb{N}$ or $\mathbb{R}_+$), the process $(X_t, t \geq 0)$ comes down from infinity in the sense that, there exists a compact set $K \subset E$ such that $\inf_{x\in E} \mathbb{P}_x(X_{t_0} \in K \mid t_0 \tau_\partial) > 0$. This property is standard for biological population processes such as Lotka-Volterra birth and death or diffusion processes [1, 3]. However, this is not the case for some classical models, such as linear birth and death processes or Ornstein-Uhlenbeck processes.

Many properties can be deduced from (1.2). For instance, this implies the existence of a constant $\lambda_0 > 0$ such that

$$\mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda_0 t}$$

and of a function $\eta : E \to (0, \infty)$ such that $\alpha(\eta) = 1$ and

$$\lim_{t\to+\infty} \sup_{x\in E} \left| e^{\lambda_0 t} \mathbb{P}_x(t < \tau_\partial) - \eta(x) \right| = 0 \tag{1.3}$$

as proved in [2, Prop. 2.3]. It also implies the existence and the exponential ergodicity of the associated Q-process, defined as the process X conditioned to never be extinct [2, Thm. 3.1]. More precisely, if (1.2) holds, then the family $(\mathbb{Q}_x)_{x\in E}$ of probability measures on Ω defined by

$$\mathbb{Q}_x(\Gamma) = \lim_{t\to+\infty} \mathbb{P}_x(\Gamma \mid t < \tau_\partial),\ \forall \Gamma \in \mathcal{F}_s,\ \forall s \geq 0, \tag{1.4}$$

is well defined and the process $(\Omega, (\mathcal{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathbb{Q}_x)_{x\in E})$ is an E-valued homogeneous Markov process. In addition, this process admits the unique invariant probability measure (sometimes refered to as the doubly limiting quasi-stationary distribution [5])

$$\beta(dx) = \eta(x)\alpha(dx)$$

and there exist constants $C', \gamma' > 0$ such that, for any $x \in E$ and all $t \geq 0$,

$$\|\mathbb{Q}_x(X_t \in \cdot) - \beta\|_{TV} \leq C' e^{-\gamma' t}. \tag{1.5}$$

The measure β

The first aim of the present note is to refine some results of [2] in order to get sharper bounds on the convergence in (1.3) and to prove that the convergence (1.4) holds in total variation norm, with uniform bounds over the initial distribution (see Theorem 2.1). Using these new results, we obtain in Corollary 2.3 that the uniform exponential convergence (1.2) implies that, for all bounded measurable function $f : E \to \mathbb{R}$ and all $T > 0$,

$$\left| \mathbb{E}_x \left(\frac{1}{T} \int_0^T f(X_t)\, dt \mid T < \tau_\partial \right) - \int_E f\, d\beta \right| \leq \frac{a\|f\|_\infty}{T}, \tag{1.6}$$

for some positive constant a. This result improves the very recent result obtained independently by He, Zhang and Zu [6, Thm. 2.1] by providing the convergence estimate in $1/T$. The interested reader might look into [6] for nice domination properties between the quasi-stationary distribution α and the probability β.

The second aim of this note is to prove that the existence of the Q-process with uniform bounds in (1.4) and its uniform exponential ergodicity (1.5) form in fact a necessary and sufficient condition for the uniform exponential convergence (1.2) toward a unique quasi-stationary distribution.

2 Main results

In this first result, we improve (1.3) and provide a uniform exponential bound for the convergence (1.4) of the conditioned process toward the Q-process.

Theorem 2.1. *Assume that (1.2) holds. Then there exists a positive constant a_1 such that*

$$\left| e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial) - \eta(x) \right| \leq a_1\, e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial) e^{-\gamma t}, \tag{2.1}$$

where λ_0 and η are the constant and function appearing in (1.3) and where $\gamma > 0$ is the constant from (1.2).

Moreover, there exists a positive constant a_2 such that, for all $t \geq 0$, for all $\Gamma \in \mathcal{F}_t$ and all $T \geq t$,

$$\|\mathbb{Q}_x(\Gamma) - \mathbb{P}_x(\Gamma \mid T < \tau_\partial)\|_{TV} \leq a_2\, e^{-\gamma(T-t)}, \tag{2.2}$$

where $(\mathbb{Q}_x)_{x\in E}$ is the Q-process defined in (1.4).

We emphasize that (2.1) is an improvement of (1.3), since the convergence is actually exponential and, in many interesting examples, $\inf_{x\in E}\mathbb{P}_x(t < \tau_\partial) = 0$. This is for example the case for elliptic diffusion processes absorbed at the boundaries of an interval, since the probability of absorption converges to 1 when the initial condition converges to the boundaries of the interval. The last theorem has a first corollary.

Corollary 2.2. *Assume that (1.2) holds. Then there exists a positive constant a_3 such that, for all $T > 0$, all probability measure μ_T on $[0, T]$ and all bounded measurable functions $f : E \to \mathbb{R}$,*

$$\begin{aligned} &\left| \mathbb{E}_x\left(\int_0^T f(X_t)\mu_T(dt) \mid T < \tau_\partial \right) - \int_E f\, d\beta \right| \\ &\qquad\qquad \leq a_3\|f\|_\infty \int_0^T \left(e^{-\gamma' t} + e^{-\gamma(T-t)} \right) \mu_T(dt). \end{aligned} \tag{2.3}$$

This follows from (2.2), the exponential ergodicity of the Q-process stated in (1.5) and the inequality

$$\begin{aligned} &\left| \mathbb{E}_x\left(\int_0^T f(X_t)\mu_T(dt) \mid T < \tau_\partial \right) - \int_E f\, d\beta \right| \\ &\qquad \leq \int_0^T \left| \mathbb{E}_x(f(X_t) \mid T < \tau_\partial) - \mathbb{E}^{\mathbb{Q}_x}(f(X_t)) \right| \mu_T(dt) \\ &\qquad\qquad + \int_0^T \left| \mathbb{E}^{\mathbb{Q}_x}(f(X_t)) - \int_E f\, d\beta \right| \mu_T(dt), \end{aligned}$$

where $\mathbb{E}^{\mathbb{Q}_x}$ is the expectation with respect to $\mathbb{Q}_x$.

In particular, choosing μ_T as the uniform distribution on $[0, T]$, we obtain a conditional ergodic theorem.

Corollary 2.3. *Assume that (1.2) holds. Then there exists a positive constant a_4 such that, for all $T > 0$ and all bounded measurable functions $f : E \to \mathbb{R}$,*

$$\left|\mathbb{E}_x\left(\frac{1}{T}\int_0^T f(X_t)\,dt \mid T < \tau_\partial\right) - \int_E f\,d\beta\right| \leq \frac{a_4\,\|f\|_\infty}{T}.$$

Considering the problem of estimating β from N realizations of the unconditioned process X, one wishes to take T as small as possible in order to obtain the most samples such that $T < \tau_\partial$ (of order $N_T = Ne^{-\lambda_0 T}$). It is therefore important to minimize the error in (2.3) for a given T. It is easy to check that $\mu_T = \delta_{t_0}$ with $t_0 = \gamma T/(\gamma+\gamma')$ is optimal with an error of the order of $\exp(-\gamma'\gamma T/(\gamma+\gamma'))$. Combining this with the Monte Carlo error of order $1/\sqrt{N_T}$, we obtain a global error of order

$$\frac{e^{\lambda_0 T/2}}{\sqrt{N}} + e^{-\gamma\gamma' T/(\gamma+\gamma')}.$$

In particular, for a fixed N, the optimal choice for T is $T \approx \frac{\log N}{\lambda_0 + 2\gamma\gamma'/(\gamma+\gamma')}$ and the error is of the order of $N^{-\zeta}$ with $\zeta = \frac{\gamma\gamma'}{2\gamma\gamma' + \lambda_0(\gamma+\gamma')}$. Conversely, for a fixed T, the best choice for N is $N \approx \exp((\lambda_0 + 2\gamma\gamma'/(\gamma+\gamma'))T)$ and the error is of the order of $\exp(-\gamma\gamma' T/(\gamma+\gamma'))$.

We conclude this section with a converse to Theorem 2.1. More precisely, we give a converse to the fact that (1.2) implies both (1.5) and (2.2).

Theorem 2.4. *Assume that there exists a Markov process $(\mathbb{Q}_x)_{x\in E}$ with state space E such that, for all $t > 0$,*

$$\lim_{T\to+\infty} \sup_{x\in E} \|\mathbb{Q}_x(X_t \in \cdot) - \mathbb{P}_x(X_t \in \cdot \mid T < \tau_\partial)\|_{TV} = 0 \tag{2.4}$$

and such that

$$\lim_{t\to+\infty} \sup_{x,y\in E} \|\mathbb{Q}_x(X_t \in \cdot) - \mathbb{Q}_y(X_t \in \cdot)\|_{TV} = 0. \tag{2.5}$$

Then the process $(\mathbb{P}_x)_{x\in E}$ admits a unique quasi-stationary distribution α and there exist positive constants γ, C such that (1.2) holds.

It is well known that the strong ergodicity (2.5) of a Markov process implies its exponential ergodicity [8, Thm. 16.0.2]. Similarly, we observe in our situation that, if (2.4) and (2.5) hold, then the combination of the above results implies that both convergences hold exponentially.

3 Proofs

3.1 Proof of Theorem 2.1

For all $x \in E$, we set

$$\eta_t(x) = \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\alpha(t < \tau_\partial)} = e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial),$$

and we recall from [2, Prop. 2.3] that $\eta_t(x)$ is uniformly bounded w.r.t. $t \geq 0$ and $x \in E$. By Markov's property

$$\begin{aligned}\eta_{t+s}(x) &= e^{\lambda_0(t+s)}\mathbb{E}_x\left(\mathbb{1}_{t<\tau_\partial}\mathbb{P}_{X_t}(s < \tau_\partial)\right)\\ &= \eta_t(x)\mathbb{E}_x\left(\eta_s(X_t) \mid t < \tau_\partial\right).\end{aligned}$$

By (1.2), there exists a constant C' independent of s such that

$$\left|\mathbb{E}_x\left(\eta_s(X_t) \mid t < \tau_\partial\right) - \int_E \eta_s d\alpha\right| \leq C' e^{-\gamma t}.$$

Since $\int \eta_s d\alpha = 1$, there exists a constant $a_1 > 0$ such that, for all $x \in E$ and $s, t \geq 0$,

$$\left|\frac{\eta_{t+s}(x)}{\eta_t(x)} - 1\right| \leq a_1 e^{-\gamma t}.$$

Hence, multiplying on both side by $\eta_t(x)$ and letting s tend to infinity, we deduce from (1.3) that, for all $x \in E$,

$$|\eta(x) - \eta_t(x)| \leq a_1 e^{-\gamma t}\eta_t(x),\ \forall t \geq 0,$$

which is exactly (2.1). We also deduce that

$$\left(1 - a_1 e^{-\gamma t}\right)\eta_t(x) \leq \eta(x) \leq \left(1 + a_1 e^{-\gamma t}\right)\eta_t(x) \tag{3.1}$$

and hence, for t large enough,

$$\frac{\eta(x)}{1 + a_1 e^{-\gamma t}} \leq \eta_t(x) \leq \frac{\eta(x)}{1 - a_1 e^{-\gamma t}}. \tag{3.2}$$

Let us now prove the second part of Theorem 2.1. For any $t \geq 0$, $\Gamma \in \mathcal{F}_t$ and $0 \leq t \leq T$,

$$\begin{aligned}\mathbb{P}_x\left(\Gamma \mid T < \tau_\partial\right) &= \frac{\mathbb{P}_x\left(\Gamma \cap \{T < \tau_\partial\}\right)}{\mathbb{P}_x(T < \tau_\partial)}\\ &= \frac{e^{\lambda_0 T}\mathbb{P}_x\left(\Gamma \cap \{T < \tau_\partial\}\right)}{\eta(x)}\,\frac{\eta(x)}{e^{\lambda_0 T}\mathbb{P}_x(T < \tau_\partial)}.\end{aligned}$$

We deduce from (2.1) that

$$\left|\frac{\eta(x)}{e^{\lambda_0 T}\mathbb{P}_x(T<\tau_\partial)}-1\right|\leq a_1 e^{-\gamma T},$$

while, for all $T>\frac{\log a_1}{\gamma}$, (3.2) entails

$$\left|\frac{e^{\lambda_0 T}\mathbb{P}_x\left(\Gamma\cap\{T<\tau_\partial\}\right)}{\eta(x)}\right|\leq\frac{\eta_T(x)}{\eta(x)}\leq\frac{1}{1-a_1e^{-\gamma T}}.$$

Hence, for all $t\geq 0$ and all $T>\frac{\log a_1}{\gamma}$,

$$\left|\mathbb{P}_x\left(\Gamma\mid T<\tau_\partial\right)-\frac{e^{\lambda_0 T}\mathbb{P}_x\left(\Gamma\cap\{T<\tau_\partial\}\right)}{\eta(x)}\right|\leq\frac{a_1e^{-\gamma T}}{1-a_1e^{-\gamma T}}. \tag{3.3}$$

Now, the Markov property implies that

$$\mathbb{P}_x\left(\Gamma\cap\{T<\tau_\partial\}\right)=\mathbb{E}_x\left(\mathbb{1}_\Gamma\mathbb{P}_{X_t}(T-t<\tau_\partial)\right),$$

and we deduce from (3.3) that, for all $T>t+\frac{\log a_1}{\gamma}$,

$$\left|e^{\lambda_0(T-t)}\mathbb{P}_{X_t}(T-t<\tau_\partial)-\eta(X_t)\right|\leq\frac{a_1e^{-\gamma(T-t)}}{1-a_1e^{-\gamma(T-t)}}\eta(X_t).$$

Thus we have

$$\begin{aligned}
&\left|\frac{e^{\lambda_0 T}\mathbb{P}_x\left(\Gamma\cap\{T<\tau_\partial\}\right)}{\eta(x)}-\frac{e^{\lambda_0 t}\mathbb{E}_x\left(\mathbb{1}_\Gamma\eta(X_t)\right)}{\eta(x)}\right|\\
&\qquad\leq\frac{e^{\lambda_0 t}}{\eta(x)}\mathbb{E}_x\left[\mathbb{1}_\Gamma\left|e^{\lambda_0(T-t)}\mathbb{P}_{X_t}(T-t<\tau_\partial)-\eta(X_t)\right|\right]\\
&\qquad\leq\frac{a_1e^{-\gamma(T-t)}}{1-a_1e^{-\gamma(T-t)}}\frac{e^{\lambda_0 t}\mathbb{E}_x(\eta(X_t))}{\eta(x)}\\
&\qquad=\frac{a_1e^{-\gamma(T-t)}}{1-a_1e^{-\gamma(T-t)}},
\end{aligned}$$

where we used the fact that $\mathbb{E}_x\eta(X_h)=e^{-\lambda_0 h}\eta(x)$ for all $h>0$ (see [2, Prop. 2.3]). This and (3.3) allows us to conclude that, for all $t\geq 0$ and all $T>t+\frac{\log a_1}{\gamma}$,

$$\left|\mathbb{P}_x\left(\Gamma\mid T<\tau_\partial\right)-\frac{e^{\lambda_0 t}\mathbb{E}_x\left(\mathbb{1}_\Gamma\eta(X_t)\right)}{\eta(x)}\right|\leq\frac{2a_1e^{-\gamma T}}{1-a_1e^{-\gamma T}}.$$

Since $\mathbb{Q}_x(\Gamma)=e^{\lambda_0 t}\mathbb{E}_x\left(\mathbb{1}_\Gamma\,\eta(X_t)\right)/\eta(x)$ (see [2, Thm. 3.1 (ii)]), we deduce that (2.2) holds true.

This concludes the proof of Theorem 2.1.

3.2 Proof of Theorem 2.4

We deduce from (2.4) and (2.5) that there exists $t_1 > 0$ and $T_1 > 0$ such that, for all $T \geq T_1$,

$$\sup_{x,y \in E} \|\mathbb{P}_x(X_{t_1} \in \cdot \mid T < \tau_\partial) - \mathbb{P}_y(X_{t_1} \in \cdot \mid T < \tau_\partial)\|_{TV} \leq 1/2.$$

In particular, for all $s \geq 0$ and all $T \geq s + T_1$,

$$\sup_{x,y \in E} \left\|\delta_x R^T_{s,s+t_1} - \delta_y R^T_{s,s+t_1}\right\|_{TV} \leq 1/2, \tag{3.4}$$

where, for all $0 \leq s \leq t \leq T$, $R^T_{s,t}$ is the linear operator defined by

$$\begin{aligned}
\delta_x R^T_{s,t} f &= \mathbb{E}_x(f(X_{t-s}) \mid T - s < \tau_\partial) \\
&= \mathbb{E}(f(X_t) \mid X_s = x,\ T < \tau_\partial) \\
&= \delta_x R^{T-s}_{0,t-s} f,
\end{aligned}$$

where we used the Markov property. Now, for any $T > 0$, the family $(R^T_{s,t})_{0 \leq s \leq t \leq T}$ is a Markov semi-group. This semi-group property and the contraction (3.4) classically imply that, for all $T \geq T_1$,

$$\sup_{x,y \in E} \left\|\delta_x R^T_{0,T} - \delta_y R^T_{0,T}\right\|_{TV} \leq (1/2)^{\lfloor T - T_1 \rfloor / t_1}.$$

Then, proceeding as in [2, Section 5.1], we deduce that (1.2) holds true. This concludes the proof of Theorem 2.4.

References

[1] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969, 2009.

[2] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probability Theory and Related Fields*, 164(1):243–283, 2016.

[3] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *ArXiv e-prints*, Apr. 2017.

[4] P. Collet, S. Martínez, and J. Martín. *Quasi-Stationary Distributions: Markov Chains, Diffusions and Dynamical Systems*. Probability and Its Applications. Springer Berlin Heidelberg, 2012.

[5] D. C. Flaspohler. Quasi-stationary distributions for absorbing continuous-time denumerable Markov chains. *Ann. Inst. Statist. Math.*, 26:351–356, 1974.

[6] G. He, H. Zhang, and Y. Zhu. On the quasi-ergodic distribution of absorbing Markov processes. *ArXiv e-prints*, Nov. 2016.

[7] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.

[8] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 1993.

[9] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for discrete-state models. *European J. Oper. Res.*, 230(1):1–14, 2013.