Characterization of Quasi-stationary Distributions

Iddo Ben-Ari, joint with Ningwei Jiang
Department of Mathematics, University of Connecticut
iddo.ben-ari.uconn.edu

UCONN
COLLEGE OF LIBERAL ARTS AND SCIENCES

Overview

Quasistationary distributions (QSDs) relate to the long-term behavior of processes which eventually reach some absorbing state (a state the process cannot leave) as limits (if exist) of the distribution of the state of the process, conditioned on not being absorbed. They are, in a sense, the mathematical version of the description of life on earth in the distant future, assuming there will be life on earth in the distant future...

Study of QSDs dates to at least as early as Wright's work [8] on gene frequencies in finite populations and Yaglom's [9] work on branching processes. There has been extensive research on QSDs with focus on concrete processes exhibiting quasistionarity (e.g. branching processes and birth and death chains) or on sufficient conditions for existence of QSDs, convergence results, and, recently, the problem of sampling from QSDs.

We present a comprehensive approach for characterizing and representing QSDs for Markov Chains under standard assumptions (our approach is complete only in the discrete-time setting). This is done through adaption and application of classical methods in the study of Markov chains to the context of QSDs.

Hypotheses and Preliminary Results

Process. $\mathbf{X} = (X_t : t \in \mathbb{Z}_+ \text{ or } t \in \mathbb{R}_+)$, a Markov chain on state space $S \cup \{\Delta\}$, with S finite or countably infinite.

Hitting times. Let $\emptyset \neq K \subsetneq S$. Define the hitting time of K

$$\tau_K := \begin{cases} \inf\{n \in \mathbb{N} : X_n \in K\} & \text{discrete time} \\ \inf\{t \in (0,\infty) : X_{t-} \notin K,\ X_t \in K\} & \text{continuous time} \end{cases}$$

For each state x, write τ_x as shorthand for $\tau_{\{x\}}$.

H1.
- S is an irreducible class.
- If $\mathbf{X}$ is a continuous-time process, also assume that the holding time at each state $x \in S$ is exponential with parameter $q_x \in (0,\infty)$.

H2. Δ is an absorbing state, and the hitting time of Δ, τ_Δ is finite a.s.

QSD. A **QSD** is a probability measure ν on S satisfying that for some probability measure μ on S

$$P_\mu(X_t \in \cdot \mid \tau_\Delta > t) \underset{t\to\infty}{\to} \nu(\cdot). \tag{1}$$

In fact, this is equivalent to $P_\nu(X_t \in \cdot \mid \tau_\Delta > t) = \nu(\cdot)$ for all $t > 0$ (e.g. [5], also for the proof of the following).

Proposition 1 (Geometric/Exponential Absorption Time) *If ν is a QSD, then there exists $\lambda_\nu \in (0,\infty)$, the* **absorption parameter**, *satisfying*

$$P_\nu(\tau_\Delta > t) = e^{-\lambda_\nu t}, \quad \text{for all } t \geq 0. \tag{2}$$

Critical Parameter. Let

$$\lambda_{\mathbf{cr}} := \sup\{\lambda \geq 0 : E_x[\exp(\lambda\tau_\Delta)] < \infty \text{ for some } x \in S\}.$$

Corollary 1
- *If $\lambda_{cr} = 0$, there are no QSDs.*
- *If ν is a QSD, then $\lambda_\nu \in (0, \lambda_{cr}]$.*

H3. $\lambda_{cr} > 0$.

MGF Regimes. Our analysis is according to the following dichotomy:
- λ_{cr} is in the **infinite MGF regime** if $E_x[\exp(\lambda_{cr}\tau_\Delta)] = \infty$ for some (equivalently, all) $x \in S$.
- A parameter $\lambda(\leq \lambda_{cr})$ is in the **finite MGF regime** if $E_x[\exp(\lambda\tau_\Delta)] < \infty$.

Infinite MGF Regime: Existence and Representation

Theorem 1 *Suppose λ_{cr} is in the infinite regime.*

1. *There exists a QSD with absorption parameter λ_{cr} if and only there exists a finite $\emptyset \neq K \subsetneq S$,*
$$E_x[\exp(\lambda_{cr}\tau_\Delta \wedge \tau_K)] < \infty \text{ for some } x \in S. \tag{3}$$
In this case there exists a unique QSD, ν_{cr}, with an absorption parameter λ_{cr}, given by
$$\nu_{cr}(x) = \frac{1}{E_x[\exp(\lambda_{cr}\tau_\Delta), \tau_\Delta < \tau_x]} \times \begin{cases} e^{\lambda_{cr}} - 1 & \text{discrete time} \\ \frac{\lambda_{cr}}{q_x - \lambda_{cr}} & \text{continuous time} \end{cases} \tag{4}$$
2. *If, in addition,*
$$E_x[\exp(\lambda_{cr}\tau_x)\tau_x, \tau_x < \tau_\Delta] < \infty \text{ for some } x \in S, \tag{5}$$
Then (1) holds for any finitely supported μ, with $\nu = \nu_{cr}$. If $\mathbf{X}$ is a discrete-time chain, the convergence is along multiples of the period of $p|_S$

Example 1 (Subcritical Branching Process) *Let $\mathbf{X}$ be a branching process with offspring distribution B with $0 < E[B] < 1$ (subcritical branching). Here $\Delta = 0$ and S is determined by the support of B. Then*

1. $e^{-\lambda_{cr}} = E[B]$.
2. *λ_{cr} is in the infinite regime and the first condition of Theorem 1 holds (e.g. with $K = \min\{k \geq 1 : P(B = k) > 0\}$.*
3. *The second condition of Theorem 1 holds if and only if $E[B\log^+ B] < \infty$. Yet, the conclusion holds without this extra assumption (e.g. [2]).*

Infinite MGF Regime: the Uniform Case

In the next theorem we drop **H3**.

Theorem 2

1. *Suppose that there exists $\bar\lambda > 0$ and a finite $\emptyset \neq K \subsetneq S$ such that*
$$E_x[\exp(\bar\lambda\tau_\Delta)] = \infty \text{ for some } x \in S \text{ and} \tag{6}$$
$$\sup_x E_x[\exp(\bar\lambda\tau_\Delta \wedge \tau_K)] < \infty. \tag{7}$$
Then $\lambda_{cr} \in (0, \bar\lambda]$, λ_{cr} is in the infinite regime, and both conditions of Theorem 1 hold.
2. *If, in addition, there exists some $x_0 \in S$ and $t_0 > 0$ such that*
$$\inf_x P(\tau_{x_0} < \tau_\Delta | \tau_\Delta > t_0) > 0. \tag{8}$$
then the convergence (1) holds for all μ, with $\nu = \nu_{cr}$. In the discrete-time setting the convergence is along multiplies of the period of $p|_S$.

This theorem was inspired by the main result in [4]. There the authors consider a continuous-time Markov chain satisfying (6),(7) and (8), but replace irreducibility with a weaker condition. They use the assumptions to prove convergence of the conditional distributions in the total variation norm, with an explicit exponential bound, indirectly proving existence and uniqueness of a QSD. Our version of the theorem is much weaker, yet it identifies this QSD as ν_{cr}.

Example 2 *When S is finite, both conditions of Theorem 2 hold.*

Finite MGF Regime: Existence Results

As the finite state space case is settled in Example 2, we add:

H0. S is infinite.

Theorem 3 *Let $\lambda > 0$ be in the finite regime.*
- *If for some $\lambda' \in (0,\lambda)$, $\lim_{x\to\infty} E_x[\exp(\lambda'\tau_\Delta)] = \infty$, then there exists a QSD with absorption parameter λ.*
- *If $\sup_x E_x[\exp(\lambda\tau_\Delta)] < \infty$ then there does not exist a QSD with absorption parameter λ.*

Finite MGF Regime: Existence Results, continued

Corollary 2 *Let*
$$\lambda_0 := \inf\{\lambda \in (0,\lambda_{cr}) : \lim_{x\to\infty} E_x[\exp(\lambda\tau_\Delta)] = \infty\} \ (\textit{convention: } \inf\emptyset = \infty).$$
There exists a QSD with absorption parameter λ for every $\lambda \in (\lambda_0, \lambda_{cr}]$.

Example 3 (Subcritical Branching, continued) *Let $\mathbf{X}$ be as in Example 1 and assume further that $P(B \in \{0,1\}) < 1$. Recall that $E[B] = e^{-\lambda_{cr}}$. In this case, $E_x[\tau_\Delta] \underset{x\to\infty}{\to} \infty$, therefore Corollary 2 yields the existence of a QSD with absorption parameter λ for every $\lambda \in (0, \lambda_{cr}]$. This is well-known (e.g. [2]).*

Example 4 (Infinity as a Natural Boundary) *In [3] the authors prove the existence of a QSD under the assumption that for all $t > 0$,*
$$\lim_{x\to\infty} P_x(\tau_\Delta \leq t) = 0.$$
As this condition implies $\lambda_0 = 0$, Corollary 2 gives a continuum of QSDs.

- *In our proof of Theorem 3 we applied a tightness argument from [3].*

Example 5 (Birth and Death Chains) *Let $\mathbf{X}$ be a continuous-time birth and death chain on $\mathbb{Z}_+ \cup \{-1\}$, with birth rates $(\lambda_k : k \in \mathbb{Z}_+ \cup \{-1\})$ and death rates $(\mu_k, k \in \mathbb{Z}_+)$. We assume $\lambda_{-1} = 0$ and $\lambda_k\mu_k \in (0,\infty)$ for $k \in \mathbb{Z}_+$. Thus here, $S = \mathbb{Z}_+$ and $\Delta = -1$. Letting $\pi_0 := 0, \pi_n := \prod_{j=1}^n \frac{\lambda_{j-1}}{\mu_j}$, $n \in \mathbb{N}$, then we assume $\sum_{n=0}^\infty \frac{1}{\lambda_n\pi_n} = \infty$, which is equivalent to $\tau_\Delta < \infty$ a.s. Let τ be the random variable on $[0,\infty]$ whose CDF is given by*
$$P(\tau \leq t) = \lim_{x\to\infty} P_x(\tau_\Delta \leq t).$$
Thus, τ is the absorption time, starting from ∞. Moreover, $E[\tau] = \sum_{n=1}^\infty \frac{1}{\lambda_n\pi_n}\sum_{i=n+1}^\infty \pi_i$. (details on the series can be found in [1]).

The following alternatives hold:

1. *$E[\tau] < \infty$. In this case the conditions of Theorem 2 hold. In particular, there exists a unique QSD which is minimal and convergence in (1) holds for any μ, with $\nu = \nu_{cr}$.*
2. *$E[\tau] = \infty$. Then either*
   - *(a) $\lambda_{cr} = 0$ and no QSDs exist.*
   - *(b) $\lambda_{cr} > 0$, and then Corollary 2 holds with $\lambda_0 = 0$. In particular, there exists a QSD with absorption parameter λ for all $\lambda \in (0, \lambda_{cr}]$.*

This result, along with identification of the QSDs for each absorption parameter were originally obtained in [7] through detailed analysis of the spectral representation of the transition function for $\mathbf{X}$. The approach we present here is based on routine analysis of moments. Identification of the QSDs can be further obtained through Theorem 4.

Finite MGF Regime: Characterization

In this section assume that $\mathbf{X}$ is a discrete-time chain, and that **H0**, **H1**,**H2** and **H3** hold.

Green's Kernels. For $\lambda > 0$ in the finite regime, define the **Green's function**, $G^\lambda(\cdot,\cdot)$:

$$G^\lambda(x,y) := \sum_{n=0}^\infty e^{\lambda n} P_x(X_n = y) = \frac{E_x[\exp(\lambda\tau_y), \tau_y < \tau_\Delta]}{1 - E_y[\exp(\lambda\tau_y), \tau_y < \tau_\Delta]}.$$

For every $x \in S$, $G^\lambda(x,\cdot)$ is a finite measure on S with total mass

$$G^\lambda(x,\mathbf{1}) = \frac{E_x[\exp(\lambda\tau_\Delta)] - 1}{e^\lambda - 1}.$$

Normalize, denoting the resulting probability measure by $\boldsymbol{K^\lambda(x,\cdot)}$:

$$K^\lambda(x,y) := \frac{G^\lambda(x,y)}{G^\lambda(x,\mathbf{1})} = \cdots = \frac{E_x[\exp(\lambda\tau_\Delta), \tau_y < \tau_\Delta]}{E_x[\exp(\lambda\tau_\Delta)] - 1} \times \frac{e^\lambda - 1}{E_y[\exp(\lambda\tau_\Delta), \tau_\Delta < \tau_y]}.$$

Martin (Entrance) Boundary.
- A sequence $\mathbf{x} = (x_n \in S : n \in \mathbb{N})$ with $\lim_{n\to\infty} x_n = \infty$ is **λ-convergent** if
$$K^\lambda(\mathbf{x},y) := \lim_{n\to\infty} K^\lambda(x_n,y) \text{ exists for all } y \in S.$$
- The λ-convergent sequences $\mathbf{x}$ and $\mathbf{x}'$ are **λ-equivalent** if $K^\lambda(\mathbf{x},\cdot) = K^\lambda(\mathbf{x}',\cdot)$, writing $[\mathbf{x}]$ for the equivalence class and $K^\lambda([\mathbf{x}],\cdot)$ for $K^\lambda(\mathbf{x},\cdot)$.
- The **Martin Boundary** $\partial^\lambda S$ is the set of equivalence classes of λ-convergent sequences.
- The **Martin Topology** is a metric on $S \cup \partial^\lambda S$ which makes it compact, with $\partial^\lambda S$ closed.
- Let
$$H^\lambda := \{[\mathbf{x}] \in \partial^\lambda S : K^\lambda([\mathbf{x}],\cdot) \text{ is a QSD}\}.$$

Proposition 2
- *All elements of H^λ are QSDs with absorption parameter λ.*
- *$[\mathbf{x}] \in H^\lambda$ if and only if $(K^\lambda(x_n,\cdot) : n \in \mathbb{N})$ is tight for every $\mathbf{x} \in [\mathbf{x}]$.*

The following Choquet-type theorem holds:

Theorem 4 *Suppose that λ is in the finite regime. If ν is a QSD with absorption parameter λ, then there exists a Borel probabilty measure F_ν on $\partial^\lambda S$ with $F_\nu(H^\lambda) = 1$ such that*

$$\nu(y) = \int_{[\mathbf{x}]\in\partial^\lambda \mathbf{S}} K^\lambda([\mathbf{x}],y)\,dF_\nu([\mathbf{x}]).$$

Remarks:

1. All QSDs with absorption parameter λ are in the convex hull of H^λ.
2. The construction and the arguments are essentially the same as for the Martin (exit) boundary (e.g. [6]) with differences in details.
3. All QSDs are excessive measures for p (measures satisfying $\nu p \leq \nu$ on $S \cup \{\Delta\}$), and representation theory for excessive measures was developed decades ago. The main contributions here are:
   - (a) Identifying the kernels which generate all QSDs for a given absorption parameter rather concretely, not as an abstract combination of excessive measures;
   - (b) Raise the awareness to the relevance of the Martin boundary in the study of QSDs.

Example 6 (Rooted Tree) *Consider a discrete-time nearest neighbor Markov chain on a rooted tree with a single absorbing state Δ whose only neighbor is the root. We will further assume that* ***H0*–*H3*** *hold.*

For every λ in the finite regime, and every sequence $(x_n : n \in \mathbb{N})$ tending to infinity along an infinite branch, the limit $\lim_{n\to\infty} K^\lambda(x_n,\cdot)$ exists and is a QSD. Thus H^λ is indexed by the infinite branches.

Discrete-time birth and death chains are one special case.

Short Bibliography

[1] William J. Anderson. *Continuous-time Markov chains*. Springer Series in Statistics: Probability and its Applications. An applications-oriented approach. Springer-Verlag, New York, 1991, pp. xii+355. ISBN: 0-387-97369-9. DOI: 10.1007/978-1-4612-3038-0. URL: https://doi.org/10.1007/978-1-4612-3038-0.

[2] Krishna B. Athreya and Samuel Karlin. "On Branching Processes with Random Environments: I: Extinction Probabilities". In: *The Annals of Mathematical Statistics* 42 (5 1971). ISSN: 0003-4851. DOI: 10.1214/aoms/1177693150.

[3] P. A. Ferrari et al. "Existence of Quasi-Stationary Distributions. A Renewal Dynamical Approach". In: *The Annals of Probability* 23.2 (1995), pp. 501 –521. DOI: 10.1214/aop/1176988277. URL: https://doi.org/10.1214/aop/1176988277.

[4] Servet Martínez, Jaime San Martín, and Denis Villemonais. "Existence and uniqueness of a quasistationary distribution for markov processes with fast return from infinity". In: *Journal of Applied Probability* 51 (3 2014). ISSN: 00219002. DOI: 10.1239/jap/1409932672.

[5] Jaime San Martín Pierre Collet Servet Martínez. "Quasi-Stationary Distributions: Markov Chains, Diffusions and Dynamical Systems". In: (2013). DOI: https://doi.org/10.1007/978-3-642-33131-2.

[6] Stanley A. Sawyer. "Martin boundaries and random walks". In: (1997). DOI: 10.1090/conm/206/02685.

[7] Erik A. van Doorn and Philip K. Pollett. "Quasi-stationary distributions for discrete-state models". In: *European Journal of Operational Research* 230.1 (2013), pp. 1–14. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2013.01.032. URL: https://www.sciencedirect.com/science/article/pii/S0377221713000799.

[8] Sewall Wright. "EVOLUTION IN MENDELIAN POPULATIONS". In: *Genetics* 16.2 (Mar. 1931), pp. 97–159. ISSN: 1943-2631. DOI: 10.1093/genetics/16.2.97. eprint: https://academic.oup.com/genetics/article-pdf/16/2/97/35081059/genetics0097.pdf. URL: https://doi.org/10.1093/genetics/16.2.97.

[9] A. M. Yaglom. "Certain limit theorems of the theory of branching random processes". In: *Dokl. Akad. Nauk SSSR, n. Ser.* 56 (1947), pp. 795–798. ISSN: 0002-3264.