Almost Sure Convergence of Measure Valued Pólya Processes

Denis Villemonais
Collaboration with Cécile Mailler

Université de Lorraine, Nancy, France

Neutron Transport Days of the Bath-Beijing-Paris Branching Structures Meeting, 17th September 2021

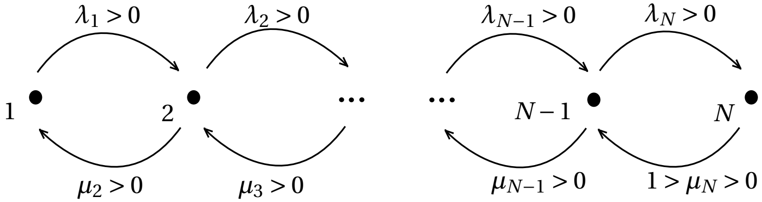
Almost Sure Convergence of Measure Valued Polya Processes

1. Quasi-stationary distributions
   a. A simple example
   b. Definition and a general result
   c. Application to the Neutron Random Walk
   d. An approximation scheme

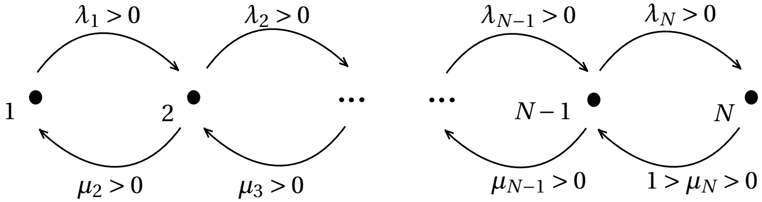
2. Measure valued Pólya processes
   a. Finite and irreducible Pólya urns
   b. Formal description of a Measure Valued Pólya Process
   c. Application to the approximation scheme

Almost Sure Convergence of Measure Valued Polya Processes

Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{1,2,\dots,N\}$ whose transition probabilities $(P_{ij})_{1\le i,j\le N}$ are given by



Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{1,2,\ldots,N\}$ whose transition probabilities $(P_{ij})_{1\le i,j\le N}$ are given by

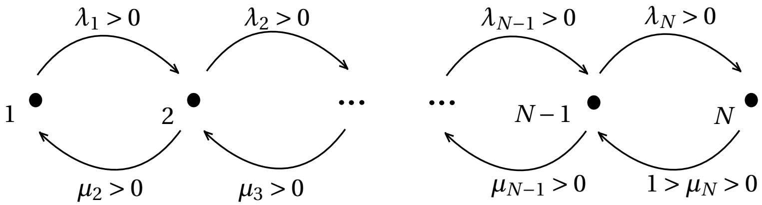


→ X admits a unique stationary distribution ν_s,

$$\nu_s P = \nu_s \quad \text{et} \quad \lim_{n\to+\infty} \mu P^n = \lim_{n\to+\infty} \mathbb{P}_\mu(X_n \in \cdot) = \nu_s,$$

for any initial distribution μ on $\{1,\ldots,N\}$.

Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{1,2,\ldots,N\}$ whose transition probabilities $(P_{ij})_{1\le i,j\le N}$ are given by



→ X admits a unique stationary distribution ν_s,

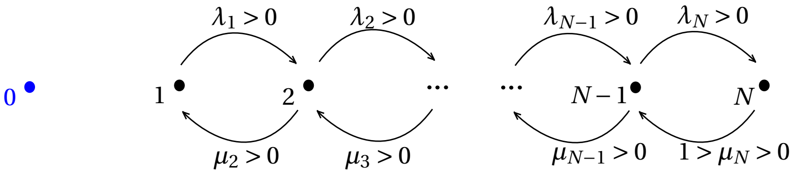
$$\nu_s P = \nu_s \quad \text{et} \quad \lim_{n\to+\infty} \mu P^n = \lim_{n\to+\infty} \mathbb{P}_\mu(X_n \in \cdot) = \nu_s,$$

for any initial distribution μ on $\{1,\ldots,N\}$.

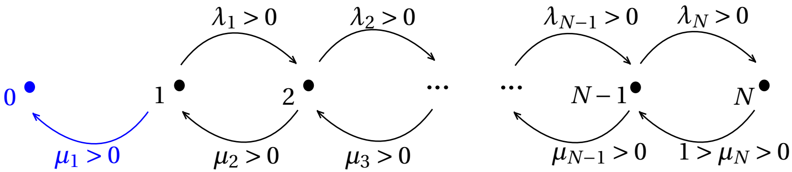
→ Well known extensions

- using spectral theory
- using coupling methods, irreducibility assumptions and so on...

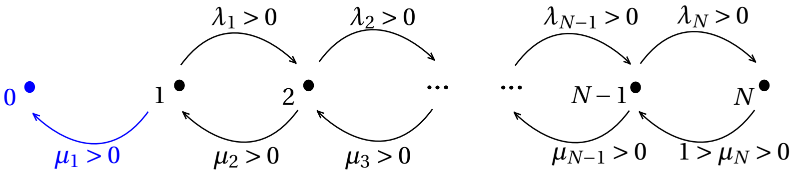
Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{0,1,2,\ldots,N\}$ whose transition probabilities $(Q_{ij})_{0\le i,j\le N}$ are given by



Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{0,1,2,\ldots,N\}$ whose transition probabilities $(Q_{ij})_{0\le i,j\le N}$ are given by



Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{0,1,2,\dots,N\}$ whose transition probabilities $(Q_{ij})_{0\le i,j\le N}$ are given by

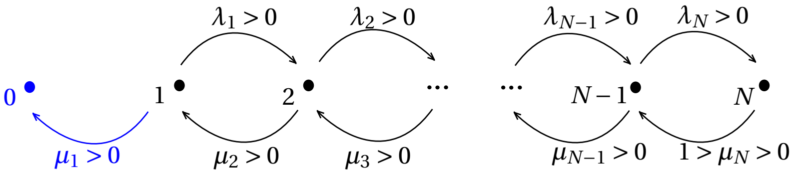


→ X admits a unique quasi-stationary distribution ν_{qs}, $P = Q_{|\{1,\dots,N\}^2}$

$$\nu_{qs}P = \theta_0 \nu_{qs} \quad \text{and} \quad \lim_{n\to+\infty} \theta_0^{-n}\mu P^n = \lim_{n\to+\infty} \theta_0^{-n}\mathbb{P}_\mu(X_n \in \cdot) = \mu(\eta)\,\nu_{qs},$$

for any initial distribution μ on $E := \{1,\dots,N\}$, $\theta_0 > 0$, $\eta : E \to (0,+\infty)$

Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{0,1,2,\dots,N\}$ whose transition probabilities $(Q_{ij})_{0\le i,j\le N}$ are given by



→ X admits a unique quasi-stationary distribution ν_{qs}, $P=Q_{|\{1,\dots,N\}^2}$

$$\nu_{qs}P=\theta_0\nu_{qs} \quad \text{and} \quad \lim_{n\to+\infty}\theta_0^{-n}\mu P^n=\lim_{n\to+\infty}\theta_0^{-n}\mathbb{P}_\mu(X_n\in\cdot)=\mu(\eta)\,\nu_{qs},$$

for any initial distribution μ on $E:=\{1,\dots,N\}$, $\theta_0>0$, $\eta:E\to(0,+\infty)$

→ It is a consequence of Perron-Frobenius theorem (Darroch Seneta 1965) and it implies that

$$\lim_{n\to+\infty}\left\|\mathbb{P}_\mu(X_n\in\cdot\mid X_n\neq 0)-\nu_{qs}\right\|_{TV}=0.$$

Almost Sure Convergence of Measure Valued Polya Processes

1. Quasi-stationary distributions
   a. A simple example
   b. Definition and a general result
   c. Application to the Neutron Random Walk
   d. An approximation scheme
2. Measure valued Pólya processes
   a. Finite and irreducible Pólya urns
   b. Formal description of a Measure Valued Pólya Process
   c. Application to the approximation scheme

Let Q be a positive kernel on E such that $Q_x(E) \le 1$ for all $x \in E$.

Consider $(X_n)_{n\ge 0}$ evolving in $E \cup \{\partial\}$, $\partial \notin E$, with transition kernel

$$\mathbb{P}_x(X_1 \in \cdot) = \delta_x Q + (1 - \delta_x Q(E))\delta_\partial \text{ and } \mathbb{P}_\partial(X_1 = \partial) = 1.$$

Setting $\tau_\partial = \inf\{n \ge 0,\ X_n = \partial\}$ the hitting time of ∂, we have

$$X_t = \partial,\ \forall\, t \ge \tau_\partial \text{ almost surely } (\textit{i.e. } \partial \text{ is absorbing}).$$

Let Q be a positive kernel on E such that $Q_x(E) \leq 1$ for all $x \in E$.

Consider $(X_n)_{n\geq 0}$ evolving in $E \cup \{\partial\}$, $\partial \notin E$, with transition kernel

$$\mathbb{P}_x(X_1 \in \cdot) = \delta_x Q + (1 - \delta_x Q(E))\delta_\partial \text{ and } \mathbb{P}_\partial(X_1 = \partial) = 1.$$

Setting $\tau_\partial = \inf\{n \geq 0,\ X_n = \partial\}$ the hitting time of ∂, we have

$$X_t = \partial,\ \forall t \geq \tau_\partial \text{ almost surely } (\textit{i.e. } \partial \text{ is absorbing}).$$

Définition

A quasi-stationary distribution (QSD) is a probability measure ν_{qs} on E such that

$$\nu_{qs} = \lim_{t\to\infty} \mathbb{P}_\mu\big(X_t \in \cdot \mid t < \tau_\partial\big)$$

for at least one initial distribution μ on E.

→ Surveys et book

- Méléard, V. 2012, Van Doorn, Pollett 2013
- Collet, Martínez, San Martín 2013

Assumption E. $\exists\ n_1 \in \mathbb{N}$, $\theta_1, \theta_2, c_1, c_2, c_3 > 0$, $\varphi_1, \varphi_2 : E \to \mathbb{R}_+$ and a probability measure ν on $K \subset E$ such that

→ (local Doblin) $\forall x \in K$,

$$\delta_x Q^{n_1} \geq c_1 \nu(\cdot \cap K) \quad \text{and} \quad \nu Q^n(E) \geq c_2 \delta_x Q^n(E),\ \forall n \in \mathbb{N}.$$

Assumption E. $\exists\ n_1 \in \mathbb{N}$, $\theta_1, \theta_2, c_1, c_2, c_3 > 0$, $\varphi_1, \varphi_2 : E \to \mathbb{R}_+$ and a probability measure ν on $K \subset E$ such that

→ (local Doblin) $\forall x \in K$,

$$\delta_x Q^{n_1} \geq c_1 \nu(\cdot \cap K) \quad \text{and} \quad \nu Q^n(E) \geq c_2 \delta_x Q^n(E),\ \forall n \in \mathbb{N}.$$

→ (Lyapunov) $\theta_1 < \theta_2$, $\varphi_1 \geq 1$, $\sup_K \varphi_1 < \infty$, $\inf_K \varphi_2 > 0$, $\varphi_2 \leq 1$,

$$\delta_x Q \varphi_1 \leq \theta_1 \varphi_1(x) + c_2 \mathbf{1}_K(x),\ \forall x \in E$$
$$\delta_x Q \varphi_2 \geq \theta_2 \varphi_2(x),\ \forall x \in E.$$

Assumption E. $\exists\ n_1 \in \mathbb{N},\ \theta_1,\theta_2,c_1,c_2,c_3 > 0,\ \varphi_1,\varphi_2 : E \to \mathbb{R}_+$ and a probability measure ν on $K \subset E$ such that

→ (local Doblin) $\forall x \in K$,

$$\delta_x Q^{n_1} \geq c_1 \nu(\cdot \cap K) \quad \text{and} \quad \nu Q^n(E) \geq c_2 \delta_x Q^n(E),\ \forall n \in \mathbb{N}.$$

→ (Lyapunov) $\theta_1 < \theta_2,\ \varphi_1 \geq 1,\ \sup_K \varphi_1 < \infty,\ \inf_K \varphi_2 > 0,\ \varphi_2 \leq 1,$

$$\delta_x Q \varphi_1 \leq \theta_1 \varphi_1(x) + c_2 \mathbf{1}_K(x),\ \forall x \in E$$
$$\delta_x Q \varphi_2 \geq \theta_2 \varphi_2(x),\ \forall x \in E.$$

Théorème (Champagnat, V. 2017+)

If Assumption E holds true, then $\exists \nu_{qs},\ \alpha \in (0,1), C > 0$ such that

$$\left|\mathbb{E}_\mu\left[f(X_n) \mid n < \tau_\partial\right] - \nu_{qs}(f)\right| \leq C\alpha^n \frac{\mu(\varphi_1)}{\mu(\varphi_2)},\ \forall n \in \mathbb{N},$$

for all μ and f such that $\mu(\varphi_1)/\mu(\varphi_2) < +\infty$ et $|f| \leq \varphi_1$.

1. The main point of the proof is to use Hairer and Mattingly [2011] to prove that

$$\left\| \delta_x S_{0,n_0 n}^{n_0 n} - \delta_y S_{0,n_0 n}^{n_0 n} \right\|_{TV} \le C\alpha^n (2 + \psi_{n_0 n}(x) + \psi_{n_0 n}(y))$$

for some time dependent Lyapunov function ψ_n and the semi-group

$$S_{n,n+1}^{n+m} f(x) = \mathbb{E}_{X_n = x}\left(f(X_{n+1}) \mid n + m < \tau_\partial \right)$$

1. The main point of the proof is to use Hairer and Mattingly [2011] to prove that

$$\left\| \delta_x S_{0,n_0 n}^{n_0 n} - \delta_y S_{0,n_0 n}^{n_0 n} \right\|_{TV} \leq C\alpha^n (2 + \psi_{n_0 n}(x) + \psi_{n_0 n}(y))$$

for some time dependent Lyapunov function ψ_n and the semi-group

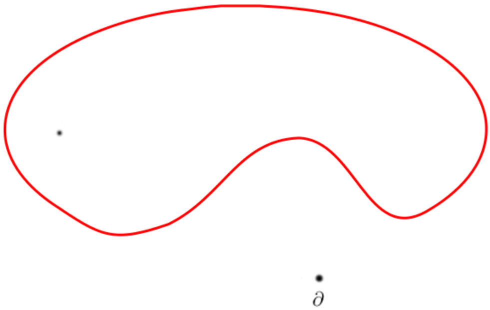
$$S_{n,n+1}^{n+m} f(x) = \mathbb{E}_{X_n = x}\left(f(X_{n+1}) \mid n+m < \tau_\partial\right)$$

2. In order to do so, we prove various estimates on $\mathbb{E}_x(\varphi_1(X_n))/\mathbb{E}_x(\varphi_2(X_n))$, on $\mathbb{P}_x(X_n \in K \mid n+m < \tau_\partial)$ and so on. For instance, one central estimate is, $\forall \theta \in (\theta_1/\theta_2, 1)$,

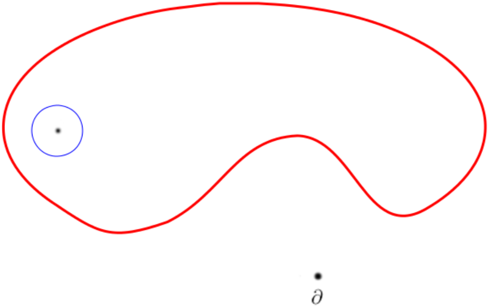
$$\mathbb{E}_x(\varphi_1(X_n) \mid n < \tau_\partial)) \leq \frac{\mathbb{E}_x(\varphi_1(X_n)\mathbf{1}_{n<\tau_\partial})}{\mathbb{E}_x(\varphi_2(X_n)\mathbf{1}_{n<\tau_\partial})} \leq \left(\theta^n \frac{\varphi_1(x)}{\varphi_2(x)}\right) \vee C_\theta.$$

Almost Sure Convergence of Measure Valued Polya Processes

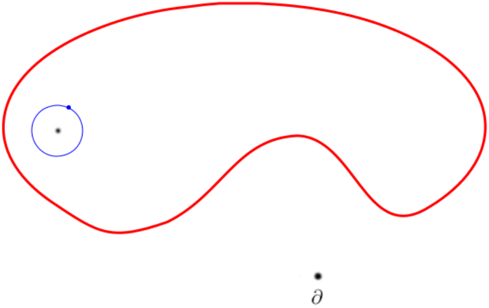
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



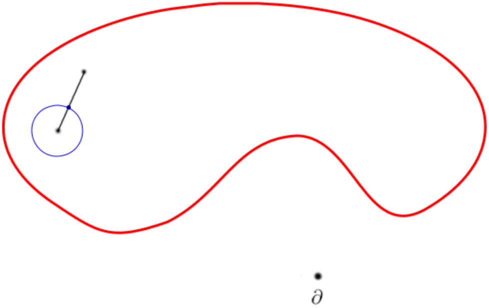
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



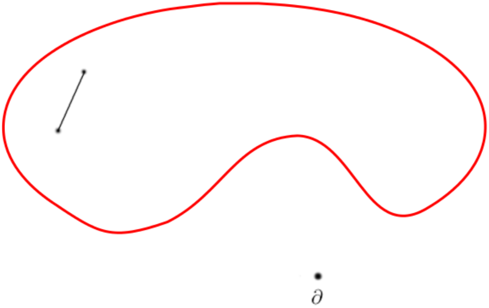
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



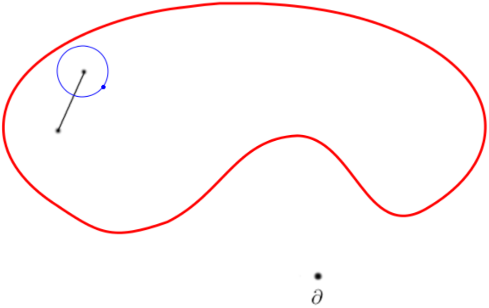
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



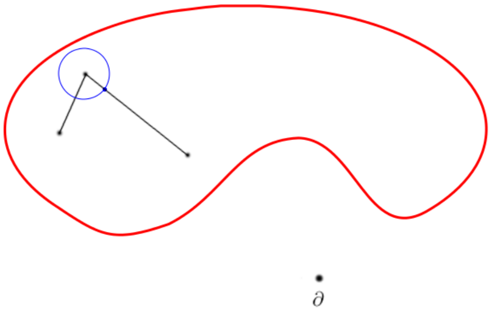
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



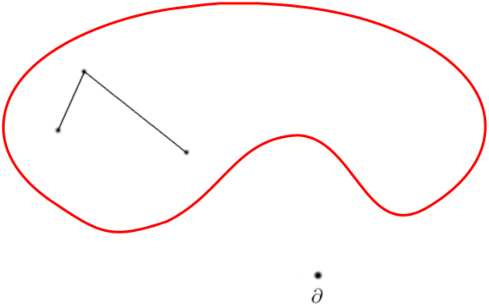
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



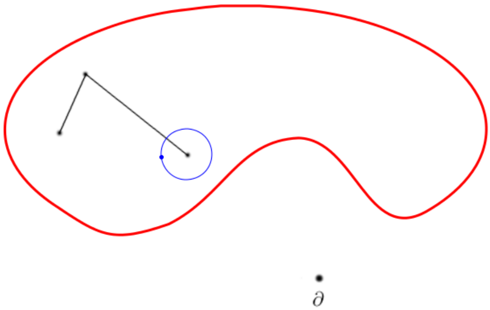
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



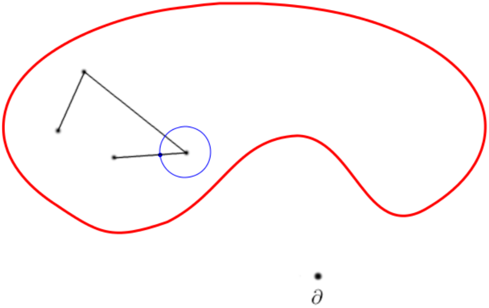
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



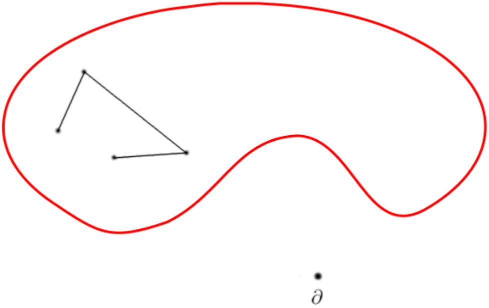
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



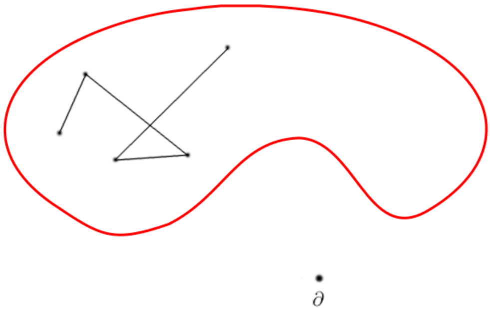
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



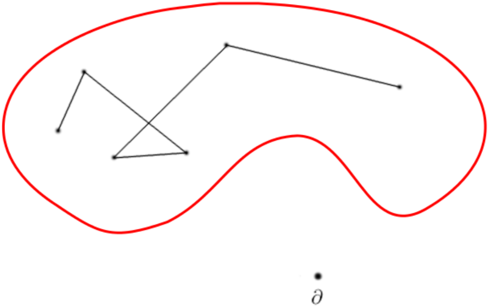
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



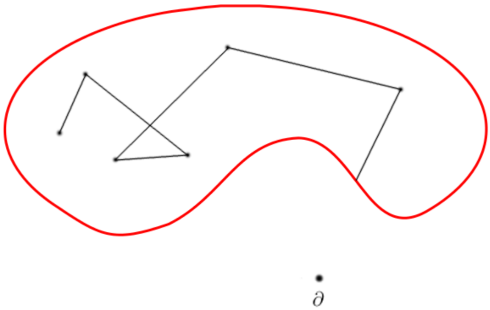
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



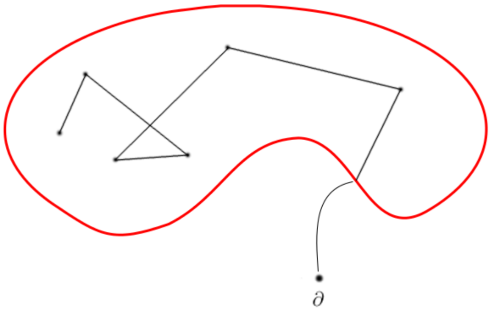
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



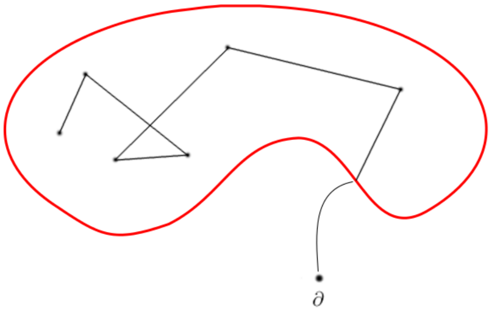
Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere.
The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



Let E be a bounded regular domain of $\mathbb{R}^d$ and S_{d-1} the unit sphere. The NRW (X,V) is a PDMP in $\mathbb{R}^d \times S_{d-1}$.



Theorem (Horton, Kyprianou, V)

Under *density conditions*, there exists $\exists \nu_{qs}$ and $\alpha \in (0,1), C > 0$ such that

$$\left\| \mathbb{P}_\mu (X_n \in \cdot \mid n < \tau_\partial) - \nu_{qs} \right\|_{TV} \leq C\alpha^n, \ \forall n \in \mathbb{N}, \ \forall \mu.$$

Almost Sure Convergence of Measure Valued Polya Processes

Let $(X_t)_{t\geq 0}$ be a Markov process evolving in a state space $E \cup \{\partial\}$ with ∂ absorbing. We fix $T > 0$ and consider the following self reinforced process $(Y_t)_{t\geq 0}$ with state space E:

Let $(X_t)_{t\geq 0}$ be a Markov process evolving in a state space $E \cup \{\partial\}$ with ∂ absorbing. We fix $T > 0$ and consider the following self reinforced process $(Y_t)_{t\geq 0}$ with state space E:

- Starting from $x \in E$, the process Y evolves according to the dynamic of X up to $\tau_1 := T \wedge \tau_\partial$.

Let $(X_t)_{t\geq 0}$ be a Markov process evolving in a state space $E\cup\{\partial\}$ with ∂ absorbing. We fix $T>0$ and consider the following self reinforced process $(Y_t)_{t\geq 0}$ with state space E:

- Starting from $x\in E$, the process Y evolves according to the dynamic of X up to $\tau_1 := T\wedge\tau_\partial$.
- Then Y_{τ_1} is chosen uniformly on the past trajectory of Y:

$$Y_{\tau_1} := Y_{U_1}, \text{ where } U_1 \sim \mathbb{U}([0,\tau_1)).$$

Let $(X_t)_{t\geq 0}$ be a Markov process evolving in a state space $E \cup \{\partial\}$ with ∂ absorbing. We fix $T > 0$ and consider the following self reinforced process $(Y_t)_{t\geq 0}$ with state space E:

- Starting from $x \in E$, the process Y evolves according to the dynamic of X up to $\tau_1 := T \wedge \tau_\partial$.
- Then Y_{τ_1} is chosen uniformly on the past trajectory of Y:

$$Y_{\tau_1} := Y_{U_1}, \text{ where } U_1 \sim \mathbb{U}([0, \tau_1)).$$

- Then Y evolves according to the dynamic of X up to the minimum of $\tau_1 + T$ and the next absorption time, and so on...

Let $(X_t)_{t\geq 0}$ be a Markov process evolving in a state space $E \cup \{\partial\}$ with ∂ absorbing. We fix $T > 0$ and consider the following self reinforced process $(Y_t)_{t\geq 0}$ with state space E:

- Starting from $x \in E$, the process Y evolves according to the dynamic of X up to $\tau_1 := T \wedge \tau_\partial$.
- Then Y_{τ_1} is chosen uniformly on the past trajectory of Y:

$$Y_{\tau_1} := Y_{U_1}, \text{ where } U_1 \sim \mathbb{U}([0,\tau_1)).$$

- Then Y evolves according to the dynamic of X up to the minimum of $\tau_1 + T$ and the next absorption time, and so on...

Typical convergence result

$$\frac{1}{t}\int_0^t \delta_{Y_s}\,\mathrm{d}s \xrightarrow[t\to+\infty]{a.s.} \nu_{qs} \qquad \text{(hopefully)}$$

Is this always true under reasonable conditions? There are no general answer as of now, but here are works on the question:

- Aldous D., Flannery B., Palacios J.L. (1988) for finite spaces;
- Benaïm M., Cloez B. (2015) for finite spaces;
- Blanchet J., Glynn P., Zheng S. (2016) for finite spaces;
- Benaïm M., Cloez B., Panloup F. (2018) for compact spaces;
- Wang A.Q., Roberts G.O., Steinsaltz D. (2020) for diffusion processes with a time inhomogeneous reinforcing mechanism;
- *Mailler C., V. (2020) for processes in non-compact spaces;*
- Benaim M., Champagnat N., V. (2021) for diffusion processes in a bounded open space.

However, none of them applies to the Neutron Random Walk.

An other natural algorithm based on a *Fleming-Viot type process*[1].
Let $N \geq 2$ and define $\mathbb{X} = (X_t^1, \ldots, X_t^N)_{t \geq 0}$ as follows:

- $\mathbb{X}$ evolves as N independent copies of X.
- At the first absorption time, the position of the absorbed particle is resampled according to the empirical distribution of the $N-1$ other particles (in E).
- Then $\mathbb{X}$ evolves as N independent copies of X up to the next absorption time, and so on.

[1] Actually not a very good name, one may prefer a *Moran particle system*

An other natural algorithm based on a *Fleming-Viot type process*[1].
Let $N \geq 2$ and define $\mathbb{X} = (X_t^1, \ldots, X_t^N)_{t\geq 0}$ as follows:

- $\mathbb{X}$ evolves as N independent copies of X.
- At the first absorption time, the position of the absorbed particle is resampled according to the empirical distribution of the $N-1$ other particles (in E).
- Then $\mathbb{X}$ evolves as N independent copies of X up to the next absorption time, and so on.

Some early references on this algorithm:

- Burdzy K., Holyst R., Ingerman D., March P. (1996, 2000)
- Del Moral P. (1996+)
- Del Moral P.& Guionnet A. (2001), & Miclo L. (2000,2003)
- Rousset M. (2006+)

[1] Actually not a very good name, one may prefer a *Moran particle system*

An other natural algorithm based on a *Fleming-Viot type process*[1].
Let $N \geq 2$ and define $\mathbb{X} = (X_t^1, ..., X_t^N)_{t \geq 0}$ as follows:

- $\mathbb{X}$ evolves as N independent copies of X.
- At the first absorption time, the position of the absorbed particle is resampled according to the empirical distribution of the $N-1$ other particles (in E).
- Then $\mathbb{X}$ evolves as N independent copies of X up to the next absorption time, and so on.

Some early references on this algorithm:

- Burdzy K., Holyst R., Ingerman D., March P. (1996, 2000)
- Del Moral P. (1996+)
- Del Moral P.& Guionnet A. (2001), & Miclo L. (2000,2003)
- Rousset M. (2006+)

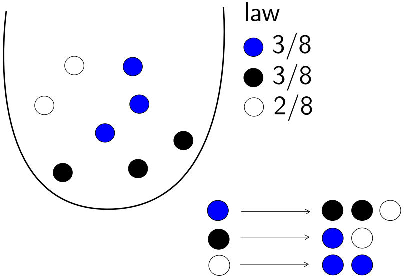
Personal contributions :

- In general state spaces, incl. hard boundaries (PhD 2011+)
- NRW: see also Oçafrain W. and V. (2018).

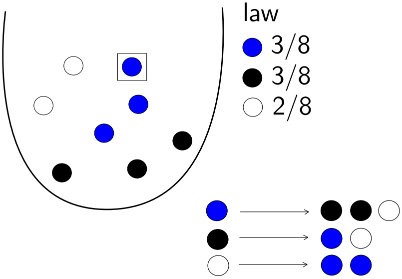
[1] Actually not a very good name, one may prefer a *Moran particle system*

Almost Sure Convergence of Measure Valued Polya Processes

State of the urn at time n

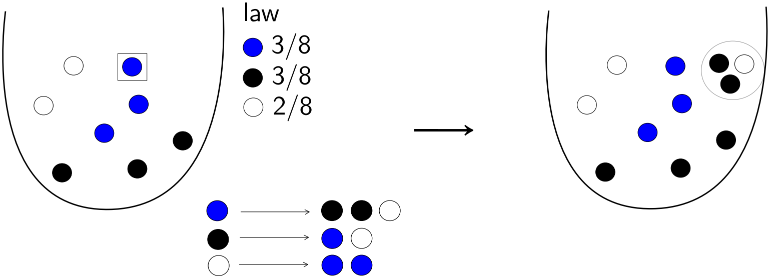


State of the urn at time n

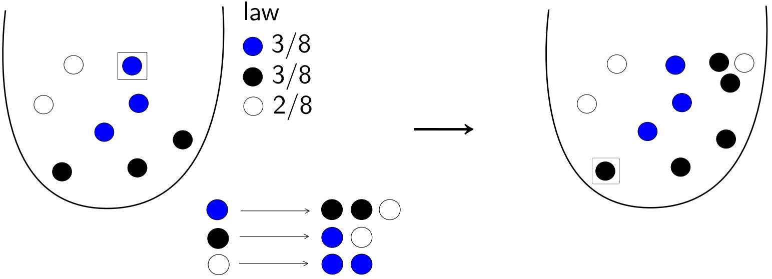


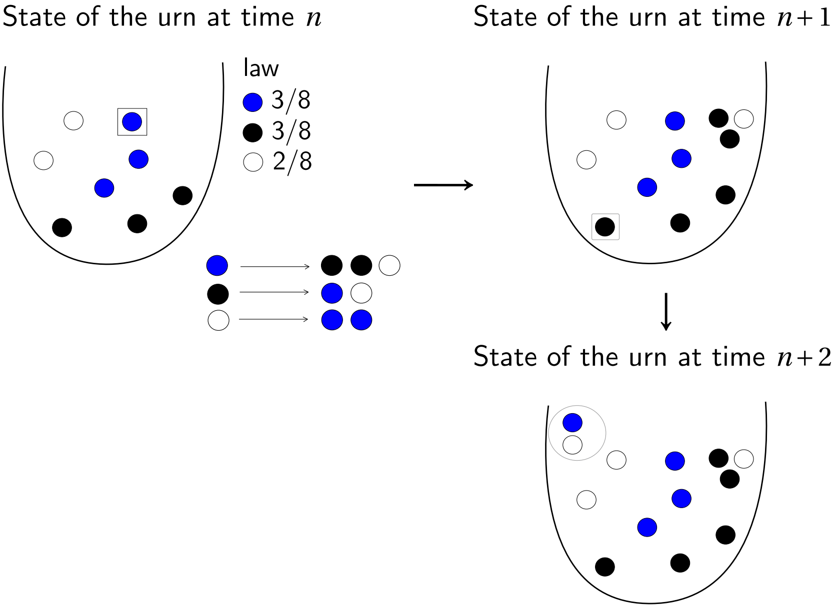
State of the urn at time n

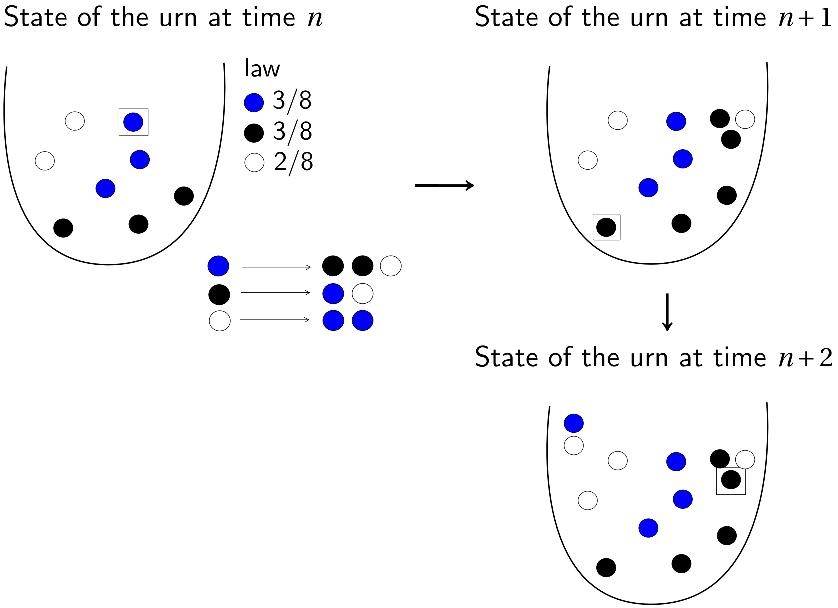
State of the urn at time $n+1$



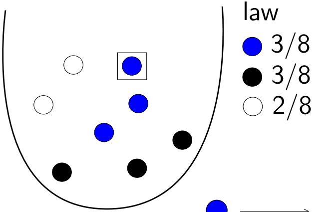
State of the urn at time n State of the urn at time $n+1$



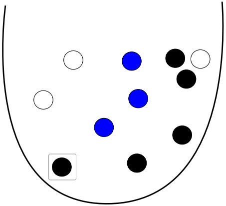

State of the urn at time n
law
3/8
3/8
2/8
State of the urn at time $n+1$
State of the urn at time $n+2$

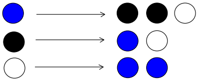

State of the urn at time n
law
3/8
3/8
2/8
State of the urn at time $n+1$
State of the urn at time $n+2$

State of the urn at time n

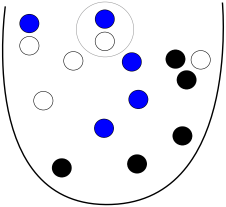


State of the urn at time $n+1$

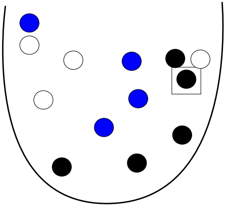




State of the urn at time $n+3$



State of the urn at time $n+2$



State of the urn at time n

$$m_n = 3\delta_{\text{blue}} + 3\delta_{\text{black}} + 2\delta_{\text{white}}.$$

State of the urn at time n

$$m_n = 3\delta_{\text{blue}} + 3\delta_{\text{black}} + 2\delta_{\text{white}}.$$

Pick a color $Y_{n+1} \in E := \{\text{blue}, \text{black}, \text{white}\}$, with probability

$$Y_{n+1} \sim \frac{m_n}{m_n(E)} = \frac{3}{8}\delta_{\text{blue}} + \frac{3}{8}\delta_{\text{black}} + \frac{2}{8}\delta_{\text{white}}.$$

State of the urn at time n

$$m_n = 3\delta_{\text{blue}} + 3\delta_{\text{black}} + 2\delta_{\text{white}}.$$

Pick a color $Y_{n+1} \in E := \{\text{blue}, \text{black}, \text{white}\}$, with probability

$$Y_{n+1} \sim \frac{m_n}{m_n(E)} = \frac{3}{8}\delta_{\text{blue}} + \frac{3}{8}\delta_{\text{black}} + \frac{2}{8}\delta_{\text{white}}.$$

We then get the state of the urn at time $n+1$ *via*

$$\boxed{m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}, \quad \text{where } R^{(n+1)} = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}.}$$

State of the urn at time n

$$m_n = 3\delta_{\text{blue}} + 3\delta_{\text{black}} + 2\delta_{\text{white}}.$$

Pick a color $Y_{n+1} \in E := \{\text{blue}, \text{black}, \text{white}\}$, with probability

$$Y_{n+1} \sim \frac{m_n}{m_n(E)} = \frac{3}{8}\delta_{\text{blue}} + \frac{3}{8}\delta_{\text{black}} + \frac{2}{8}\delta_{\text{white}}.$$

We then get the state of the urn at time $n+1$ *via*

$$\boxed{m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}, \quad \text{where } R^{(n+1)} = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}.}$$

For instance, if $Y_{n+1} = \text{blue}$, then

$$m_{n+1} = 3\delta_{\text{blue}} + 5\delta_{\text{black}} + 3\delta_{\text{white}}.$$

→ Aldous, Flannery, Palacios 1988, Janson 2004, Pemantle 2007
→ Bandyopadhyay, Thacker 2017, Mailler, Marckert 2017

Almost Sure Convergence of Measure Valued Polya Processes

1. Quasi-stationary distributions
   a. A simple example
   b. Definition and a general result
   c. Application to the Neutron Random Walk
   d. An approximation scheme

2. Measure valued Pólya processes
   a. Finite and irreducible Pólya urns
   b. Formal description of a Measure Valued Pólya Process*
   c. Application to the approximation scheme

Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:

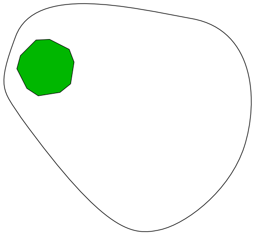
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



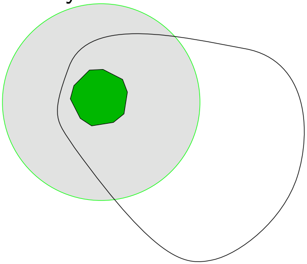
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



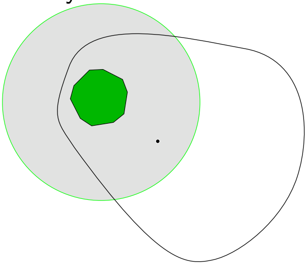
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



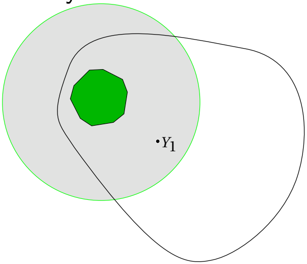
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.$$

R may be random and sampling according to $m_n P / m_n P(E)$:



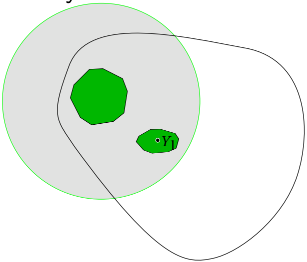
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



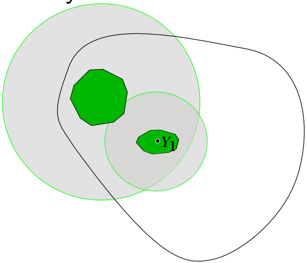
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



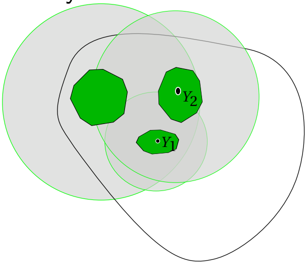
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



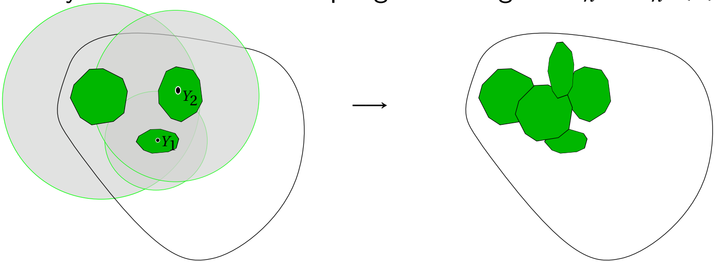
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



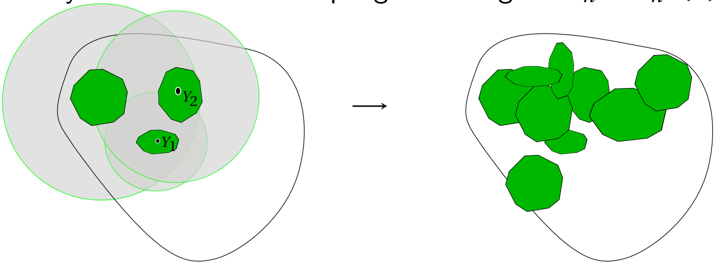
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



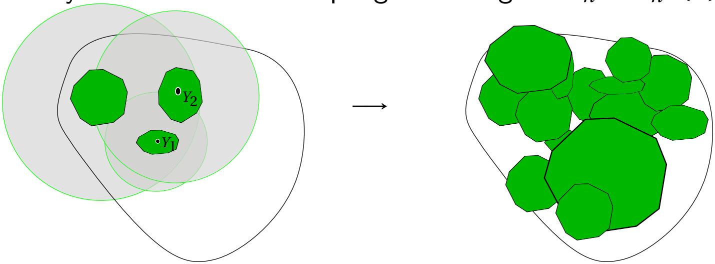
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



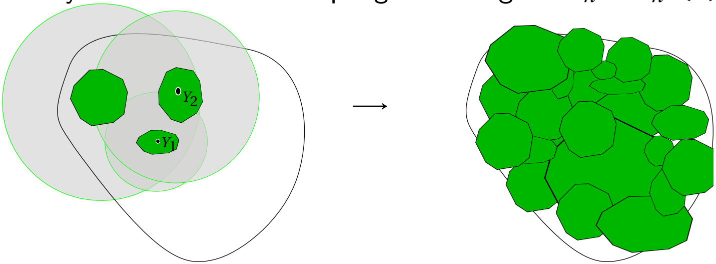
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



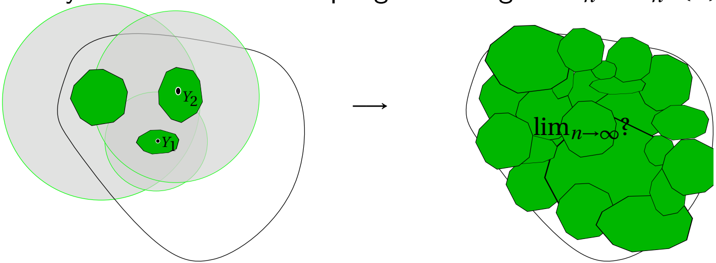
Let E be a Polish space, $\mathscr{M}_+$ the set of positive measures on E and

- $m_0 \in \mathscr{M}_+$ the initial urn,
- $R^{(i)}$ (iid random) kernels from E to E.
- P a kernel from E to E

The urn m_n being defined, we pick and set

$$\boxed{Y_{n+1} \sim \frac{m_n P}{m_n P(E)} \quad \text{et} \quad m_{n+1} = m_n + \delta_{Y_{n+1}} R^{(n+1)}.}$$

R may be random and sampling according to $m_n P / m_n P(E)$:



Hypothesis C. There exist $c_1, c_2, \theta > 0$ and a locally bounded function V with compact level sets, such that

- $0 < c_1 \leq \inf_{x \in E} R_x(E) \leq \sup_{x \in E} R_x(E) \leq 1$, where $R_x := \mathbb{E}(R_x^{(1)})$

Hypothesis C. There exist $c_1, c_2, \theta > 0$ and a locally bounded function V with compact level sets, such that

- $0 < c_1 \leq \inf_{x \in E} R_x(E) \leq \sup_{x \in E} R_x(E) \leq 1$, where $R_x := \mathbb{E}(R_x^{(1)})$
- $\theta \in (0, c_1)$ and

$$R_x \cdot V \leq \theta V(x) + c_2, \quad \forall x \in E.$$

Hypothesis C. There exist $c_1, c_2, \theta > 0$ and a locally bounded function V with compact level sets, such that

- $0 < c_1 \le \inf_{x \in E} R_x(E) \le \sup_{x \in E} R_x(E) \le 1$, where $R_x := \mathbb{E}(R_x^{(1)})$
- $\theta \in (0, c_1)$ and

$$R_x \cdot V \le \theta V(x) + c_2, \quad \forall x \in E.$$

- the sub Markov kernel $R - I$ admits a QSD ν_{qs}. In addition, the convergence toward the QSD is uniform in total variation norm in $\{\alpha \in \mathcal{M}_1(E) \mid \alpha \cdot V^{1/q} \le C\}$, $C > 0$ for some $q > 1$.

Hypothesis C. There exist $c_1, c_2, \theta > 0$ and a locally bounded function V with compact level sets, such that

- $0 < c_1 \leq \inf_{x \in E} R_x(E) \leq \sup_{x \in E} R_x(E) \leq 1$, where $R_x := \mathbb{E}(R_x^{(1)})$
- $\theta \in (0, c_1)$ and

$$R_x \cdot V \leq \theta V(x) + c_2, \quad \forall x \in E.$$

- the sub Markov kernel $R - I$ admits a QSD ν_{qs}. In addition, the convergence toward the QSD is uniform in total variation norm in $\{\alpha \in \mathcal{M}_1(E) \mid \alpha \cdot V^{1/q} \leq C\}$, $C > 0$ for some $q > 1$.
- R is Feller and some (technical) bounds on moments of $R_x^{(1)}$.

(Note the similarities with Assumption E)

Hypothesis C. There exist $c_1, c_2, \theta > 0$ and a locally bounded function V with compact level sets, such that

- $0 < c_1 \leq \inf_{x \in E} R_x(E) \leq \sup_{x \in E} R_x(E) \leq 1$, where $R_x := \mathbb{E}(R_x^{(1)})$
- $\theta \in (0, c_1)$ and

$$R_x \cdot V \leq \theta V(x) + c_2, \quad \forall x \in E.$$

- the sub Markov kernel $R - I$ admits a QSD ν_{qs}. In addition, the convergence toward the QSD is uniform in total variation norm in $\{\alpha \in \mathcal{M}_1(E) \mid \alpha \cdot V^{1/q} \leq C\}$, $C > 0$ for some $q > 1$.
- R is Feller and some (technical) bounds on moments of $R_x^{(1)}$.

(Note the similarities with Assumption E)

Theorem (Mailler, V. 2020)

If $m_0 \cdot V < \infty$, then the sequence $(m_n / m_n(E))_{n \geq 0}$ converges almost-surely to ν_{qs} for the topology of weak convergence.

Almost Sure Convergence of Measure Valued Polya Processes

Recall that we are given a Markov process X evolving in $E \cup \{\partial\}$ and $T > 0$ is fixed. We consider the initial composition measure $m_0 = \delta_{x_0}$ and define the random kernel

$$R_x^{(1)} \simeq \int_0^{T \wedge \tau_\partial} \delta_{X_s} \, \mathrm{d}s, \text{ where } (X_t)_{t \geq 0} \sim \mathbb{P}_x.$$

Then the law of the empirical distribution of the reinforced process at the n^{th} resampling time is distributed as $m_n / m_n(E)$.

Recall that we are given a Markov process X evolving in $E \cup \{\partial\}$ and $T > 0$ is fixed. We consider the initial composition measure $m_0 = \delta_{x_0}$ and define the random kernel

$$R_x^{(1)} \simeq \int_0^{T \wedge \tau_\partial} \delta_{X_s} \, \mathrm{d}s, \text{ where } (X_t)_{t \geq 0} \sim \mathbb{P}_x.$$

Then the law of the empirical distribution of the reinforced process at the n^{th} resampling time is distributed as $m_n / m_n(E)$.

Example. Assume that X is solution to the SDE

$$dX_t = dB_t + b(X_t)\, dt, \; X_0 \in \mathbb{R}^d$$

killed at rate κ, with $\|\kappa\|_\infty < +\infty$, let Y be the reinforced diffusion.

Recall that we are given a Markov process X evolving in $E \cup \{\partial\}$ and $T > 0$ is fixed. We consider the initial composition measure $m_0 = \delta_{x_0}$ and define the random kernel

$$R_x^{(1)} \simeq \int_0^{T \wedge \tau_\partial} \delta_{X_s} \, \mathrm{d}s, \text{ where } (X_t)_{t \geq 0} \sim \mathbb{P}_x.$$

Then the law of the empirical distribution of the reinforced process at the n^{th} resampling time is distributed as $m_n / m_n(E)$.

Example. Assume that X is solution to the SDE

$$dX_t = dB_t + b(X_t)\, dt, \; X_0 \in \mathbb{R}^d$$

killed at rate κ, with $\|\kappa\|_\infty < +\infty$, let Y be the reinforced diffusion.

Theorem (Mailler, V. 2020)

If $\limsup_{x \to +\infty} \frac{\langle b(x), x \rangle}{|x|} < -\frac{3}{2} \|\kappa\|_\infty^{1/2}$, then

$$\frac{1}{t} \int_0^t \delta_{Y_s} \, \mathrm{d}s \xrightarrow[t \to +\infty]{weak} \nu_{QSD} \text{ a.s.}$$

Difficulty: We observe that

$$Q_x(E) = \mathbb{E}\left(R_x^{(1)}(E)\right) = \mathbb{E}_x(\tau_1 \wedge T) \leq T\,\mathbb{P}_x(\tau_1 \leq T),$$

so that, when the underlying is the NRW, then the condition on the minimum of $Q_x(E)$ does not hold true.

Difficulty: We observe that

$$Q_x(E) = \mathbb{E}\left(R_x^{(1)}(E)\right) = \mathbb{E}_x(\tau_1 \wedge T) \leq T\,\mathbb{P}_x(\tau_1 \leq T),$$

so that, when the underlying is the NRW, then the condition on the minimum of $Q_x(E)$ does not hold true.

A possible direction. The same difficulty holds when X is an elliptic diffusion process killed at the boundary of a bounded regular open set D. In this case, one can manage to prove by coupling methods that $\frac{1}{t}\int_0^t \delta_{Y_s}\,\mathrm{d}s$ is tight,

Difficulty: We observe that

$$Q_x(E) = \mathbb{E}\left(R_x^{(1)}(E)\right) = \mathbb{E}_x(\tau_1 \wedge T) \le T\,\mathbb{P}_x(\tau_1 \le T),$$

so that, when the underlying is the NRW, then the condition on the minimum of $Q_x(E)$ does not hold true.

A possible direction. The same difficulty holds when X is an elliptic diffusion process killed at the boundary of a bounded regular open set D. In this case, one can manage to prove by coupling methods that $\frac{1}{t}\int_0^t \delta_{Y_s}\,\mathrm{d}s$ is tight,and obtain:

Theorem (Benaïm, Champagnat, V. 2021)

Let Y be a reinforced elliptic diffusion process, resampled according to its historical empirical distribution when it hits the boundary, then

$$\frac{1}{t}\int_0^t f(Y_s)\,\mathrm{d}s \xrightarrow[t\to+\infty]{a.s.} \nu_{QS},$$

where ν_{QS} is the unique quasi-stationary distribution of the underlying process, and f is any bounded continuous function.

The END