Springer Proceedings in Mathematics & Statistics

Janos Englander
Brian C. Rider *Editors*

# Advances in Superprocesses and Nonlinear PDEs

Springer

# Springer Proceedings in Mathematics & Statistics

Volume 38

For further volumes:
http://www.springer.com/series/10533

# Springer Proceedings in Mathematics & Statistics

This book series features volumes composed of select contributions from workshops and conferences in all areas of current research in mathematics and statistics, including OR and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

Janos Englander • Brian Rider
Editors

# Advances in Superprocesses and Nonlinear PDEs

Springer

*Editors*
Janos Englander
Department of Mathematics
University of Colorado
Boulder, Colorado, USA

Brian Rider
Department of Mathematics
University of Colorado
Boulder, Colorado, USA

ISSN 2194-1009 ISSN 2194-1017 (electronic)
ISBN 978-1-4614-6239-2 ISBN 978-1-4614-6240-8 (eBook)
DOI 10.1007/978-1-4614-6240-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013931995

Mathematics Subject Classification (2010): 60B15, 60H25, 60J80, 60F05, 60J85, 60J68, 60F25, 60K35, 92B05, 60J25, 35J60



Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

This book grew out of the conference *Advances in Superprocesses and Nonlinear PDEs* held between June 24 and June 26, 2010 at the University of Colorado Boulder. The main speakers at the meetings were

Zhen-Qing Chen (U. Washington)
Donald Dawson (Carleton, Ottawa)
Eugene B. Dynkin (Cornell)
Steve N. Evans (UC Berkeley)
Patrick J. Fitzsimmons (UC San Diego)
Klaus Fleischmann (Weierstrass Institute, Berlin)
Simon C. Harris (Bath, UK)
Andreas E. Kyprianou (Bath, UK)
Rinaldo Schinazi (University of Colorado, Colorado Springs)
Dan Stroock (MIT)

One of the motivations of the conference was recent advances in the theory of superprocesses. The last 10 years have witnessed intensive research on superprocesses, with important progress made on superprocesses over flows, backbone constructions, superprocesses in random media, interacting and branching-coalescing superprocesses, superprocesses with immigration, scaling limit theorems and self-intersection local times.

The meeting was also dedicated to the 60th birthday of our colleague, Sergei Kuznetsov.

Professor Kuznetsov is one of the top experts on measure-valued branching processes (or superprocesses) and their connection to nonlinear partial differential operators. His research interests range from stochastic processes and partial differential equations to mathematical statistics, time series analysis and statistical software. He has published over 90 papers in international journals.

Here we mention just two of his remarkable results. In 1980, Kuznetsov proved that every Markov process in a Borel state space (i.e., a measurable space isomorphic to a Borel subset in a Polish space) has a transition function. His most well-known contribution to probability theory though is the so-called Kuznetsov-measure.

*Duality* is a very important notion in probability: the stationary Markov processes with transition functions $p$, $\hat{p}$ are duals with respect to a given $\sigma$-finite measure $m$ if

$$m(\mathrm{d}x)p_t(x,\mathrm{d}y) = m(\mathrm{d}y)\hat{p}_t(y,\mathrm{d}x).$$

A closely related fact is that each Markov process can be considered in two time directions (this is the way Kolmogorov's forward and backward equations are deduced). In fact, Dynkin suggested that the functions $p$ and $\hat{p}$ may be interpreted as forward and backward transition functions of a *single* Markov process with random birth time $\alpha$ and death time $\beta$. This approach was applied also to nonstationary transition functions $p(s,x;t,\mathrm{d}y)$, $\hat{p}(s,x;t,\mathrm{d}y)$ and measures $m$ depending on the time interval $(s,t)$. The process $\{X_t\}_{t\in(\alpha,\beta)}$ can be given by its two-dimensional distributions as

$$m_{st}(\mathrm{d}x,\mathrm{d}y) = P(\alpha < s, X_s \in \mathrm{d}x, X_t \in \mathrm{d}y, t < \beta).$$

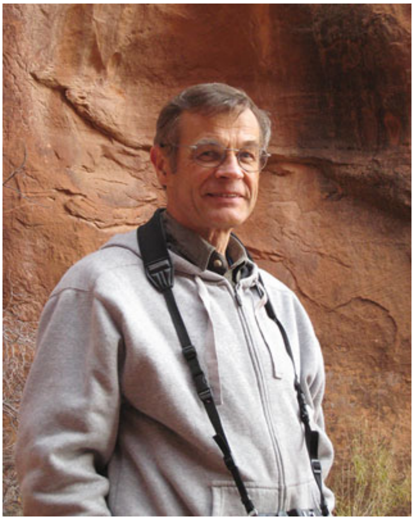
The problem, however, is that the family $\{m_{st}\}$ should satisfy a normalization condition that guarantees that $P$ is a probability measure; in the stationary case, this condition holds if $m$ is a probability measure, invariant for both processes. Since the definition of duality requires only $\sigma$-finiteness of $m$, this assumption is too restrictive. This problem was solved by Kuznetsov in 1973, who managed to get rid of the condition: the measure $P$ (called the Kuznetsov measure) and the corresponding $m$ are both just $\sigma$-finite. In the theory of Markov processes, considering a process with random birth and death times with the help of the Kuznetsov-measure has proven to be a very useful alternative to working with dual processes.

Sergei obtained his Ph.D. in 1976 in the former Soviet Union under the guidance of Eugene Dynkin, who contributed the first chapter in this volume, and ever since that time Sergei has been the main research collaborator of his former advisor. This extremely fruitful collaboration resulted in 17 papers so far, in premier journals in probability and functional analysis. Sergei joined the Department of Mathematics at the University of Colorado at Boulder in 1998.

Finally, we are grateful to the National Science Foundation for their support of the meeting and to Springer for inviting this proceedings volume into their series. We offer our apologies for the unusually long editing time.

Boulder, Colorado, USA — Janos Englander
Boulder, Colorado, USA — Brian Rider

# Contents



Professor Sergei Kuznetsov

# Markov Processes and Their Applications to Partial Differential Equations: Kuznetsov's Contributions

**E.B. Dynkin**

**Abstract** We describe some directions of research in probability theory and related problems of analysis to which S. E. Kuznetsov has made fundamental contributions.

A Markov process (understood as a random path $X_t, 0 \leq t < \infty$ such that past before $t$ and future after $t$ are independent given $X_t$) is determined by a probability measure $P$ on a path space. This measure can be constructed starting from a transition function and probability distribution of $X_0$ . For a number of applications, it is also important to consider a path in both, forward and backward directions which leads to a concept of dual processes. In 1973, Kuznetsov constructed, as a substitute for such a pair of processes, a single random process $(X_t, \mathbb{P})$ determined on a random time interval $(\alpha, \beta)$. The corresponding forward and backward transition functions define a dual pair of processes. A $\sigma$-finite measure $\mathbb{P}$ became, under the name "Kuznetsov measure," an important tool for research on Markov processes and their applications.

In 1980, Kuznetsov proved that every Markov process in a Borel state space has a transition function (a problem that was open for many years). In 1992, he used this result to obtain simple necessary and sufficient conditions for existence of a unique decomposition of excessive functions into extreme elements—a significant extension of a classical result on positive superharmonic functions.

Intimate relations between the Brownian motion and differential equations involving the Laplacian $\Delta$ were known for a long time. Applications of probabilistic tools to classical potential theory and to study of linear PDEs are more recent. Even more recent is application of such tools to nonlinear PDEs. In a series of publications, starting from 1994, Dynkin and Kuznetsov investigated a class of semilinear elliptic equations by using super-Brownian motion and more general measure-valued Markov processes called superdiffusions. The main directions

E.B. Dynkin (✉)
Professor Emeritus Department of Mathematics Malott Hall Cornell University,
Ithaca, NY 14853, 4201
e-mail: ebd1@cornell.edu

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*,
Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_1, 

of this work were (a) description of removable singularities of solutions and (b) characterization of all positive solutions. One of the principal tools for solving the second problem was the fine trace of a solution on the boundary invented by Kuznetsov.

The same class of semilinear equations was the subject of research by Le Gall who applied a path-valued process Brownian snake instead of the super-Brownian motion. A slightly more general class of equations was studied by analysts including H. Brezis, M. Marcus and L. Veron. In the opinion of Brezis: "it is amazing how useful for PDEs are the new ideas coming from probability. This is an area where the interaction of probability and PDEs is most fruitful and exiting".

**Keywords** Markov processes • Transition function • Excessive functions • Superdiffusions • Semilinear PDE • Fine trace • Kuznetsov measures

## 1 Dual Markov Processes and Kuznetsov Measures

The idea of duality plays an important role in application of stochastic analysis to potential theory and partial differential equations.

Markov processes $(X_t, P)$ and $(\hat{X}_t, \hat{P})$ with stationary transition functions $p, \hat{p}$ are in duality relative to a given $\sigma$-finite measure $m$ if

$$m(dx)p_t(x,dy) = m(dy)\hat{p}_t(y,dx) \tag{1}$$

The processes can be defined on different spaces $\Omega$ and $\hat{\Omega}$ and there exists no relation between their paths. However the real source of duality is the fact that each Markov process can be considered in two time directions (this is the way the Kolmogorov forward and backward differential equations are deduced).

An alternative approach was suggested in [Dyn72] (see also [Dyn75] and [Dyn76]): functions $p$ and $\hat{p}$ are interpreted as a forward and a backward transition functions of a single Markov process with random birth time $\alpha$ and death time $\beta$. This approach was applied also to nonstationary transition functions $p(s,x;t,dy), \hat{p}(s,x;t,dy)$ and measures $m$ depending on the time interval $(s,t)$. The process $X_t, t \in (\alpha, \beta)$ could be described by its two-dimensional distributions $m_{st}(dx,dy) = P\{\alpha < s, X_s \in dx, X_t \in dy, t < \beta\}$. The family $m_{st}$ (called determining function) satisfies certain conditions including a normalization condition that guarantees that $P$ is a probability measure. In the stationary case, this condition holds if $m$ is a probability measure invariant for both dual processes. Since the definition of duality requires only $\sigma$-finiteness of $m$, this is a too restrictive assumption. It was removed by Kuznetsov in [Kuz73]. Of course, the measure $\mathbb{P}$ corresponding to a $\sigma$-finite $m$ is only $\sigma$-finite. However considering a process $(X_t, \mathbb{P})$ with random birth and death times instead of a pair of dual Markov processes proved to be very useful in the theory of Markov processes and its applications. The name "Kuznetsov measure" for $\mathbb{P}$ is commonly used in the literature. In [Mey00],[1] P.-A. Meyer called its construction "the most remarkable result... of the Dynkin school which was in advance of its time."

## 2 Markov Property and the Existence of a Transition Function

A stochastic process $(X_t, P)$ is Markovian if, for every $t$, the past $X_s, s < t$ and future $X_u, u > t$ are independent given $X_t$. This property is preserved under the time reversal. However, interactions between Markov processes and mathematical analysis are based on a possibility to start a process from a given point $x$ of the state space $E$ at a given time $t$ which requires the existence of a transition function, that is of a family of (sub)-probability measures

$$p(s,x;t,\cdot) \quad s < t, x \in E, \tag{2}$$

subject to the Chapman–Kolmogorov equation, such that

$$P\{X_t \in \Gamma | X_s\} = p(s,x;t,\Gamma) \qquad P-a.s. \tag{3}$$

Markov property implies the existence of a quasi-transition function, subject to a weaker condition: the Chapman–Kolmogorov equation holds only off an exceptional set depending on three times $s < t < u$. A transition function exists for such important processes as diffusions and Markov chains. (In fact, these processes are defined in terms of certain transition functions.) A natural question—does it exist for any Markov process—was asked by Kolmogorov in 1930s. In [Wal72] Walsh proved that transition function exists for time reversed strong Markov right continuous processes.

In [Kuz80], Kuznetsov proved that Markov property implies the existence of transition function if the state space $E$ is Borel (i.e., a measurable space isomorphic to a Borel subset in a separable complete metric space). In [Kuz86] he established the existence of a stationary transition function for every stochastically continuous Markov process in $E$ with a stationary quasi-transition function and with one-dimensional distributions absolutely continuous with respect to a reference measure $\nu$. In [Kuz92], he applied this result to investigate an existence of a stationary transition function $\hat{p}_t(x,dy)$ dual to a given stationary transition function $p_t(x,dy)$ relative to a given measure $m$ that is related to $p$ by (1).

[1] Based on his talk given in 1997 at Université Pierre et Marie Curie (Paris VI).

The measure $\nu$ must be excessive.[2] Under this condition on $\nu$, it is sufficient that $p$ is normal[3] and separates points.[4]

A substitute for a transition function for a random field $X_t, t \in T$ is its specification defined as a family of compatible conditional distributions for a value of the field given its restriction to subsets of the parameter set $T$. The existence of a specification for a wide class of random fields was proved by Kuznetsov in [Kuz84].

## 3 Existence and Uniqueness of Decomposition of Excessive Functions into Extremes

A class of positive functions $f$ similar to the class of classical positive superharmonic functions is associated with every stationary Markov transition function $p$. This class, called excessive functions, is defined by the conditions

$$\int_E p_t(x,dy)f(y) \leq f(x) \quad \text{for every } x \in E \, , \int_E p_t(x,dy)f(y) \uparrow f(x) \quad \text{as } t \downarrow 0. \tag{4}$$

Every positive superharmonic function $f$ in a smooth domain $D \subset R^d$ has a unique representation

$$f(x) = \int_D g(x,y)\mu(dy) + \int_{\partial D} k(x,y)\nu(dy) \tag{5}$$

where $g(x,y)$ is the Green function, $k(x,y)$ is the Poisson kernel and $\mu, \nu$ are finite measures. To every $x \in D$ there corresponds an extreme function $g_x(\cdot) = g(x,\cdot)$ and to every $x \in \partial D$ there corresponds an extreme function $k_x(\cdot) = k(x,\cdot)$. In the early 1940s, Martin [Mar41] deduced a similar formula for non-smooth domains by replacing $\partial D$ and $k(x,y)$ by what we call now the Martin boundary and the Martin kernel. In the 1950s Doob [Doo59] interpreted Martin's results in terms of the final behavior of the Brownian motion in $D$ and he obtained analogous results for Markov chains in a countable state space. In the 1960s Hunt [Hun68] deduced the Martin representation of excessive functions by investigating the initial behavior of approximate Markov chains—a concept introduced by him for this purpose. In 1969 Dynkin [Dyn69a, Dyn69b] suggested to use instead the space of paths $X_t$ defined for all positive and negative integers $t$ and a $\sigma$-finite measure $\mathbb{P}$ on this space that can be considered as a precursor of the Kuznetsov measure. Martin boundary theory was extended to some classes of Markov processes with continuous time parameter by Kunita and T. Watanabe [KW65].

[2] Which means for every $\Gamma, t$, $\int_E \nu(dx)p_t(x,\Gamma) \leq \nu(\Gamma)$ and for every $\Gamma$, $\int_E \nu(dx)p_t(x,\Gamma) \uparrow \nu(\Gamma)$ as $t \downarrow 0$.

[3] That is $p_t(x,E) \uparrow 1$ as $t \downarrow 0$ for every $x \in E$.

[4] That is, if $p(\cdot,x,\cdot) = p(\cdot,y,\cdot)$, then $x = y$.

The problem—to describe the most general conditions on a Markov process under which the related excessive functions have a unique decomposition into extreme elements—was discussed in several papers of Dynkin. In his talk at the International Congress of Mathematicians (Nice, 1970) he suggested to start from the description of time-dependent excessive measures (also called entrance laws). He proved in [Dyn72] that such measures can be uniquely decomposed into extremes and he deduced from here an analogous property of time-dependent excessive functions assuming the existence of a transition density. In 1974, Kuznetsov [Kuz74] established: if this condition is violated, then there exists a time-dependent excessive function not decomposable into extremes. In [Kuz92b], Kuznetsov combined these ideas with the existence of a dual transition function proved in [Kuz92] to get the following final result. Every $\gamma$-integrable excessive function has a unique decomposition into extremes if and only if, for every $x \in E$, the measure

$$U_x(\cdot) = \int_0^\infty e^{-t} p_t(x,\cdot)\, dt, \quad x \in E$$

is absolutely continuous with respect to $U_\gamma(\cdot) = \int_E \gamma(dx) U_x(\cdot)$.

## 4 Superdiffusions and Semilinear PDEs

A diffusion is a model of a random motion of a single particle. It is characterized by a second-order elliptic differential operator $L$. A special case is the Brownian motion corresponding to the Laplacian $\Delta$. A superdiffusion describes a random evolution of a cloud of particles. It is closely related to equations involving an operator $Lu - \psi(u)$. Here $\psi$ belongs to a class of functions which contains, in particular $\psi(u) = u^\alpha$ with $\alpha > 1$. Fundamental contributions to the analytic theory of equations

$$Lu = \psi(u) \tag{6}$$

and the corresponding parabolic equations were made by Keller [Kel57], Osserman [Oss57], Brezis and Strauss [BS], Loewner and Nirenberg [LN74], Brezis and Véron [BV], Baras and Pierre [BP84, BP84b], and Marcus and Véron [MV, MV98b, MV07]. A relation between the Eq. (6) and superdiffusions was established, first, by S. Watanabe [Wat68]. Dawson [Daw75] and Perkins [Per89, Per91] obtained deep results on the path behavior of the super-Brownian motion.

These equations were investigated in a series of papers of Dynkin and Kuznetsov by a combination of probabilistic and analytic tools. A systematic presentation of the results can be found in monographs [Dyn02] and [Dyn05]. In our exposition we concentrate on a special case of equations

$$Lu = u^\alpha, \qquad 1 < \alpha \le 2 \tag{7}$$

in a bounded smooth domain $D \subset R^d$.

### 4.1 Removable Boundary Singularities

A subset $\Gamma$ of the boundary $\partial D$ is called a removable boundary singularity if every positive solution of (7) in $D$ that is equal to zero on the boundary off $\Gamma$ is equal to 0 in $D$. In [GV91], Gmira and Véron proved that single points are removable if and only if the dimension $d \geq \frac{\alpha+1}{\alpha-1}$. Later in [Le94], Le Gall gave a complete characterization of the class of removable sets in the case $\alpha = 2$ and $L = \Delta$. Namely, a set is removable if and only if it has certain capacity zero. Probabilistically, it is removable if and only if it is polar for the corresponding superdiffusion $X$.[5]

In [DK96], Dynkin and Kuznetsov established similar result for an arbitrary $\alpha \in (1,2]$. The proof consists of two parts. The first is probabilistic. For a set $\Gamma$ with nonzero capacity we construct a required solution as a log-potential of a certain linear additive functional of $X$. The more difficult is the second analytic part: to show that, if the capacity of $\Gamma$ is 0, then every solution $u \geq 0$ such that $u = 0$ on $\partial D$ off $\Gamma$ vanishes in $D$. Here analytic tools used by Le Gall are insufficient[6] and need to be supplemented by a new construction suggested by Kuznetsov. It involves a sequence of truncating functions $h_n$ vanishing near $\Gamma$ , with a Sobolev-type norm tending to zero. By using an estimate of $||uh_n||_\alpha$ we get that $u = 0$.

### 4.2 Positive Solutions and Their Boundary Traces

Every positive solution of a linear equation $Lu = 0$ in a bounded smooth domain $D$ has a unique representation

$$u(x) = \int_{\partial D} k(x,y)\nu(dy) \tag{8}$$

where $k$ is the Poisson kernel and $\nu$ is a finite measure on $\partial D$. We call $\nu$ the boundary trace of $u$.[7] Formula (8) establishes a 1–1 correspondence between the set of all positive solutions and the set of all finite measures on $\partial D$.

The situation is much more complicated for Eq. (7). A positive solution can explode on a part of the boundary and even if this part is empty, not every finite measure can serve as the boundary trace. A description of all positive solutions in the case $L = \Delta, \alpha = 2, d = 2$ was given by Le Gall in [Le97]. He defined a trace of $u$ as a pair $(\Gamma, \nu)$ where $\Gamma$ is a closed subset of $\partial D$ on which $u$ blows up rapidly, and $\nu$ is a $\sigma$-finite measure on $\partial D \setminus \Gamma$. He also described the class of pairs $(\Gamma, \nu)$ that are in 1–1 correspondence with solutions.

---

[5] Intuitively, it is not hit by the random cloud modeled by $X$.

[6] Conditions $\alpha = 2, L = \Delta$ simplify the situation drastically.

[7] If $u$ is continuous in $D \cup \partial D$, then its boundary value is the density of $\nu$ with respect to the surface area.

The definition of trace in [Le97] is applicable for all dimensions $d$ however, if $\alpha = 2$ and $d > 2$, then there exist infinite many solutions with the same trace. A principal difference between these cases is that there exist no polar sets except the empty set in the first case and such sets exist in the second case. For Eq. (7), the first case (called subcritical) takes place if and only if $\alpha < (d+1)/(d-1)$. In this case positive solutions can be characterized by a boundary trace introduced in [DK98a] and in a slightly different form in [MV]. However, this is not true for $\alpha \geq (d+1)/(d-1)$. A breakthrough in investigating the latter case was an idea of a fine trace suggested by Kuznetsov in [Kuz98]. There he defined a special topology on the boundary associated with the equation (the so-called fine topology), and he constructed a "fine trace" $(\Gamma, \nu)$ where the set $\Gamma$ is finely closed and $\nu$ is still a $\sigma$-finite measure that does not charge polar sets. He then described the class of possible fine traces and showed that a wide class of solutions, called $\sigma$-moderate solutions, can be uniquely characterized by their fine traces. In [DK98] these results were extended to Eq. (6) under weak conditions on a monotone increasing convex nonlinear term $\psi(u)$. The crucial problem: "are all solutions $\sigma$-moderate?" (stated in the Epilog in [Dyn02]) was solved positively by Mselati in [Ms04] for $L = \Delta, \alpha = 2$ by using the Brownian snake. In a series of articles by Dynkin and Kuznetsov, Mselati's result was extended, by using superdiffusions instead of the Brownian snake to the equation $\Delta u = u^{\alpha}$ with $1 < \alpha \leq 2$. A self-contained presentation of the proofs can be found in [Dyn05].

## References


BP84. P. Baras and M. Pierre, *Problems paraboliques semi-linéares avec donnees measures*, Applicable Analysis **18** (1984), 111–149.

BP84b. ——, *Singularités éliminable pour des équations semi-linéares*, Ann. Inst. Fourier Grenoble **34** (1984), 185–206.

BS. H. Brezis and V. Strauss, *Semilinear second-order Elliptic equations in $L^1$*, J. Math. Soc. Japan **25** (1973), 565–590.

BV. H. Brezis and L. Véron, *Removable singularities of some removable equations*, Arch. Rat. Mech. Anal. **75** (1980), 1–6.

Daw75. D. A. Dawson, *Stochastic evolution equations and related measure processes*, J. Multivariate Anal. **3** (1975), 1–52.

Doo59. J. L. Doob, *Discrete Potential Theory and Boundaries*, J. Math. Mech **8** (1959), 433–458.

Dyn69a. E. B. Dynkin, *The boundary theory for Markov processes (discrete case)*, Russian Math. Surveys **24, 2** (1969), 89–157.

Dyn69b. ——, *Exit space of a Markov process*, Russian Math. Surveys **24, 4** (1969), 89–152.

Dyn72. ——, *Integral representation of excessive measures and excessive functions*, Russian Math. Surveys **27,1** (1972), 43–84.

Dyn75. ——, *Markov representation of stochastic systems*, Russian Math. Surveys **30,1** (1975), 64–104.

Dyn76. ——, *On a new approach to Markov processes*, Proceedings of the Third Japan-USSR Symposium on Probability Theory (1976), 42–62. Lecture Notes in Mathematics, vol. 550.

Dyn02. ——, *Diffusions, Superdiffusions and Partial Differential Equations*, 2002. American Math Society, Colloquium Publications, Vol. 50.

Dyn05. ——, *Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations*, 2004. American Math Society, University Lecture Series, Vol. 34.

DK96. E. B. Dynkin and S.E. Kuznetsov, *Superdiffusions and removable singularities for quasilinear partial differential equations*, Comm. Pure Appl.Math. **49** (1996), 125–176.

DK98a. ——, *Trace on the boundary for solutions of nonlinear partial differential equations*, Trans. Amer. Math. Soc. **350** (1998), 4499–4519.

DK98. ——, *Fine topology and fine trace on the boundary associated with a class of quasilinear partial differential equations*, Comm. Pure Appl.Math. **51** (1998), 897–936.

GV91. A. Gmira and L. Véron, *Boundary singularities of solutions of some nonlinear elliptic equations*, Duke Math. J. **64** (1991), 271–324.

Hun68. G. A. Hunt, *Markov Chains and Martin Boundaries*, Illinois J. Math. **4** (1968), 233–278, 365–410.

Kel57. J. B. Keller, *On the solutions of* $\Delta u = f(u)$, Comm. Pure Appl. Math. **10** (1957), 503–510.

KW65. H. Kunita and T. Watanabe, *Markov processes and Martin boundaries*, Illinois J. Math. **9** (1965), 386–391.

Kuz73. S. E. Kuznetsov, *Construction of Markov processes with random birth and death times*, Teoriya Veroyatn. i ee Primen. **18** (1973), 596–601. English Transl. in Theor. Prob. Appl., Vol 18 (1974).

Kuz74. ——, *On decomposition of excessive functions*, Dokl. AN SSSR **214** (1974), 276–278. English Transl. in Soviet Math. Dokl., Vol 15 (1974).

Kuz80. ——, *Any Markov process in a Borel space has a transition function*, Teoriya Veroyatn. i ee Primen. **25** (1980), 389–393. English Transl. in Theor. Prob. Appl., Vol 25 (1980).

Kuz84. ——, *Specifications and stopping theorem for random fields*, Teoriya Veroyatn. i ee Primen. **29** (1984), 65–77. English Transl. in Theor. Prob. Appl., Vol 29 (1985).

Kuz86. ——, *On the existence of a homogeneous transition function*, Teoriya Veroyatn. i ee Primen. **31** (1986), 290–300. English Transl. in Theor. Prob. Appl., Vol 31 (1987).

Kuz92. ——, *More on existence and uniqueness of decomposition of excessive functions and measures to extremes*, Seminaire de Probabilites XXVI. Lecture Notes in Mathematics **1526** (1992), 445–472.

Kuz92b. ——, *On existence of a dual semigroup*, Seminaire de Probabilites XXVI. Lecture Notes in Mathematics **1526** (1992), 473–484.

Kuz98. ——, $\sigma$*-moderate solutions of* $Lu = u^\alpha$ *and fine trace on the boundary*, C.R. Acad. Sci. Paris, Série I **326** (1998), 1189–1194.

Oss57. R. Osserman, *On the inequality* $\Delta u \geq f(u)$, *Pacific J. Math.* **7** (1957), 1641–1647.

Le94. J.-F. Le Gall, *The Brownian snake and solutions of* $\Delta u = u2$ *in a domain*, Probab. Theory Rel. Fields **102** (1995), 393–402.

Le97. ——, *A probabilistic Poisson representation for positive solutions of* $\Delta u = u2$ *in a planar domain*, Comm. Pure Appl. Math. **50** (1997), 69–103.

LN74. C. Loevner and L. Nirenberg, *Partial differential equations invariant under conformal or projective transformations*, (1974), 245–272. Contributions to Analysis, Academic Press, Orlando, FL.

Mar41. R. S. Martin, *Minimal positive harmonic functions*, Trans. Amer. Math. Soc. **49** (1941), 137–172.

Mey00. P.-A. Meyer, *Dynkin and theory of Markov processes*, Selected Papers of E. B. Dynkin with Commentary (2000), 763–772. Amer. Math. Soc., Providence, RI.

Ms04. B. Mselaty, *Classification and probabilistic representation of the positive solutions of a semilinear elliptic equation*, Mem. Amer. Math. Soc. **168** (2004).

MV. M. Marcus and L. Véron, *The boundary trace of positive solutions of semilinear elliptic equations, I. The subcritical case*, Arch. Rat. Mech. Anal. **144** (1998), 201–231.

MV98b. ——, *The boundary trace of positive solutions of semilinear elliptic equations: The supercritical case*, J. Math. Pures Appl. **77** (1998), 481–524.

MV07. ——, *The precise boundary trace of positive solutions of* $\Delta u = u^q$ *in the supercritical case*, **446** (2007), 345–383. Perspectives in nonlinear partial differential equations, Contemp. Math., Amer. Math. Soc., Providence, RI.

Per89. E. A. Perkins, *The Hausdorff measure of the closed support of super-Brownian motion*, Ann. Inst. H. Poincaré **25** (1989), 205–224.

Per91. ——, *On the continuity of measure-valued processes*, (1991), 261–268. Seminar on Stochastic Processes, Boston/Basel/Berlin.

Wal72. J. B. Walsh, *Transition functions of Markov processes*, Sminaire de Probabilits **VI** (1972), 215–232. Lecture Notes in Math., Vol. 258, Springer, Berlin, 1972.

Wat68. S. Watanabe, *A limit theorem on branching processes and continuous state branching processes*, J. Math. Kyoto Univ. **8** (1968), 141–167.

# Stochastic Equations on Projective Systems of Groups

**Steven N. Evans and Tatyana Gordeeva**

**Abstract** We consider stochastic equations of the form $X_k = \phi_k(X_{k+1})Z_k$, $k \in \mathbb{N}$, where $X_k$ and $Z_k$ are random variables taking values in a compact group $G_k$, $\phi_k : G_{k+1} \to G_k$ is a continuous homomorphism, and the noise $(Z_k)_{k\in\mathbb{N}}$ is a sequence of independent random variables. We take the sequence of homomorphisms and the sequence of noise distributions as given and investigate what conditions on these objects result in a unique distribution for the "solution" sequence $(X_k)_{k\in\mathbb{N}}$ and what conditions permit the existence of a solution sequence that is a function of the noise alone (i.e., the solution does not incorporate extra input randomness "at infinity"). Our results extend previous work on stochastic equations on a single group that was originally motivated by Tsirelson's example of a stochastic differential equation that has a unique solution in law but no strong solutions.

**Keywords** Group representation • Uniqueness in law • Strong solution • Extreme point • Lucas theorem • Toral automorphism

**Mathematics Subject Classification (2010):** 60B15, 60H25.

---

SNE supported in part by NSF grant DMS-0907630. TG supported in part by a VIGRE grant awarded to the Department of Statistics, University of California at Berkeley.

S.N. Evans (✉) • T. Gordeeva
Department of Statistics #3860, University of California at Berkeley,
367 Evans Hall, Berkeley, CA 94720-3860, U.S.A
e-mail: evans@stat.Berkeley.EDU; gordeeva@stat.Berkeley.EDU

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_2, 

## 1 Introduction

The following stochastic process was considered by Yor in [Yor92] in order to clarify the structure underpinning Tsirelson's celebrated example [Cir75] of a stochastic differential equation that does not have a strong solution even though all solutions have the same law.

Let $\mathbb{T}$ be the usual circle group; that is, $\mathbb{T}$ can be thought of as the interval $[0,1)$ equipped with addition modulo 1. Suppose for each $k \in \mathbb{N}$ that $\mu_k$ is a Borel probability measure on $\mathbb{T}$. Write $\mu = (\mu_k)_{k\in\mathbb{N}}$. We say that sequence of $\mathbb{T}$-valued random variables $(X_k)_{k\in\mathbb{N}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *solves the stochastic equation associated with* $\mu$ if

$$\mathbb{P}[f(X_k) \,|\, (X_j)_{j>k}] = \int_{\mathbb{T}} f(X_{k+1} + z)\, \mu_k(dz)$$

for all bounded Borel function $f : \mathbb{T} \to \mathbb{R}$, where we use the notation $\mathbb{P}[\cdot\,|\,\cdot]$ for condition expectations with respect to $\mathbb{P}$. In other words, if for each $k \in \mathbb{N}$ we define a $\mathbb{T}$-valued random variable $Z_k$ by requiring

$$X_k = X_{k+1} + Z_k, \tag{1}$$

then $(X_k)_{k\in\mathbb{N}}$ solves the stochastic equation associated with $\mu$ if and only if for all $k \in \mathbb{N}$ the distribution of $Z_k$ is $\mu_k$ and $Z_k$ is independent of $(X_j)_{j>k}$.

Yor addressed the existence of solutions $(X_k)_{k\in\mathbb{N}}$ that are *strong* in the sense that the random variable $X_k$ is measurable with respect to $\sigma((Z_j)_{j\geq k})$ for each $k \in \mathbb{N}$; that is, speaking somewhat informally, a solution is strong if it can be reconstructed from the "noise" $(Z_j)_{j\in\mathbb{N}}$ without introducing additional randomness "at infinity." It turns out that strong solutions exist if and only if

$$\lim_{m\to\infty} \lim_{n\to\infty} \prod_{\ell=m}^{n} \left| \int_{\mathbb{T}} \exp(2\pi i h x)\, \mu_\ell(dx) \right| > 0$$

for all $h \in \mathbb{Z}$ or, equivalently,

$$\sum_{k=1}^{\infty} \left[ 1 - \left| \int_{\mathbb{T}} \exp(2\pi i h x)\, \mu_k(dx) \right| \right] < \infty.$$

Yor's investigation was extended in [AUY08], where the group $\mathbb{T}$ is replaced by an arbitrary, possibly non-abelian, compact Hausdorff group. As one would expect, the role of the complex exponentials $\exp(2\pi i h \cdot)$, $h \in \mathbb{Z}$, in this more general setting is played by group representations. Interesting new phenomena appear when the group is non-abelian due to the fact that there are irreducible representations which are no longer one-dimensional. Several of the results in [AUY08] are framed in terms of properties of the set of extremal solutions (i.e., solutions that cannot be written as mixtures of others), and the structure of such solutions was elucidated further in [HY10].

We further extend the work in [Yor92, AUY08] by considering the following more general setup.

Fix a sequence $(G_k)_{k\in\mathbb{N}}$ of compact Hausdorff groups with countable bases. Suppose for each $k \in \mathbb{N}$ that there is a continuous homomorphism $\phi_k : G_{k+1} \to G_k$. Define a compact subgroup $H \subseteq G := \prod_{k\in\mathbb{N}} G_k$ by

$$H := \{g = (g_k)_{k\in\mathbb{N}} \in G : g_k = \phi_k(g_{k+1}) \text{ for all } k \in \mathbb{N}\}, \tag{2}$$

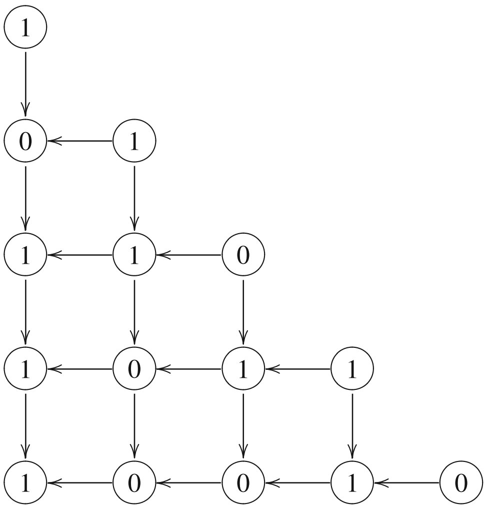
For example, if we take $G_k = \mathbb{T}$ for all $k \in \mathbb{N}$, then the homomorphism $\phi_k$ is necessarily of the form $\phi_k(x) = N_k x$ for some $N_k \in \mathbb{Z}$ and

$$H = \{g = (g_k)_{k\in\mathbb{N}} \in G : g_k = N_k g_{k+1} \text{ for all } k \in \mathbb{N}\}.$$

For a more interesting example, fix a compact group abelian group $\Gamma$, put $G_k := G_{1,k} \times G_{2,k-1} \cdots \times G_{k,1}$, where each group $G_{i,j}$ is a copy of $\Gamma$, and define the homomorphism $\phi_k$ by

$$\phi_k(g_{1,k+1}, g_{2,k}, \ldots, g_{k+1,1}) := (g_{1,k+1} + g_{2,k}, g_{2,k} + g_{3,k-1}, \ldots, g_{k,2} + g_{k+1,1})$$

(where we write the group operation in $\Gamma$ additively). Note that in this case $H$ is isomorphic to the infinite product $\Gamma^{\mathbb{N}}$, because an element $h = (h_{i,j})_{(i,j)\in\mathbb{N}\times\mathbb{N}}$ is uniquely specified by the values $(h_{i,1})_{i\in\mathbb{N}}$ and there are no constraints on these elements. The following picture shows a piece of an element of $H$ when $\Gamma$ is the group $\{0,1\}$ equipped with addition modulo 2.



Assume for each $k \in \mathbb{N}$ that $\mu_k$ is a Borel probability measure $G_k$ and write $\mu = (\mu_k)_{k\in\mathbb{N}}$. We say that sequence of random variables $(X_k)_{k\in\mathbb{N}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $X_k$ takes values in $G_k$, *solves the stochastic equation associated with* $\mu$ if

$$\mathbb{P}[f(X_k) \,|\, (X_j)_{j>k}] = \int_{G_k} f(\phi_k(X_{k+1})z)\, \mu_k(dz)$$

for all bounded Borel function $f : G_k \to \mathbb{R}$. In other words, if for each $k \in \mathbb{N}$ we define a $G_k$-valued random variable $Z_k$ by requiring

$$X_k = \phi_k(X_{k+1})Z_k, \tag{3}$$

then $(X_k)_{k\in\mathbb{N}}$ solves the stochastic equation if and only if for all $k \in \mathbb{N}$ the distribution of $Z_k$ is $\mu_k$ and $Z_k$ is independent of $(X_j)_{j>k}$. In particular, if $(X_k)_{k\in\mathbb{N}}$ solves the stochastic equation, then the sequence of random variables $(Z_k)_{k\in\mathbb{N}}$ is independent.

Certain special cases of this setup when $G_k = \Gamma$, $k \in \mathbb{N}$, for some fixed group $\Gamma$ and $\phi_k = \psi$, $k \in \mathbb{N}$ for a fixed automorphism $\psi : \Gamma \to \Gamma$ were considered in [Tak09, Raj11].

Note that whether or not a sequence $(X_k)_{k\in\mathbb{N}}$ solves the stochastic equation associated with $\mu$ is solely a feature of the distribution of the sequence, and so we say that a probability measure on the product group $\prod_{k\in\mathbb{N}} G_k$ is a solution of the stochastic equation if it is the distribution of a sequence that solves the equation and write $\mathcal{P}_\mu$ for the set of such measures.

In keeping with the terminology above, we say that a solution $(X_k)_{k\in\mathbb{N}}$ is *strong* if $X_k$ is measurable with respect to $\sigma((Z_j)_{j\geq k})$ for each $k \in \mathbb{N}$. Note that whether or not a solution is strong also depends only on its distribution, and so we define strong elements of $\mathcal{P}_\mu$ in the obvious manner and denote the set of such probability measures by $\mathcal{P}_\mu^{\text{strong}}$.

Because applying the homomorphism $\phi_k$ to $X_{k+1}$ can degrade the "signal" present in $X_{k+1}$ (e.g., $\phi_k$ need not be invertible), the question of whether or not strong solutions exist will involve the interaction between the homomorphisms $(\phi_k)_{k\in\mathbb{N}}$ and distributions $(\mu_k)_{k\in\mathbb{N}}$ of the noise random variables and it introduces new phenomena not present in [Yor92, AUY08].

An outline of the rest of this chapter is as follows. In Sect. 2 we examine the compact, convex set of solutions and show that strong solutions are extreme points of this set. We show that the subgroup $H$ acts transitively on the extreme points of the set of solutions and we relate the existence of strong solutions to properties of the set of extreme points. In Sect. 3, we obtain criteria for the existence of strong solutions in terms of the the representations of the group $G_k$ and the corresponding Fourier transforms of the probability measures $\mu_k$. In Sect. 3, we determine the relationship between the existence of strong solutions and the phenomenon of "freezing" wherein almost all sample paths of the random noise sequence agree with some sequence of constants for all sufficiently large indices. Finally, in Sects. 5 and 6, respectively, we investigate the example considered above of random variables indexed by the nonnegative quadrant of the two-dimensional integer lattice and another example where each group $G_k$ is the two-dimensional torus and each homomorphisms $\phi_k$ is a fixed ergodic toral automorphism.

## 2 Extreme Points of $\mathcal{P}_\mu$ and Strong Solutions

It is natural to first inquire whether $\mathcal{P}_\mu$ is nonempty and, if so, whether it consists of a single point that is, whether there exist probability measures that solve the stochastic equation associated with $\mu$ and, if so, whether there is a single such measure. The question of existence is easily disposed of by Proposition 2.1 below. Note that because the group $G = \prod_{k\in\mathbb{N}} G_k$ is compact and metrizable, the set of probability measures on $G$ equipped with the topology of weak convergence is also compact and metrizable.

**Proposition 2.1.** *For any sequence $\mu$, the set $\mathcal{P}_\mu$ is non-empty.*

*Proof.* Construct on some probability space a sequence $(Z_k)_{k\in\mathbb{N}}$ of independent random variables such that $Z_k$ has distribution $\mu_k$. For each $N \in \mathbb{N}$, define random variables $X_1^{(N)},\ldots,X_{N+1}^{(N)}$ recursively by

$$X_{N+1}^{(N)} := e_{N+1} := \text{identity in } G_{N+1}$$

and

$$X_k^{(N)} = \phi_k(X_{k+1}^{(N)})Z_k, \quad 1 \le k \le N,$$

so that for $1 \le k \le N$ the random variable $\phi_k(X_{k+1}^{(N)})^{-1}X_k^{(N)}$ has distribution $\mu_k$ and is independent of $X_{k+1}^{(N)}, X_{k+2}^{(N)},\ldots,X_N^{(N)}$.

Write $\mathbb{P}_N$ for the distribution of the sequence $(X_1^{(N)},\ldots,X_N^{(N)},e_{N+1},e_{N+2},\ldots)$. Because the space of probability measures on the group $\prod_{k\in\mathbb{N}} G_k$ equipped with the weak topology is compact and metrizable, there exists a subsequence $(N_n)_{n\in\mathbb{N}}$ and a probability measure $\mathbb{P}_\infty$ such that $\mathbb{P}_{N_n} \to \mathbb{P}_\infty$ weakly as $n \to \infty$. It is clear that $\mathbb{P}_\infty \in \mathcal{P}_\mu$. □

The question of uniqueness (i.e., whether or not $\#\mathcal{P}_\mu = 1$) is more demanding and will occupy much of our attention in the remainder of this chapter.

As a first indication of what is involved, consider the case where each measure $\mu_k$ is simply the unit point mass at the identity $e_k$ of $G_k$. In this case $(X_k)_{k\in\mathbb{N}}$ solves the stochastic equation if $X_k = \phi_k(X_{k+1})$ for all $k \in \mathbb{N}$. Recall the definition of the compact subgroup $H \subseteq G := \prod_{k\in\mathbb{N}} G_k$ from (2). It is clear that $\mathcal{P}_\mu$ coincides with the set of probability measures that are supported on $H$, and hence $\#\mathcal{P}_\mu = 1$ if and only if $H$ consists of just the single identity element. Note that if $\#H > 1$ and $(X_k)_{k\in\mathbb{N}}$ is a solution with distribution $\mathbb{P} \in \mathcal{P}_\mu$ that is not a point mass, then $X_k$ is certainly not a function of $(Z_j)_{j\ge k} = (e_j)_{j\ge k}$ and the solution $(X_k)_{k\in\mathbb{N}}$ is not strong. Moreover, the probability measures $\mathbb{P} \in \mathcal{P}_\mu$ that are distributions of strong solutions $(X_k)_{k\in\mathbb{N}}$ are the point masses at elements of $H$ and $\mathcal{P}_\mu$ is the closed convex hull of this set of measures.

An elaboration of the argument we have just given establishes the following result.

**Proposition 2.2.** *If $H$ is nontrivial (i.e, contains elements other than the identity), then $\mathcal{P}_\mu \setminus \mathcal{P}_\mu^{\mathrm{strong}} \neq \emptyset$. In particular, if $H$ is nontrivial and $\#\mathcal{P}_\mu = 1$, then $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$.*

*Proof.* Suppose that all solutions are strong. Let $(X_k)_{k\in\mathbb{N}}$ be a strong solution.

By extending the underlying probability space if necessary, construct an $H$-valued random variable $(U_k)_{k\in\mathbb{N}}$ that is independent of $(X_k)_{k\in\mathbb{N}}$ and is not almost surely constant. Note that $(U_k)_{k\in\mathbb{N}}$ is not $\sigma((X_k)_{k\in\mathbb{N}})$-measurable and hence, *a fortiori*, $(U_k)_{k\in\mathbb{N}}$ is not $\sigma((Z_k)_{k\in\mathbb{N}})$-measurable.

Observe that

$$\phi_k(U_{k+1}X_{k+1})Z_k = \phi_k(U_{k+1})\,\phi_k(X_{k+1})Z_k = U_k X_k,$$

because $\phi_k(U_{k+1}) = U_k$ for all $k \in \mathbb{N}$ by definition of $H$. Hence, $(U_k X_k)_{k\in\mathbb{N}}$ is also a solution. Thus, $(U_k X_k)_{k\in\mathbb{N}}$ is a strong solution by our assumption that all solutions are strong. In particular, $U_k X_k$ is $\sigma((Z_j)j \geq k)$-measurable for all $k \in \mathbb{N}$. However, $U_k = (U_k X_k)X_k^{-1}$ is $\sigma((Z_j)_{j\geq k})$-measurable, and we arrive at a contradiction. □

*Remark 2.3.* Consider the particular setting of [AUY08], where $G_k = \Gamma$, $k \in \mathbb{N}$, for some fixed group $\Gamma$, each homomorphism $\phi_k$ is the identity, and $H = \{(g, g, \ldots) : g \in \Gamma\}$. In this case, one can choose the sequence $(U_k)_{k\in\mathbb{N}}$ in the proof of Proposition 2.2 to be $(U, U, \cdots)$, where $U$ is distributed according to Haar measure on $\Gamma$; that is, $(U_k)_{k\in\mathbb{N}}$ is distributed according to Haar measure on $H$. Each marginal distribution of the solution $(X_k)_{k\in\mathbb{N}}$ is then Haar measure on $G_k = \Gamma$. In our more general setting it will not generally be the case that if $(U_k)_{k\in\mathbb{N}}$ is distributed according to Haar measure on $H$, then $X_k$ will be distributed according to Haar measure on $G_k$ for each $k \in \mathbb{N}$. For example, fix a compact group $\Gamma$, put $G_k = \Gamma^{\mathbb{N}}$ for all $k \in \mathbb{N}$ and define $\phi_k : G_{k+1} \to G_k$ by $\phi_k(g_1, g_2, g_3, \ldots) = (g_1, g_1, g_2, g_2, g_3, g_3, \ldots)$ for all $k \in \mathbb{N}$. It is clear that $H = \{((g, g, \ldots), (g, g, \ldots), \ldots) : g \in \Gamma\}$, so that $\{x_k : (x_1, x_2, \ldots) \in H\} \subseteq G_k$ is just the diagonal subgroup $\{(g, g, \ldots) : g \in \Gamma\}$ of the group $G_k$. Hence, for example, if $\mu_k$ is the point mass at the identity of $G_k$ for each $k \in \mathbb{N}$, the possible solutions $(X_k)_{k\in\mathbb{N}}$ are just arbitrary random elements of $H$, and it is certainly not possible to construct a solution such that the marginal distribution of $X_k$ is Haar measure on $G_k$ for some $k \in \mathbb{N}$.

From now on, we let $X_k : G \to G_k$, $k \in \mathbb{N}$, denote the random variable defined by $X_k((x_j)_{j\in\mathbb{N}}) := x_k$ and define $Z_k : G \to G_k$, $k \in \mathbb{N}$, by $Z_k := \phi_n(X_{k+1})^{-1}X_k$.

**Notation 2.4.** Given a sequence of random variables $S = (S_1, S_2, \ldots)$ and $k \in \mathbb{N}$, set $\mathcal{F}_k^S := \sigma((S_j)_{j\geq k})$. Similarly, set $\mathcal{F}^S := \mathcal{F}_1^S$ and $\mathcal{F}_\infty^S := \bigcap_{k\in\mathbb{N}} \mathcal{F}_k^S$.

**Notation 2.5.** For any sequence $\mu = (\mu_k)_{k\in\mathbb{N}}$, the set of solutions $\mathcal{P}_\mu$ is clearly a compact convex subset. Let $\mathcal{P}_\mu^{\mathrm{ex}}$ denote the extreme points of $\mathcal{P}_\mu$.

**Lemma 2.6.** *A probability measure $\mathbb{P} \in \mathcal{P}_\mu$ belongs to $\mathcal{P}_\mu^{\mathrm{ex}}$ if and only if the remote future $\mathcal{F}_\infty^X$ is trivial under $\mathbb{P}$.*

*Proof.* Our proof follows that of an analogous result in [AUY08].

Suppose that $\mathbb{P} \in \mathcal{P}_\mu$ and the $\sigma$-field $\mathcal{F}_\infty^X$ is not trivial under $\mathbb{P}$.

Fix a set $A \in \mathcal{F}_\infty^X$ with $0 < \mathbb{P}(A) < 1$. Then,

$$\mathbb{P}(\cdot) = \mathbb{P}(A)\mathbb{P}(\cdot\,|\,A) + (1-\mathbb{P}(A))\mathbb{P}(\cdot\,|\,A^c).$$

Observe that $\mathbb{P}(\cdot\,|\,A) \neq \mathbb{P}(\cdot\,|\,A^c)$, since $\mathbb{P}(A\,|\,A) = 1 \neq \mathbb{P}(A\,|\,A^c) = 0$.

Note for each $k \in \mathbb{N}$ and $B \subseteq G_k$ that

$$\begin{aligned}\mathbb{P}\{X_k\,\phi_k(X_{k+1})^{-1} \in B\,|\,A\} &= \frac{\mathbb{P}(\{X_k\,\phi_k(X_{k+1})^{-1} \in B\} \cap A)}{\mathbb{P}(A)} \\ &= \frac{\mu_k(B)\mathbb{P}(A)}{\mathbb{P}(A)} = \mu_k(B)\end{aligned}$$

because $\mathbb{P} \in \mathcal{P}_\mu$ and hence $X_k\,\phi_k(X_{k+1})^{-1}$ is independent of $\mathcal{F}_\infty^X$ under $\mathbb{P}$. Similarly, if $C \in \mathcal{F}_{k+1}^X$,

$$\begin{aligned}\mathbb{P}(\{X_k\,\phi_k(X_{k+1})^{-1} \in B\} \cap C\,|\,A) &= \frac{\mu_k(B)\mathbb{P}(C\cap A)}{\mathbb{P}(A)} \\ &= \mathbb{P}\{X_k\,\phi_k(X_{k+1})^{-1} \in B\,|\,A\}\mathbb{P}(C\,|\,A)\end{aligned}$$

Thus, $\mathbb{P}(\cdot\,|\,A) \in \mathcal{P}_\mu$. The analogous argument establishes $\mathbb{P}(\cdot\,|\,A^c) \in \mathcal{P}_\mu$. Since $\mathbb{P}(\cdot\,|\,A) \neq \mathbb{P}(\cdot\,|\,A^c)$, the probability measure $\mathbb{P}$ cannot belong to $\mathcal{P}_\mu^{\mathrm{ex}}$.

Now assume that $\mathbb{P} \in \mathcal{P}_\mu$ and $\mathcal{F}_\infty^X$ is trivial under $\mathbb{P}$. To show $\mathbb{P}$ is an extreme point, it suffices to show that if $\mathbb{P}' \in \mathcal{P}_\mu$ is absolutely continuous with respect to $\mathbb{P}$, then $\mathbb{P} = \mathbb{P}'$.

Note that a solution $X$ is a time-inhomogeneous Markov chain (indexed in backwards time with index set starting at infinity) with the following transition probability:

$$\mathbb{P}\{X_k \in A\,|\,X_{k+1}\} = \mu_k\{g \in G_k : \phi_k(X_{k+1})g \in A\}.$$

Since $\mathbb{P}$ and $\mathbb{P}'$ are the distributions of Markov chains with common transition probabilities and $\mathbb{P}'$ is absolutely continuous with respect to $\mathbb{P}$, it follows that for any measurable set $A$ the random variables $\mathbb{P}(A\,|\,\mathcal{F}_\infty^X)$ and $\mathbb{P}'(A\,|\,\mathcal{F}_\infty^X)$ are equal $\mathbb{P}$-a.s. Because $\mathcal{F}_\infty^X$ is trivial under both $\mathbb{P}$ and $\mathbb{P}'$, it must be the case that $\mathbb{P}(A) = \mathbb{P}'(A)$. □

**Corollary 2.7.** *All strong solutions $\mathbb{P} \in \mathcal{P}_\mu$ are extreme; that is, $\mathcal{P}_\mu^{\mathrm{strong}} \subseteq \mathcal{P}_\mu^{\mathrm{ex}}$.*

*Proof.* By definition, if $\mathbb{P} \in \mathcal{P}_\mu$ is strong, then $X_k \in \mathcal{F}_k^Z$ for all $k \in \mathbb{N}$. Thus, $\mathcal{F}_k^X = \mathcal{F}_k^Z$ for all $k \in \mathbb{N}$ and hence $\mathcal{F}_\infty^X = \mathcal{F}_\infty^Z$. The last $\sigma$-field is trivial by the Kolmogorov zero-one law. □

*Remark 2.8.* There can be extreme solutions that are not strong. For example, suppose that the $G_k = \Gamma$, $k \in \mathbb{N}$, for some nontrivial group $\Gamma$, each $\phi_k$ is the identity map, and each $\mu_k$ is the Haar measure on $\Gamma$. It is clear that $\mathcal{P}_\mu$ consists of just the measure $\bigotimes_{k\in\mathbb{N}} \mu_k$ (i.e., Haar measure on $G$), and so this solution is extreme. However, it follows from Proposition 2.2 that this solution is not strong.

It is clear that if $\mathbb{P} \in \mathcal{P}_\mu$ and $h = (h_k)_{k\in\mathbb{N}} \in H$, then the distribution of the sequence $(h_k X_k)_{k\in\mathbb{N}}$ also belongs to $\mathbb{P} \in \mathcal{P}_\mu$. Moreover, if $\mathbb{P} \in \mathcal{P}_\mu^{\mathrm{ex}}$, then it follows from Lemma 2.6 that the distribution of the sequence $(h_k X_k)_{k\in\mathbb{N}}$ also belongs to $\mathcal{P}_\mu^{\mathrm{ex}}$. Similarly, if $\mathbb{P} \in \mathcal{P}_\mu^{\mathrm{strong}}$, then the distribution of the sequence $(h_k X_k)_{k\in\mathbb{N}}$ also belongs to $\mathcal{P}_\mu^{\mathrm{strong}}$. We record these observations for future reference.

**Lemma 2.9.** *The collection of maps $T_h : \mathcal{P}_\mu \to \mathcal{P}_\mu$, $h \in H$, defined by $T_h(\mathbb{P})(\cdot) = \mathbb{P}\{(h_k X_k)_{k\in\mathbb{N}} \in \cdot\}$ constitute a a group action of $H$ on $\mathcal{P}_\mu$. The set $\mathcal{P}_\mu^{\mathrm{ex}}$ of extreme solutions and the set $\mathcal{P}_\mu^{\mathrm{strong}}$ of strong solutions are both invariant for this action.*

It follows from the next result that either $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$ or $\mathcal{P}_\mu^{\mathrm{strong}} = \mathcal{P}_\mu^{\mathrm{ex}}$. For the purposes of the proof and later it is convenient to introduce the following notation.

**Notation 2.10.** For $k, \ell \in \mathbb{N}$ with $k < \ell$, define $\phi_k^\ell : G_\ell \to G_k$ by

$$\phi_k^\ell = \phi_k \circ \phi_{k+1} \circ \cdots \circ \phi_{\ell-1},$$

and adopt the convention that $\phi_k^k$ is the identity map from $G_k$ to itself.

**Theorem 2.11.** *The group action $(T_h)_{h\in H}$ is transitive on $\mathcal{P}_\mu^{\mathrm{ex}}$.*

*Proof.* For $k \in \mathbb{N}$, define $X'_k : \prod_{k\in\mathbb{N}} (G_k \times G_k \times G_k) \to G_k$ (resp. $X''_k : \prod_{k\in\mathbb{N}} (G_k \times G_k \times G_k) \to G_k$) and $Y_k : \prod_{k\in\mathbb{N}} (G_k \times G_k \times G_k) \to G_k$) by $X'_k((x'_j, x''_j, y_j)_{j\in\mathbb{N}}) = x'_k$ (resp. $X''_k((x'_j, x''_j, y_j)_{j\in\mathbb{N}}) = x''_k$ and $Y_k((x'_j, x''_j, y_j)_{j\in\mathbb{N}}) = y_k$).

Suppose that $\mathbb{P}', \mathbb{P}'' \in \mathcal{P}_\mu$. Write $\mathbb{P}'_z(\cdot)$ (resp. $\mathbb{P}''_z(\cdot)$) for the regular conditional probability of $\mathbb{P}'\{X \in \cdot \,|\, Z = z\}$ (resp. $\mathbb{P}''\{X \in \cdot \,|\, Z = z\}$).

Define a probability measure $\mathbb{Q}$ on $\prod_{k\in\mathbb{N}} (G_k \times G_k \times G_k)$ by

$$\mathbb{Q}\{(X', X'', Y) \in A' \times A'' \times B\} = \int_G \mathbb{P}'_z(A') \mathbb{P}''_z(A'') 1_B(z) \, (\bigotimes_{k\in\mathbb{N}} \mu_k)(dz).$$

By construction, $\phi_k(X'_{k+1})^{-1} X'_k = \phi_k(X''_{k+1})^{-1} X''_k = Y_k$ for all $k \in \mathbb{N}$, $\mathbb{Q}$-a.s., the distribution of the pair $(X', Y)$ under $\mathbb{Q}$ is the same as that of the pair $(X, Z)$ under $\mathbb{P}'$, and the distribution of the pair $(X'', Y)$ under $\mathbb{Q}$ is the same as that of the pair $(X, Z)$ under $\mathbb{P}''$. In particular, the distributions of $X'$ and $X''$ under $\mathbb{Q}$ are, respectively, $\mathbb{P}'$ and $\mathbb{P}''$.

Suppose for some $k \in \mathbb{N}$ that $\Phi' : G \to \mathbb{R}$ and $\Phi'' : G \to \mathbb{R}$ are both bounded $\mathcal{F}^X_{k+1}$-measurable functions and $\Psi : G_k \to \mathbb{R}$ is a bounded Borel function. Then, $\Phi' \circ X' : \prod_{j\in\mathbb{N}}(G_j \times G_j \times G_j) \to \mathbb{R}$ is $\mathcal{F}^{X'}_{k+1}$-measurable and $\Phi'' \circ X'' : \prod_{j\in\mathbb{N}}(G_j \times G_j \times G_j) \to \mathbb{R}$ is $\mathcal{F}^{X''}_{k+1}$-measurable, and hence, by the construction of $\mathbb{Q}$ (using the notations $\nu[\cdot]$ and $\nu[\cdot\,|\,\cdot]$ for expectation and conditional expectation with respect to a probability measure $\nu$),

$$\begin{aligned}\mathbb{Q}[\Phi' \circ X' \, \Phi'' \circ X'' \,|\, \mathcal{F}^Y] &= \mathbb{Q}[\Phi' \circ X' \,|\, \mathcal{F}^Y]\, \mathbb{Q}[\Phi'' \circ X'' \,|\, \mathcal{F}^Y] \\ &= \mathbb{P}'_Y[\Phi' \circ X]\, \mathbb{P}''_Y[\Phi'' \circ X]\end{aligned}$$

is $\mathcal{F}^Y_{k+1}$-measurable. Thus, by the construction of $\mathbb{Q}$ and the independence of the elements of the sequence $(Y_j)_{j\in\mathbb{N}}$ under $\mathbb{Q}$,

$$\begin{aligned}\mathbb{Q}[\Phi' \circ X' \, \Phi'' \circ X'' \, \Psi \circ Y_k] &= \mathbb{Q}[\mathbb{Q}[\Phi' \circ X' \, \Phi'' \circ X'' \, \Psi \circ Y_k \,|\, \mathcal{F}^Y]] \\ &= \mathbb{Q}[\mathbb{Q}[\Phi' \circ X' \, \Phi'' \circ X'' \,|\, \mathcal{F}^Y]\Psi \circ Y_k] \\ &= \mathbb{Q}[\mathbb{P}'_Y[\Phi' \circ X]\, \mathbb{P}''_Y[\Phi'' \circ X]]\, \mathbb{Q}[\Psi \circ Y_k] \\ &= \mathbb{Q}[\Phi' \circ X' \, \Phi'' \circ X'']\, \mathbb{Q}[\Psi \circ Y_k].\end{aligned}$$

Therefore, by a standard monotone class argument, $Y_k$ is independent of $\mathcal{F}^{(X',X'')}_{k+1}$. Consequently, the sub-$\sigma$-fields $\mathcal{F}_Y$ and $\mathcal{F}^{(X',X'')}_\infty$ are independent.

Suppose now that $\mathbb{P}', \mathbb{P}'' \in \mathcal{P}^{\mathrm{ex}}_\mu$. Observe for $k < n$ that

$$\begin{aligned}X'_k(X''_k)^{-1} &= \left[\phi^n_k(X'_n) \prod_{m=k}^{n-1} \phi^m_k(Y_m)\, Y_k\right] \left[\phi^n_k(X''_n) \prod_{m=k}^{n-1} \phi^m_k(Y_m)\, Y_k\right]^{-1} \quad \mathbb{Q}-\text{a.s.} \\ &= \phi^n_k(X'_n)\, \phi^n_k(X''_n)^{-1},\end{aligned} \tag{4}$$

and so there exists a $G$-valued random variable $W \in \mathcal{F}^{X',X''}_\infty$ such that $W_k = X'_k(X''_k)^{-1}$, $\mathbb{Q}$-a.s. From the above, $W$ is independent of the sub-$\sigma$-field $\mathcal{F}_Y$. By construction, $W$ takes values in the subgroup $H$.

Let $\mathbb{Q}(\cdot\,|\,W = h)$ be the regular conditional probability for $\mathbb{Q}$ given $W = h \in H$, so that

$$\mathbb{Q}(\cdot) = \int_H \mathbb{Q}(\cdot\,|\,W = h)\, \mathbb{Q}\{W \in dh\}. \tag{5}$$

It follows that

$$\mathbb{Q}\{X'_k = \phi_k(X'_{k+1})\, Y_k, \forall k \in \mathbb{N} \,|\, W = h\} = 1$$

for $\mathbb{Q}\{W \in dh\}$-almost every $h \in H$. Moreover, because $W$ is independent of $\mathcal{F}_Y$ it follows that $\mathbb{Q}\{Y \in \cdot\} = \mathbb{Q}\{Y \in \cdot\,|\,W = h\} = \bigotimes_{k\in\mathbb{N}} \mu_k$ for $\mathbb{Q}\{W \in dh\}$-almost every $h \in H$. Thus, $\mathbb{Q}\{X' \in \cdot\,|\,W = h\} \in \mathcal{P}_\mu$ for $\mathbb{Q}\{\varepsilon \in dh\}$-almost every $h \in H$ and, by (5),

$$\mathbb{P}'(\cdot) = \mathbb{Q}\{X' \in \cdot\} = \int_H \mathbb{Q}\{X' \in \cdot\,|\,W = h\}\, \mathbb{Q}\{W \in dh\}.$$

This would contradict the extremality of $\mathbb{P}'$ unless

$$\mathbb{P}'(\cdot) = \mathbb{Q}\{X' \in \cdot \,|\, W = h\}, \text{ for } \mathbb{Q}\{W \in dh\}\text{-almost every } h \in H.$$

Similarly,

$$\mathbb{P}''(\cdot) = \mathbb{Q}\{X'' \in \cdot \,|\, W = h\}, \text{ for } \mathbb{Q}\{W \in dh\}\text{-almost every } h \in H.$$

By (4),

$$\mathbb{Q}\{X'_k = h_k X''_k \, \forall k \in \mathbb{N} \,|\, W = h\} = 1, \text{ for } \mathbb{Q}\{W \in dh\}\text{-almost every } h \in H.$$

Therefore,

$$\mathbb{P}' = T_h(\mathbb{P}''), \text{ for } \mathbb{Q}\{W \in dh\}\text{-almost every } h \in H. \qquad \square$$

**Notation 2.12.** Given $\mathbb{P}^0 \in \mathcal{P}_\mu^{\mathrm{ex}}$, let $H_\mu^{\mathrm{stab}}(\mathbb{P}^0) := \{h \in H : T_h(\mathbb{P}^0) = \mathbb{P}^0\}$ be the stabilizer subgroup of the point $\mathbb{P}^0$ under the group action $(T_h)_{h \in H}$.

*Remark 2.13.* It follows from the transitivity of $H$ on $\mathcal{P}_\mu^{\mathrm{ex}}$ that for any two probability measures $\mathbb{P}', \mathbb{P}'' \in \mathcal{P}_\mu^{\mathrm{ex}}$ the subgroups $H_\mu^{\mathrm{stab}}(\mathbb{P}')$ and $H_\mu^{\mathrm{stab}}(\mathbb{P}'')$ are conjugate.

**Corollary 2.14.** *A necessary and sufficient condition for* $\#\mathcal{P}_\mu = 1$ *is that* $H_\mu^{\mathrm{stab}}(\mathbb{P}^0) = H$ *for some, and hence all,* $\mathbb{P}^0 \in \mathcal{P}_\mu^{\mathrm{ex}}$.

*Proof.* This is immediate from Theorem 2.11 and the observation that $\#\mathcal{P}_\mu = 1$ if and only if $\#\mathcal{P}_\mu^{\mathrm{ex}} = 1$. $\square$

**Corollary 2.15.** *If* $H_\mu^{\mathrm{stab}}(\mathbb{P}^0)$ *is nontrivial for some, and hence all,* $\mathbb{P}^0 \in \mathcal{P}_\mu^{\mathrm{ex}}$, *then* $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$.

*Proof.* As we observed prior to the statement of Theorem 2.11, it is a consequence of that result that either $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$ or $\mathcal{P}_\mu^{\mathrm{strong}} = \mathcal{P}_\mu^{\mathrm{ex}}$.

Suppose that $\mathbb{P}^0 \in \mathcal{P}_\mu^{\mathrm{strong}}$ is such that $H_\mu^{\mathrm{stab}}(\mathbb{P}^0)$ is nontrivial. By working on an extended probability space, we may assume that there is an $H_\mu^{\mathrm{stab}}(\mathbb{P}^0)$-valued random variable $(U_k)_{k \in \mathbb{N}}$ that is independent of $(X_k)_{k \in \mathbb{N}}$ and is not almost surely constant. The distribution of the solution $(U_k X_k)_{k \in \mathbb{N}}$ is also $\mathbb{P}^0$ and, in particular, this solution is strong. However, this implies that

$$\begin{aligned}
\sigma(U_k X_k) &\subseteq \sigma((\phi_j(U_{j+1} X_{j+1})^{-1} U_j X_j)_{j \ge k}) \\
&= \sigma((\phi_j(X_{j+1})^{-1} X_j)_{j \ge k}) \\
&= \mathcal{F}_k^Z
\end{aligned}$$

for all $k \in \mathbb{N}$, and hence $U_k$ is $\mathcal{F}_k^Z$-measurable for all $k \in \mathbb{N}$, because $X_k$ is $\mathcal{F}_k^Z$-measurable by the assumption that $\mathbb{P}^0 \in \mathcal{P}_\mu^{\text{strong}}$. However, because the sequence $(U_k)_{k\in\mathbb{N}}$ is independent of the sequence of $(X_k)_{k\in\mathbb{N}}$ and not almost surely constant, it follows that $(U_k)_{k\in\mathbb{N}}$ is not $\sigma((X_k)_{k\in\mathbb{N}})$-measurable and hence *a fortiori* $(U_k)_{k\in\mathbb{N}}$ is not $\sigma((Z_k)_{k\in\mathbb{N}})$-measurable. We thus arrive at a contradiction. □

## 3 Representation Theory and the Existence of Strong Solutions

**Notation 3.1.** Let $\mathcal{G}$ be the set of all unitary, finite-dimensional representations of the compact group $G = \prod_{k\in\mathbb{N}} G_k$.

Any irreducible representations of $G$ are equivalent to a tensor product representation of the form

$$(g_k)_{k\in\mathbb{N}} \mapsto \rho^{(k_1)}(g_{k_1}) \otimes \cdots \otimes \rho^{(k_n)}(g_{k_n}),$$

where $\{k_1,\ldots,k_n\}$ is a finite subset of $\mathbb{N}$ and $\rho^{(k_j)}$ is an (necessarily finite-dimensional) irreducible representation of $G_{k_j}$ for $1 \le j \le n$. Furthermore, an arbitrary element of $\mathcal{G}$ is equivalent to a (finite) direct sum of irreducible representations.

**Notation 3.2.** For $k \in \mathbb{N}$ write $\iota_k : G_k \mapsto G$ for the map that sends $h \in G_k$ to $(e_1,\ldots,e_{k-1},h,e_{k+1},\ldots)$, where, as above, $e_j$ is the identity element of $G_j$ for $j \in \mathbb{N}$.

Consider an arbitrary representation $\rho \in \mathcal{G}$. It is clear from the above that if $\mathbb{P} \in \mathcal{P}_\mu^{\text{strong}}$, then $\rho \circ \iota_k(X_k)$ is $\mathcal{F}_k^Z$-measurable for all $k \in \mathbb{N}$. Note that $\rho \circ \iota_k$ is a representation of $G_k$ and all representations of $G_k$ arise this way. On the other hand, because, by the Peter–Weyl theorem, the closure in the uniform norm of the (complex) linear span of matrix entries of the irreducible representations of $G_k$ is the vector space of continuous complex-valued functions on $G_k$, it follows that if $\rho \circ \iota_k(X_k)$ is $\mathcal{F}_k^Z$-measurable for all $k \in \mathbb{N}$ for an arbitrary representation $\rho \in \mathcal{G}$, then $\mathbb{P} \in \mathcal{P}_\mu^{\text{strong}}$. This observation leads to the following definition and theorem.

**Notation 3.3.** Set

$$\mathcal{H}_\mu^{\text{strong}} := \{\rho \in \mathcal{G} : \exists \mathbb{P} \in \mathcal{P}_\mu^{\text{ex}} \text{ such that } \rho \circ \iota_k(X_k) \text{ is } \mathcal{F}_k^Z\text{-measurable } \mathbb{P}\text{-a.s. } \forall k \in \mathbb{N}\}.$$

**Theorem 3.4.** *The set $\mathcal{P}_\mu^{\text{strong}}$ of strong solutions is non-empty (and hence equal to $\mathcal{P}_\mu^{\text{ex}}$) if and only if $\mathcal{H}_\mu^{\text{strong}} = \mathcal{G}$.*

*Proof.* The result is immediate from the discussion preceding the statement of the theorem once we note that if $\mathbb{P}'$ and $\mathbb{P}''$ both belong to $\mathcal{P}_\mu^{\text{ex}}$ then, by Theorem 2.11, there exists $h \in H$ such that $\mathbb{P}''$ is the distribution of $hX = (h_k X_k)_{k\in\mathbb{N}}$ under $\mathbb{P}'$ and so $\rho \circ \iota_k(X_k)$ is $\mathcal{F}_k^Z$-measurable $\mathbb{P}''$-a.s. if and only if $\rho \circ \iota_k(h_k X_k)$ is $\mathcal{F}_k^Z$-measurable $\mathbb{P}'$-a.s. (recall that $Z_k = \phi(X_{k+1})^{-1} X_k = \phi(h_k X_{k+1})^{-1} h_k X_k$ when $h \in H$); therefore, $\rho \circ \iota_k(X_k)$ is $\mathcal{F}_k^Z$-measurable $\mathbb{P}''$-a.s. if and only if $[\rho \circ \iota_k(h_k)][\rho \circ \iota_k(X_k)]$ is $\mathcal{F}_k^Z$-measurable $\mathbb{P}'$-a.s., which is in turn equivalent to $\rho \circ \iota_k(X_k)$ being $\mathcal{F}_k^Z$-measurable $\mathbb{P}'$-a.s. by the invertibility of the matrix $\rho \circ \iota_k(h_k)$. Thus,

$$\mathcal{H}_\mu^{\text{strong}} = \{\rho \in \mathcal{G} : \ \rho \circ \iota_k(X_k) \text{ is } \mathcal{F}_k^Z\text{-measurable } \mathbb{P}\text{-a.s. } \forall k \in \mathbb{N} \}$$

for any $\mathbb{P} \in \mathcal{P}_\mu^{\text{ex}}$. □

Theorem 3.4 is still somewhat unsatisfactory as a criterion for the existence of strong solutions because it requires a knowledge of the set $\mathcal{P}_\mu^{\text{ex}}$ of extreme solutions. We would prefer a criterion that was directly in terms of the sequence $(\mu_k)_{k \in \mathbb{N}}$. In order to (partly) remedy this situation, we introduce the following objects.

**Notation 3.5.** Fix $\rho \in \mathcal{G}$. For $k, \ell \in \mathbb{N}$ with $k \le \ell$, set

$$R_k^\ell := \int_{G_\ell} \rho \circ \iota_k \circ \phi_k^\ell(z) \, \mu_\ell(dz).$$

Let

$$\mathcal{H}_\mu^{\text{det}} := \{\rho \in \mathcal{G} : \lim_{m \to \infty} \lim_{n \to \infty} \left|\det(R_k^n R_k^{n-1} \cdots R_k^m)\right| > 0 \ \forall k \in \mathbb{N}\}$$

and

$$\mathcal{H}_\mu^{\text{norm}} := \{\rho \in \mathcal{G} : \lim_{m \to \infty} \lim_{n \to \infty} \|R_k^n R_k^{n-1} \cdots R_k^m\| > 0 \ \forall k \in \mathbb{N}\},$$

where $\|\cdot\|$ is the $\ell^2$ operator norm on the appropriate space of matrices.

**Proposition 3.6.** *Fix* $\mathbb{P} \in \mathcal{P}_\mu$.

*(i) If* $\rho \in \mathcal{H}_\mu^{\text{det}}$, *then*

$$\mathbb{P}[\rho \circ \iota_k(X_k) \,|\, \mathcal{F}_\infty^X \vee \mathcal{F}_k^Z] = \rho \circ \iota_k(X_k)$$

*for all* $k \in \mathbb{N}$. *In particular, if* $\mathbb{P} \in \mathcal{P}_\mu^{\text{ex}}$, *then* $\rho \circ \iota_k(X_k)$ *is* $\mathcal{F}_k^Z$*-measurable for all* $k \in \mathbb{N}$.

*(ii) If* $\rho \notin \mathcal{H}_\mu^{\text{norm}}$, *then*

$$\mathbb{P}[\rho \circ \iota_k(X_k) \,|\, \mathcal{F}_\infty^X \vee \mathcal{F}_k^Z] = 0$$

*for some* $k \in \mathbb{N}$. *In particular, if* $\mathbb{P} \in \mathcal{P}_\mu^{\text{ex}}$, *then* $\rho \circ \iota_k(X_k)$ *is not* $\mathcal{F}_k^Z$*-measurable for some* $k \in \mathbb{N}$.

*Proof.* The proof follows that of an analogous result in [AUY08] with modifications required by the greater generality in which we are working.

Consider claim (i). Fix $\rho \in \mathcal{H}_\mu^{\text{det}}$ and $k \in \mathbb{N}$. For $\ell > k$ we have

$$\rho \circ \iota_k(X_k) = \rho \circ \iota_k \circ \phi_k^\ell(X_\ell) \, \rho \circ \iota_k \circ \phi_k^{\ell-1}(Z_{\ell-1}) \cdots \rho \circ \iota_k \circ \phi_k^k(Z_k). \tag{6}$$

For $k \le m \le n$ put

$$\Xi_n^m := \rho \circ \iota_k \circ \phi_k^n(Z_m) \cdots \rho \circ \iota_k \circ \phi_k^m(Z_m).$$

Note that

$$\mathbb{P}[\Xi_n^m] = R_k^n \cdots R_k^m.$$

For any $p \ge k$, the matrix $\rho \circ \iota_k \circ \phi_k^p$ is unitary, and so $\|\rho \circ \iota_k \circ \phi_k^p(h)\| = 1$ for all $h \in G_p$. By Jensen's inequality, $\|R_k^p\| \le 1$. In particular, $|\det(R_k^p)| \le 1$. Hence,

$$\lim_{m\to\infty} \lim_{n\to\infty} |\det(\mathbb{P}[\Xi_n^m])|$$

exists and is given by

$$\sup_m \inf_{n\ge m} |\det(R_k^n)| \cdots |\det(R_k^m)|.$$

Moreover, there are constants $\varepsilon > 0$ and $M \in \mathbb{N}$ such that $|\det(\mathbb{P}[\Xi_n^m])| \ge \varepsilon$ whenever $n \ge m \ge M$. It follows from Cramer's rule that the matrices $\mathbb{P}[\Xi_n^m]$ are invertible with uniformly bounded entries for $n \ge m \ge M$.

Set $\Phi_n^m := \mathbb{P}[\Xi_n^m]^{-1}\Xi_n^m$ for $n \ge m \ge M$. The matrices $\Phi_n^m$ have uniformly bounded entries and

$$\mathbb{P}\left[\Phi_{n+1}^m \,|\, \sigma((Z_p)_{p=m}^n)\right] = \Phi_n^m,$$

so that $(\Phi_n)_{n\ge m}$ is a bounded matrix-valued martingale with respect to the filtration $(\sigma((Z_p)_{p=m}^n))_{n\ge m}$. Thus, $\lim_{n\to\infty} \Phi_n^m =: \Phi_\infty^m$ exists and is $\mathcal{F}_m^Z$-measurable $\mathbb{P}$-a.s. for each $m \ge M$. Consequently, $\lim_{n\to\infty} \Xi_n^m =: \Xi_\infty^m$ also exists and is $\mathcal{F}_m^Z$-measurable $\mathbb{P}$-a.s. for each $m \ge M$. Part (i) is now clear from (6).

Now consider part (ii). Fix $\rho \notin \mathcal{H}_\mu^{\text{norm}}$ and $k \in \mathbb{N}$ such that

$$\lim_{m\to\infty} \lim_{n\to\infty} \left\| R_k^n R_k^{n-1} \cdots R_k^m \right\| = 0.$$

It follows from (6) that for $n \ge m \ge k$

$$\mathbb{P}\left[\rho \circ \iota_k(X_k) \,|\, \mathcal{F}_n^X \vee \sigma((Z_j)_{j=k}^m)\right] = \rho \circ \iota_k \circ \phi_k^n(X_n) R_k^{n-1} \cdots R_k^{m+1}$$
$$\rho \circ \iota_k \circ \phi_m^k(Z_m) \cdots \rho \circ \iota_k \circ \phi_k^k(Z_k).$$

Since $\rho(g)$ is a unitary matrix for all $g \in G$, the norm of the right-hand side is at most $\|R_k^{n-1} \cdots R_k^{m+1}\|$, which, by assumption, converges to 0 as $n \to \infty$ followed by $m \to \infty$. Thus, by the reverse martingale convergence theorem and the martingale convergence theorem,

$$\mathbb{P}\left[\rho \circ \iota_k(X_k) \,|\, \mathcal{F}_\infty^X \vee \mathcal{F}_k^Z\right] = \lim_{m\to\infty} \lim_{n\to\infty} \mathbb{P}\left[\rho \circ \iota_k(X_k) \,|\, \mathcal{F}_n^X \vee \sigma((Z_j)_{j=k}^m)\right] = 0. \qquad \square$$

The following result is immediate from Theorem 3.4 and Proposition 3.6.

**Theorem 3.7.** *The following containments hold*

$$\mathcal{H}_\mu^{\mathrm{norm}} \supseteq \mathcal{H}_\mu^{\mathrm{strong}} \supseteq \mathcal{H}_\mu^{\mathrm{det}}.$$

*Thus,* $\mathcal{H}_\mu^{\mathrm{det}} = \mathcal{G}$ *implies that* $\mathcal{P}_\mu^{\mathrm{strong}} \neq \emptyset$ *and* $\mathcal{H}_\mu^{\mathrm{norm}} \neq \mathcal{G}$ *implies that* $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$*.*

The following is a straightforward equivalent of Theorem 3.7 and we omit the proof.

**Corollary 3.8.** *If*

$$\lim_{m\to\infty}\lim_{n\to\infty}\left|\det\left(\prod_{\ell=m}^{n}\int_{G_\ell}\rho\circ\phi_k^\ell(z)\,\mu_\ell(dz)\right)\right| > 0$$

*for all irreducible representations* $\rho$ *of* $G_k$ *for all* $k \in \mathbb{N}$*, then* $\mathcal{P}_\mu^{\mathrm{strong}} \neq \emptyset$*. If*

$$\lim_{m\to\infty}\lim_{n\to\infty}\left\|\prod_{\ell=m}^{n}\int_{G_\ell}\rho\circ\phi_k^\ell(z)\,\mu_\ell(dz))\right\| = 0$$

*for some irreducible representation* $\rho$ *of* $G_k$ *for some* $k \in \mathbb{N}$*, then* $\mathcal{P}_\mu^{\mathrm{strong}} = \emptyset$*.*

Under a further assumption, we get a representation theoretic necessary and sufficient condition for the existence of strong solutions.

**Definition 3.9.** A Borel probability measure $\nu$ on a compact Hausdorff group $\Gamma$ is *conjugation invariant* if

$$\int_\Gamma f(g^{-1}xg)\,\nu(dx) = \int_\Gamma f(x)\,\nu(dx)$$

for all $g \in \Gamma$ and bounded Borel functions $f : \Gamma \to \mathbb{R}$.

*Remark 3.10.* Note that if $\Gamma$ is abelian, then any Borel probability measure $\nu$ on $\Gamma$ is conjugation invariant.

**Corollary 3.11.** *Suppose that each probability measure* $\mu_k$*,* $k \in \mathbb{N}$*, is conjugation invariant. Then,*

$$\mathcal{H}_\mu^{\mathrm{norm}} = \mathcal{H}_\mu^{\mathrm{strong}} = \mathcal{H}_\mu^{\mathrm{det}}$$

*and* $\mathcal{P}_\mu^{\mathrm{strong}} \neq \emptyset$ *if and only if each of these sets is* $\mathcal{G}$ *or, equivalently,*

$$\lim_{m\to\infty}\lim_{n\to\infty}\left|\prod_{\ell=m}^{n}\int_{G_\ell}\chi\circ\phi_k^\ell(z)\,\mu_\ell(dz)\right| > 0$$

*for each character* $\chi$ *of an irreducible representation of* $G_k$ *for all* $k \in \mathbb{N}$*.*

*Proof.* The result is immediate from Corollary 3.8 and Lemma 3.12 below. □

The following lemma is well known, but we include a proof for the sake of completeness.

**Lemma 3.12.** *If $\nu$ is a conjugation invariant Borel probability measure on a compact Hausdorff group $\Gamma$ and $\rho$ is an irreducible representation of $\Gamma$ with character $\chi$, then*

$$\int_\Gamma \rho(x)\,\nu(dx) = \int_\Gamma \chi(x)\,\nu(dx) \times I,$$

*where $I$ is the identity matrix.*

*Proof.* Let $\lambda$ be the normalized Haar measure on $\Gamma$. By assumption,

$$\int_\Gamma \rho(x)\,\nu(dx) = \int_\Gamma \int_\Gamma \rho(g^{-1}xg)\,\lambda(dg)\,\nu(dx).$$

Now, for $x, y \in \Gamma$ we have

$$\begin{aligned}\int_\Gamma \rho(g^{-1}xg)\,\lambda(dg)\,\rho(y) &= \int_\Gamma \rho(g^{-1}xgy)\,\lambda(dg)\\ &= \int_\Gamma \rho(yh^{-1}xh)\,\lambda(dh)\\ &= \rho(y)\int_\Gamma \rho(h^{-1}xh)\,\lambda(dh),\end{aligned}$$

and so the matrix $\int_\Gamma \rho(g^{-1}xg)\,\lambda(dg)$ commutes with the matrix $\rho(y)$ for all $y \in \Gamma$. It follows from Schur's lemma that $\int_\Gamma \rho(g^{-1}xg)\,\lambda(dg) = cI$ for some constant $c$, and taking traces of both sides gives $c = \chi(x)$. □

## 4 Freezing

Recall that the Hilbert–Schmidt norm of a matrix $A$ is given by $\|A\|_{HS} := \mathrm{tr}(A^*A)^{\frac{1}{2}}$, where $A^*$ is the adjoint of $A$ (this norm is also called the Frobenius norm and the Schur norm). Write $d(\rho)$ for the dimension of a unitary representation $\rho \in \mathcal{G}$, and note that $\|\rho(x)\|_{HS}^2 = \mathrm{tr}(I) = d(\rho)$. If $\nu$ is a probability measure on $G$, then $\|\int_G \rho(x)\,\nu(dx)\|_{HS}^2 \le d(\rho)$ by Jensen's inequality.

**Notation 4.1.** Set

$$\mathcal{H}_\mu^{\text{freeze}} := \left\{\rho \in \mathcal{G} : \sum_{m=k}^{\infty}\left[d(\rho) - \left\|\int_{G_k} \rho \circ \iota_k \circ \phi_k^m(z)\,\mu_m(dz)\right\|_{HS}^2\right] < \infty\ \forall k \in \mathbb{N}\right\}.$$

**Proposition 4.2.** *The sets* $\mathcal{H}_\mu^{\text{freeze}}$ *and* $\mathcal{H}_\mu^{\text{det}}$ *are equal, and so* $\mathcal{H}_\mu^{\text{freeze}} = \mathcal{H}_\mu^{\text{det}} = \mathcal{G}$ *implies that* $\mathcal{P}_\mu^{\text{strong}} \neq \emptyset$. *Moreover, if each probability measure* $\mu_k$, $k \in \mathbb{N}$, *is conjugation invariant, then*

$$\mathcal{H}_\mu^{\text{norm}} = \mathcal{H}_\mu^{\text{strong}} = \mathcal{H}_\mu^{\text{det}} = \mathcal{H}_\mu^{\text{freeze}}$$

*and* $\mathcal{P}_\mu^{\text{strong}} \neq \emptyset$ *if and only if each of these sets is* $\mathcal{G}$ *or, equivalently,*

$$\lim_{m\to\infty}\lim_{n\to\infty}\left|\prod_{\ell=m}^{n}\int_{G_\ell}\chi\circ\phi_k^\ell(z)\,\mu_\ell(dz)\right| > 0$$

*for each character* $\chi$ *of an irreducible representation of* $G_k$ *for all* $k \in \mathbb{N}$.

*Proof.* It suffices to show that $\mathcal{H}_\mu^{\text{freeze}} = \mathcal{H}_\mu^{\text{det}}$, because the remainder of the result will then follow from Theorem 3.7 and Corollary 3.11.

Fix $\rho \in \mathcal{G}$. Write $0 \le \lambda_k^\ell(1) \le \cdots \le \lambda_k^\ell(d(\rho))$ for the eigenvalues of the matrix

$$\left(\int_{G_k}\rho(z)\,\mu_k^\ell(dz)\right)^*\left(\int_{G_k}\rho(z)\,\mu_k^\ell(dz)\right).$$

Observe that

$$\lim_{m\to\infty}\lim_{n\to\infty}\prod_{\ell=m}^{n}\left|\det\int_{G_k}\rho(z)\,\mu_k^\ell(dz)\right| > 0$$

$$\Longleftrightarrow$$

$$\lim_{m\to\infty}\lim_{n\to\infty}\prod_{\ell=m}^{n}\left|\det\int_{G_k}\rho(z)\,\mu_k^\ell(dz)\right|^2 > 0$$

$$\Longleftrightarrow$$

$$\lim_{m\to\infty}\lim_{n\to\infty}\prod_{\ell=m}^{n}\lambda_k^\ell(1)\cdots\lambda_k^\ell(d(\rho)) > 0$$

$$\Longleftrightarrow$$

$$\sum_{m=k}^{\infty}\left[(1-\lambda_k^m(1))+\cdots+(1-\lambda_k^m(d(\rho)))\right] < \infty$$

$$\Longleftrightarrow$$

$$\sum_{m=k}^{\infty}\left[d(\rho)-(\lambda_k^m(1))+\cdots+\lambda_k^m(d(\rho)))\right] < \infty$$

$$\Longleftrightarrow$$

$$\sum_{m=k}^{\infty}\left[d(\rho)-\left\|\int_{G_k}\rho\circ\iota_k\circ\phi_k^m(z)\,\mu_m(dz)\right\|_{HS}^2\right] < \infty,$$

as required. □

Given Proposition 4.2, the reader may wonder why we introduced the set $\mathcal{H}_\mu^{\mathrm{freeze}}$. The equivalence established in Proposition 4.2 makes the proof of the following result considerably more transparent.

**Proposition 4.3.** *Suppose that each group $G_k$, $k \in \mathbb{N}$, is finite. Then, $\mathcal{H}_\mu^{\mathrm{det}} = \mathcal{H}_\mu^{\mathrm{freeze}} = \mathcal{G}$ if and only if for some (equivalently, all) $\mathbb{P} \in \mathcal{P}_\mu$ there are constants $c_{k,m} \in G_k$, $k,m \in \mathbb{N}$, $k \le m$, such that*

$$\mathbb{P}\{\phi_k^m(Z_m) \neq c_{k,m}\,\mathrm{i.o.}\} = 0$$

*for all $k \in \mathbb{N}$.*

*Proof.* Write $\mu_k^m$ for the probability measure on $G_k$ that is the push-forward of the probability measure $\mu_m$ on $G_m$ by the map $\phi_k^m : G_m \to G_k$. For simplicity, we write $\mu_k^m(g)$ instead of $\mu_k^m(\{g\})$ for $g \in G_k$. It is clear that $\mathbb{P}\{\phi_k^m(Z_m) \neq c_{k,m}\,\mathrm{i.o.}\} = 0$ $k \le m$ for all $k \in \mathbb{N}$ for some family of constants $c_{k,m} \in G_k$, $k,m \in \mathbb{N}$, if and only if $\mathbb{P}\{\phi_k^m(Z_m) \neq c_{k,m}^*\,\mathrm{i.o.}\} = 0$ where $c_{k,m}^*$ is any family with the property

$$\mu(c_{k,m}^*) = \max\{\mu_k^m(g) : g \in G_k\}$$

and, by the Borel–Cantelli lemma, this in turn occurs if and only if

$$\sum_{m=k}^{\infty} \mu(G_k \backslash \{c_{k,m}^*\}) < \infty$$

for all $k \in \mathbb{N}$.

Now,

$$\left(\sum_{g\in G_k} \mu_k^m(g)^2\right)^{1/2} \ge \max_{g\in G_k} \mu_k^m(g) = \mu_k^m(c_{k,m}) = \mu_k^m(c_{k,m}) \sum_{g\in G_k} \mu_k^m(g) \ge \sum_{g\in G_k} \mu_k^m(g)^2.$$

By Parseval's equality,

$$\sum_{g\in G_k} \mu_k^m(g)^2 = \frac{1}{\#G_k} \sum_{\rho\in\hat{G}_k} d(\rho) \left\| \sum_{g\in G_k} \rho(g)\mu_k^m(g) \right\|_{HS}^2,$$

and hence

$$\begin{aligned} 1 &- \left( \frac{1}{\#G_k} \sum_{\rho\in\hat{G}_k} d(\rho) \left\| \sum_{g\in G_k} \rho(g)\mu_k^m(g) \right\|_{HS}^2 \right) \\ &\ge \mu_k^m(G_k \backslash \{c_{k,m}\}) \\ &\ge 1 - \left( \frac{1}{\#G_k} \sum_{\rho\in\hat{G}_k} d(\rho) \left\| \sum_{g\in G_k} \rho(g)\mu_k^m(g) \right\|_{HS}^2 \right)^{1/2}. \end{aligned}$$

Note for a sequence of constant $(a_n)_{n\in\mathbb{N}} \subset [0,1]$ that $\sum_{n\in\mathbb{N}}(1-a_n) < \infty$ if and only if $\sum_{n\in\mathbb{N}}(1-a_n^2) < \infty$. Note also that

$$1 = \frac{1}{\#G_k} \sum_{\rho\in\hat{G}_k} d(\rho)^2.$$

Thus,

$$\sum_{m=k}^{\infty} \mu(G_k \backslash \{c_{k,m}^*\}) < \infty$$

for all $k \in \mathbb{N}$ if and only if

$$\sum_{m=k}^{\infty} \frac{1}{\#G_k} \sum_{\rho\in\hat{G}_k} d(\rho) \left[ d(\rho) - \left\| \sum_{g\in G_k} \rho(g)\mu_k^m(g) \right\|_{HS}^2 \right] < \infty$$

for all $k \in \mathbb{N}$, which is in turn equivalent to

$$\sum_{m=k}^{\infty} \sum_{\rho\in\hat{G}_k} \left[ d(\rho) - \left\| \sum_{g\in G_k} \rho(g)\mu_k^m(g) \right\|_{HS}^2 \right] < \infty$$

for all $\rho \in \hat{G}_k$ for all $k \in \mathbb{N}$.

A decomposition of the representation $\rho \circ \iota_k$ of $G_k$ for some $\rho \in \mathcal{G}$ into irreducibles shows that the last condition is equivalent to the one in the statement. □

*Remark 4.4.* It follows from Proposition 4.2 and Proposition 4.3 that if each group $G_k$, $k \in \mathbb{N}$, is finite and for some (equivalently, all) $\mathbb{P} \in \mathcal{P}_\mu$ there are constants $c_{k,m} \in G_k$, $k,m \in \mathbb{N}$, $k \le m$, such that

$$\mathbb{P}\{\phi_k^m(Z_m) \neq c_{k,m} \,\text{i.o.}\} = 0$$

for all $k \in \mathbb{N}$, then $\mathcal{P}_\mu^{\text{strong}} \neq \emptyset$. Moreover, these two conditions are equivalent when each probability measure $\mu_k$, $k \in \mathbb{N}$, is conjugation invariant. Also, for the special case when $G_k = \Gamma$, $k \in \mathbb{N}$, for some fixed finite group $\Gamma$ and each homomorphism $\phi_k : \Gamma \to \Gamma$ is the identity, it follows from Corollary 2.6 of [HY10] that the two conditions are equivalent. It would be interesting to know the status of the reverse implication in general.

## 5 Groups Indexed by the Lattice

Recall from the Introduction the example of our general setup where $G_k := G_{1,k} \times G_{2,k-1} \cdots \times G_{k,1}$ with each group $G_{i,j}$ a copy of some fixed compact abelian group $\Gamma$ and the homomorphism $\phi_k$ is given by

$$\phi_k(g_{1,k+1}, g_{2,k}, \ldots, g_{k+1,1}) := (g_{1,k+1} + g_{2,k}, g_{2,k} + g_{3,k-1}, \ldots, g_{k,2} + g_{k+1,1}).$$

We will consider the particular case where $\Gamma$ is $\mathbb{Z}_p$, the group of integers modulo some prime number $p$.

Because $\mathbb{Z}_p$ is abelian, all its irreducible representations of $G$ are one-dimensional. The irreducible representations are the trivial one and those of the form $\rho(g) = \prod_{n=1}^{m} \exp\left(\frac{2\pi i z_n}{p} g_{i_n, j_n}\right)$ for some $m$, pairs $(i_1, j_1), \ldots, (i_m, j_m) \in \mathbb{N}^2$, and $1 \le z_n \le p-1$.

The homomorphism $\phi_k^\ell$ maps $(g_{1,\ell}, \ldots, g_{\ell,1}) \in G_\ell$ to $(h_{1,k}, \ldots, h_{k,1}) \in G_k$ where

$$h_{i,k+1-i} = \sum_{j=0}^{\ell-k} \binom{\ell-k}{j} g_{i+j,\ell+1-i-j} \in \mathbb{Z}_p.$$

Set $f(m,n) := \binom{m}{n} \mod p$. When we restrict to $G_k$, the representation $\rho \circ \iota_k$ is of the form $\prod_{i=1}^{k} \exp\left(\frac{2\pi z_i}{p} g_{i,k+1-i}\right)$ with $0 \le z_i \le p-1$. We therefore need to evaluate

$$R_k^\ell = \int_{G_\ell} \prod_{i=1}^{k} \prod_{j=0}^{\ell-k} \exp\left(\frac{2\pi z_i}{p} f(\ell-k, j) g_{i+j,\ell+1-i-j}\right) \mu_\ell(dg_\ell)$$

to determine whether or not $\mathcal{P}_\mu^{\text{strong}} = \emptyset$. The following theorem of Lucas (see [Gra97]) gives the value of $f$.

**Theorem 5.1.** *Let $m, n$ be nonnegative integers and $p$ a prime number. Suppose*

$$m = m_k p^k + \ldots + m_1 p + m_0$$

*and*

$$n = n_k p^k + \ldots + n_1 p + n_0.$$

*Then,*

$$\binom{m}{n} = \prod_{i=0}^{k} \binom{m_i}{n_i} \mod p.$$

*Equivalently, if $m_0$ and $n_0$ are the least nonnegative residues of $m$ and $n$ mod $p$, then* $\binom{m}{n} = \binom{\lfloor m/p \rfloor}{\lfloor n/p \rfloor}\binom{m_0}{n_0}$.

Rather than use Theorem 5.1 directly to construct interesting examples, we consider a consequence of it for the case $p = 2$. Suppose that $\mu_k = \mu_{1,k} \otimes \cdots \otimes \mu_{k,1}$ where $\mu_{i,k+1-i}\{1\} = \pi_k = 1 - \mu_{i,k+1-i}\{0\}$ for some $0 \le \pi_k \le 1$.

Define $x = (x_{m,\ell+1-m})_{m=1}^{\ell} \in G_\ell = G_{1,\ell} \times \cdots \times G_{\ell,1} \cong \mathbb{Z}_2^\ell$ by

$$x := \sum_{i=1}^{k} \sum_{j=0}^{\ell-k} z_i f(\ell-k, j) e^{(i+j,\ell+1-i-j)},$$

where the arithmetic is performed modulo 2 and $e^{(m,\ell+1-m)} \in G_\ell$ is the vector with $e^{(m,\ell+1-m)}_{m,\ell+1-m} = 1$ and $e^{(m,\ell+1-m)}_{n,\ell+1-n} = 0$ for $n \neq m$. Then,

$$\int_{G_\ell} \prod_{i=1}^{k} \prod_{j=0}^{\ell-k} \exp\left(\frac{2\pi z_i}{p} f(\ell-k,j) g_{i+j,\ell+1-i-j}\right) \mu_\ell(dg_\ell) = (1-2\pi_\ell)^{M(k,\ell,z)},$$

where

$$M(k,\ell,z) := \#\{1 \le m \le \ell : x_{m,\ell+1-m} = 1\}.$$

Observe that if $x_{m,\ell+1-m} = 1$, then

$$\sum_{j=0}^{\ell-k} f(\ell-k,j) e^{(i+j,\ell+1-i-j)}_{m,\ell+1-m} = 1$$

for some $1 \le i \le k$ with $z_i = 1$. Now

$$\begin{aligned}
&\#\{1 \le m \le \ell : \sum_{j=0}^{\ell-k} f(\ell-k,j) e^{(i+j,\ell+1-i-j)}_{m,\ell+1-m} = 1\} \\
&= \#\{1 \le m \le \ell : f(\ell-k,m-i) = 1, i \le m \le i+\ell-k\} \\
&= \#\{i \le m \le i+\ell-k : f(\ell-k,m-i) = 1\} \\
&= \#\{0 \le m \le \ell-k : f(\ell-k,m) = 1\}.
\end{aligned}$$

As remarked in [Gra97], a consequence of the following theorem of Kummer from 1852 that the number of the binomial coefficients $\binom{m}{n}$, $0 \le n \le m$, which are odd is $2^{N(m)}$, where $N(m)$ is the number of times that the digit 1 appears in the base 2 representation of $m$.

**Theorem 5.2.** *Let $m,n$ be nonnegative integers and $p$ a prime number. The greatest power of $p$ that divides $\binom{m}{n}$ is given by the number of "carries" that are necessary when we add $m$ and $n-m$ in base $p$.*

Thus,

$$M(k,\ell,z) \le k 2^{N(\ell-k)}$$

and $M(k,\ell,z) = 2^{N(\ell-k)}$ when $\#\{1 \le i \le k : z_i = 1\} = 1$.

Therefore, if we assume $\pi_n \to 0$ as $n \to \infty$, then we are interested in whether

$$\lim_{\ell\to\infty} \prod_{r=1}^{\ell} (1-2\pi_{h+r})^{2^{N(r)}} \neq 0$$

for all $h \in \mathbb{N}$ or, equivalently, whether

$$\sum_{r=1}^{\infty} 2^{N(r)} \pi_{h+r} < \infty$$

for all $h \in \mathbb{N}$.

For example, fix a positive integer $a$ and an increasing function $b : \mathbb{N} \to \mathbb{N}$ such that $a \leq b(m) < m$ and $\lim_{m\to\infty} b(m) = \infty$. Suppose that $\pi_n = 0$ unless $2^m + 2^{b(m)} - 2^a \leq n \leq 2^m + 2^{b(m)}$ for some $m \in \mathbb{N}$. Note for any $h \in \mathbb{N}$ that

$$\sum_{r=1}^{\infty} 2^{N(r)} \pi_{h+r} = \sum_{s=k+1}^{\infty} 2^{N(s-h)} \pi_s$$

and this sum is finite if and only if

$$\sum_{n=1}^{\infty} 2^{b(\log_2 n)} \pi_n$$

is finite.

Thus, $\mathcal{P}_{\mu}^{\text{strong}} \neq \emptyset$ if and only if $\sum_{n=1}^{\infty} 2^{b(\log_2 n)} \pi_n < \infty$ in this case. On the other hand, $\mathbb{P}\{Z_k \neq 0 \text{ i.o.}\} > 0$ (equivalently, $\mathbb{P}\{Z_k \neq 0 \text{ i.o.}\} = 1$) if and only if $\sum_{n=1}^{\infty} n\pi_n < \infty$. Therefore, when $\lim_{m\to\infty} m - b(m) = \infty$ it is possible to construct $(\pi_n)_{n\in\mathbb{N}}$ such that almost surely infinitely many "bits" are "corrupted" and yet strong solutions still exist.

## 6 Automorphisms of the Torus

Consider the torus group $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. We write an element $x \in \mathbb{T}^2$ as a column vector $x = (x_1, x_2)^\top \in [0,1)^2$, where $\top$ denotes the transpose of a vector.

Any $2 \times 2$ $\mathbb{Z}$-valued matrix $S$ defines a homomorphism $x \mapsto Sx$ from $\mathbb{T}^2$ to itself if we do ordinary matrix multiplication modulo $\mathbb{Z}^2$. If the matrix $S$ has determinant 1, then this homomorphism is invertible. Such a transformation is called a *linear toral automorphism*.

Note that if

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then the eigenvalues of $S$ are

$$\frac{1}{2}(a + d \pm \sqrt{a^2 + 4bc - 2ad + d^2}) = \frac{1}{2}(a + d \pm \sqrt{(a+d)^2 - 4}),$$

Thus, the eigenvalues are real and distinct unless $a+d$ is 0, $\pm 1$ or $\pm 2$, in which case the pairs of eigenvalues are, respectively $\{\pm i\}$, $\{\frac{1}{2}(1 \pm i\sqrt{3})\}$, $\{\frac{1}{2}(-1 \pm i\sqrt{3})\}$, $\{1,1\}$, and $\{-1,-1\}$. Note that in each of the latter cases the eigenvalues lie on the unit circle.

**Definition 6.1.** A *ergodic toral automorphism* is a linear toral automorphism given by a matrix $S$ with no eigenvalues on the unit circle.

For some of the more probabilistic properties of ergodic toral automorphisms, see [Kat71]. Such mappings are the prototypical examples of Anosov systems that have been the subject of intensive study dynamical systems world (see [Fra69]).

A hyperbolic linear toral automorphism has two real eigenvalues $\lambda_1 > 1 > \lambda_1^{-1} = \lambda_2$. These eigenvalues are irrational and the corresponding (right) eigenvectors $v^1$ and $v^2$ have irrational slope (see, e.g., Sect. 5.6 of [LT93]).

**Theorem 6.2.** *Suppose for every $i \in \mathbb{N}$ that the group $G_i$ is a copy of $\mathbb{T}^2$ and that the homomorphism $\phi_i$ is a fixed ergodic toral automorphism given by a matrix $S$. Suppose the noise distribution $\mu_k$ is a fixed measure $\mu^*$ that satisfies $\mu^*(A) \geq \varepsilon\lambda(A \cap B)$ for every Borel set $A$, where $\varepsilon > 0$, $\lambda$ is normalized Haar measure, and $B$ is a fixed Borel set $B$ with $\lambda(B) > 0$. Then, $\mathcal{P}_\mu^{strong} = \emptyset$.*

*Proof.* We need to evaluate $R_k^\ell = \int_{\mathbb{T}^2} \rho \cdot \iota_k \cdot \phi_k^\ell(z)\mu_\ell(dz)$. Let $\nu$ be the measure defined by $\nu(A) = \varepsilon\lambda(A \cap B)$ a Borel set $A$, where $\varepsilon$, $\lambda$ and $B$ are as in the statement. Observe that

$$
\begin{aligned}
|R_k^\ell| &\leq \int_{\mathbb{T}^2 G_\ell} |\rho \cdot \iota_k \cdot \phi_k^\ell(z)|\,(\mu_\ell - \nu)(dz) + \int_{\mathbb{T}^2} |\rho \cdot \iota_k \cdot \phi_k^\ell(z)|\,\nu(dz)| \\
&\leq \int_{\mathbb{T}^2} (\mu_\ell - \nu)(dz) + \left| \int_{\mathbb{T}^2} \rho \cdot \iota_k \cdot \phi_k^\ell(z)\,\nu(dz) \right|,
\end{aligned}
$$

and note that the last term on the right-hand side is $\left|\int_{\mathbb{T}^2} \rho \cdot \iota_k(z)\,(\nu \cdot \phi_k^\ell)^{-1})(dz)\right|$.

As noted in Sect. 5.6 of [LT93], any ergodic toral automorphism $S$ exhibits *topological mixing*: for any Borel sets $A, B \subseteq \mathbb{R}^2$, $\lim_{n\to\infty} \frac{\lambda(S^n B) \cap A}{\lambda(B)} = \lambda(A)$. Because $\phi_k^\ell$ is an ergodic toral automorphism, so is $(\phi_k^\ell)^{-1}$. Therefore, $\lim_{\ell\to\infty} \left|\int_{\mathbb{T}^2} \rho \cdot \iota_k(z)(\nu \cdot \phi_k^\ell)^{-1}(dz)\right| = |\int_{\mathbb{T}^2} \rho \cdot \iota_k(z)\varepsilon\lambda(dz)| = 0$. Consequently, $|R_k^\ell| \leq \int_{\mathbb{T}^2}(\mu_\ell - \nu)(dz) = 1 - \varepsilon\lambda(B)$ for every nontrivial representation $\rho$, and hence

$$
\lim_{m\to\infty} \lim_{n\to\infty} |R_k^n R_k^{n-1} \cdots R_k^m| = 0 \; \forall k \in \mathbb{N},
$$

showing that $\mathcal{P}_\mu^{strong} = \emptyset$. ☐

Every finite-dimensional unitary representation of $G_i$ is of the form

$$
x \mapsto e^{2\pi i(z \cdot x)},
$$

where $z$ is a vector $(z_1, z_2) \in \mathbb{Z}^2$ and $z \cdot x$ is the usual inner product. Hence, if we lift this representation to a representation of $G$ we have

$$
R_k^\ell = \int_{\mathbb{T}^2} e^{2\pi i(z \cdot S^{\ell-k} x)}\, \mu_\ell(dx).
$$

Suppose that the probability measure $\mu_\ell$ is concentrated on the set of multiples of the eigenvector $v^2$ associated with the eigenvalue $\lambda_2 \in (0,1)$. Then,

$$R_k^\ell = \int_{\mathbb{R}} e^{2\pi i (t\lambda_2^{\ell-k} z \cdot v^2)} \, \nu_\ell(dt)$$

for some probability measure $\nu_\ell$ on $\mathbb{R}$. It is clear that under appropriate hypotheses

$$\lim_{m\to\infty} \lim_{n\to\infty} |R_k^n R_k^{n-1} \cdots R_k^m| > 0 \; \forall k \in \mathbb{N}$$

and hence, by Corollary 3.8, $\mathcal{P}_\mu^{\text{strong}} \neq \emptyset$. For example, if $\nu_\ell = \nu$ for all $\ell \in \mathbb{N}$ for some fixed probability measure $\nu$ on $\mathbb{R}$, then it suffices that $\int_{\mathbb{R}} |t| \, \nu(dt) < \infty$. In particular, it is possible to construct examples where $\mu_1 = \mu_2 = \ldots$ is a measure that has all of $\mathbb{T}^2$ as its closed support and yet $\mathcal{P}_\mu^{\text{strong}} \neq \emptyset$.

## References


AUY08. Jirô Akahori, Chihiro Uenishi, and Kouji Yano, *Stochastic equations on compact groups in discrete negative time*, Probab. Theory Related Fields **140** (2008), no. 3–4, 569–593. MR 2365485 (2009d:60173)

Cir75. B. S. Cirel′son, *An example of a stochastic differential equation that has no strong solution*, Teor. Verojatnost. i Primenen. **20** (1975), no. 2, 427–430. MR 0375461 (51 #11654)

Fra69. John Franks, *Anosov diffeomorphisms on tori*, Trans. Amer. Math. Soc. **145** (1969), 117–124. MR 0253352 (40 #6567)

Gra97. Andrew Granville, *Arithmetic properties of binomial coefficients. I. Binomial coefficients modulo prime powers*, Organic mathematics (Burnaby, BC, 1995), CMS Conf. Proc., vol. 20, Amer. Math. Soc., Providence, RI, 1997, pp. 253–276. MR 1483922 (99h:11016)

HY10. Takao Hirayama and Kouji Yano, *Extremal solutions for stochastic equations indexed by negative integers and taking values in compact groups*, Stochastic Process. Appl. **120** (2010), no. 8, 1404–1423. MR 2653259 (2011j:60222)

Kat71. Yitzhak Katznelson, *Ergodic automorphisms of $T^n$ are Bernoulli shifts*, Israel J. Math. **10** (1971), 186–195. MR 0294602 (45 #3672)

LT93. Ding Jun Luo and Li Bang Teng, *Qualitative theory of dynamical systems*, Advanced Series in Dynamical Systems, vol. 12, World Scientific Publishing Co. Inc., River Edge, NJ, 1993. MR 1249274 (94k:58043)

Raj11. Chandiraraj Robinson Edward Raja, *A stochastic difference equation with stationary noise on groups*, 2011, *Canad. J. Math.*, to appear. http://dx.doi.org/10.4153/CJM-2011-094-6.

Tak09. Yoichiro Takahashi, *Time evolution with and without remote past*, Advances in discrete dynamical systems, Adv. Stud. Pure Math., vol. 53, Math. Soc. Japan, Tokyo, 2009, pp. 347–361. MR 2582432 (2011b:60019)

Yor92. Marc Yor, *Tsirel′ son's equation in discrete time*, Probab. Theory Related Fields **91** (1992), no. 2, 135–152. MR 1147613 (93d:60104)

# Modeling Competition Between Two Influenza Strains

**Rinaldo B. Schinazi**

**Abstract** We use spatial and nonspatial models to argue that competition alone may explain why two influenza strains do not usually coexist during a given flu season. The more virulent strain is likely to crowd out the less virulent one. This can be seen as a consequence of the Exclusion Principle of Ecology. We exhibit, however, a spatial model for which coexistence is possible.

**Keywords** Competition models • Stochastic process • Influenza • Swine strain • Exclusion principle • Ecology

## 1 Introduction

The seasonal flu strain was a lot less prevalent during the 2009/2010 influenza season than during the previous years, see Fluview (the weekly CDC influenza report). On the other hand, some time during spring 2009, the new so-called swine strain appeared. There seems to be a relation between these two events. In this chapter we propose to explain this phenomenon using competition models. We will use spatial and nonspatial models to show that in a given flu season coexistence of two strains is unlikely due to competition alone. We will also show that geometry and space may be critical for coexistence. Our models deal with competition over only one flu season. In the real world, because of mutations the fight between two strains may go on for several flu seasons before one strain outcompetes the other. This picture is consistent with the very skinny shape of the phylogenetic tree for influenza; see, for instance Koelle et al. [5] and van Nimwegen [8]. In this chapter, the two competing strains are assumed not to undergo mutations, and therefore the time scale we focus on is one flu season.

R.B. Schinazi (✉)
University of Colorado at Colorado Springs, Colorado Springs CO80933-7150, USA
e-mail: rschinaz@uccs.edu

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_3, 

A competing explanation of the non-coexistence of the two influenza strains is cross immunity. For instance, immunity may explain why older generations have not been as much affected as the younger ones in the swine flu epidemic. It may be due to some previous exposure to a similar strain, see the discussion in Greenbaum et al. [3]. However, using a cross-immunity argument to explain why the swine strain crowds out the seasonal one may be more difficult. The hypothesis would be that the swine strain must confer some immunity against the seasonal flu. But, clearly the seasonal strain does not confer any immunity against the swine strain: after all even young people (the group most severely affected by the swine strain) have usually been exposed to the seasonal strain and do not seem to be protected against the swine strain. Hence, for this argument to work the swine strain must confer some immunity against the seasonal strain, but the seasonal strain cannot confer any immunity against the swine strain. In contrast to this cross immunity hypothesis we argue in this chapter that even in models for which there is no immunity at all (every individual that recovers is immediately susceptible again!), coexistence of two competing strains is rather unlikely.

## 2 The ODE Model

Our first model is a system of ordinary differential equations. Let $u_1(t)$ and $u_2(t)$ be the density of individuals infected at time $t$ with strains 1 and 2, respectively . We set

$$\begin{aligned} u_1' &= \lambda_1 u_1 u_0 - \delta_1 u_1 \\ u_2' &= \lambda_2 u_2 u_0 - \delta_2 u_2 \end{aligned}$$

where $u_0(t)$ is the density of susceptible individuals at time $t$. In words, individuals infected with strain $i$ infect susceptible individuals at rate $\lambda_i$ and get healthy at rate $\delta_i$, for $i = 1,2$. We assume that the only possible states are 0, 1 and 2. Hence, at any time $t > 0$ we have $u_0(t) + u_1(t) + u_2(t) = 1$. In particular, as soon as an infected individual gets healthy, it is back in the susceptible pool.

Let 1 be the seasonal and 2 be the swine strains. Some reports indicate that the swine strain may be more virulent than the seasonal strain, see Fraser et al. [2]. Under that assumption,

$$\frac{\lambda_1}{\delta_1} < \frac{\lambda_2}{\delta_2}.$$

Assume also that at some point in time the ODE model is at the equilibrium $(0, 1 - \frac{\delta_2}{\lambda_2})$. That is, there is no seasonal strain and the swine strain is in equilibrium. Now introduce a little bit of seasonal strain (small $u_1$). Will the seasonal strain be able to grow? Using that $u_1$ is almost 0 and that $u_2$ is almost $1 - \frac{\delta_2}{\lambda_2}$ we make the approximation

$$u_0 = 1 - u_1 - u_2 \sim 1 - \left(1 - \frac{\delta_2}{\lambda_2}\right) = \frac{\delta_2}{\lambda_2}.$$

Hence,

$$u_1' \sim \lambda_1 u_1 \frac{\delta_2}{\lambda_2} - \delta_1 u_1 = u_1 \left( \lambda_1 \frac{\delta_2}{\lambda_2} - \delta_1 \right).$$

Since we are assuming that $\frac{\lambda_1}{\delta_1} < \frac{\lambda_2}{\delta_2}$ we get $u_1' < 0$. That is, under these assumptions and according to this model, the seasonal flu will not take hold.

In fact this system of ODE is a particular case of a well-known competition model. For the general version of this model, it is known that one of the strains will vanish; see Exercise 3.3.5 in Hofbauer and Sigmund [4]. The point is that we have two populations (the population of individuals infected with strain 1 and the population of individuals infected with strain 2) that compete for a single resource (the susceptible individuals). It turns out that in such a model, one population will drive the other one out. This is a particular case of the so-called "Exclusion Principle" of Ecology: if the number of populations is larger than the number of resources all the populations cannot subsist in the long run, see 5.4 in Hofbauer and Sigmund [4].

## 3 The Spatial Stochastic Model

In the preceding model there is no space structure, and all the individuals in the population can be seen as neighbors of each other. In this section, we go to the other extreme where there is a rigid space structure and each individual has a fixed number of neighbors.

We now describe a spatial stochastic model called the multitype contact process, see Neuhauser [7]. Let $S$ be the integer lattice $\mathbf{Z}^d$ ($d$ is the dimension) or the homogeneous tree $\mathbf{T}_d$ for which each site has $d+1$ neighbors. The system is described by a configuration $\xi \in \{0,1,2\}^S$, where $\xi(x) = 0$ means that site $x$ is occupied by a susceptible individual, $\xi(x) = 1$ means that $x$ is occupied by an individual infected by strain 1 and $\xi(x) = 2$ means that $x$ is occupied by an individual infected by strain 2. If $S$ is $\mathbf{Z}^d$, then each site has $2d$ neighbors, if $S$ is $\mathbf{T}_d$, then each site has $d+1$ neighbors. For $x \in S$ and $\xi \in \{0,1,2\}^S$, let $n_1(x,\xi)$ and $n_2(x,\xi)$ denote the number of neighbors of $x$ that are infected by strain 1 and strain 2, respectively.

The multitype contact process $\xi_t$ with birth rates $\lambda_1, \lambda_2$ makes transitions at $x$ when the configuration of the process is $\xi$

$$1 \to 0 \text{ at rate } 1$$

$$2 \to 0 \text{ at rate } 1$$

$$0 \to 1 \text{ at rate } \lambda_1 n_1(x,\xi),$$

$$0 \to 2 \text{ at rate } \lambda_2 n_2(x,\xi),$$

In words, a susceptible individual gets infected by an infected neighbor at rates $\lambda_1$ or $\lambda_2$, depending on which strain the neighbor is infected with. An infected individual gets healthy (and is immediately susceptible again) at rate 1. Note that compared to the ODE model, we are assuming in this model that $\delta_1 = \delta_2 = 1$. This is so because most of the mathematical results have been proved under the assumption $\delta_1 = \delta_2$. We take this common value to be 1 to minimize the number of parameters.

The multitype contact process is a generalization of the basic contact process which has only one type. The transitions of the basic contact process are given by

$$1 \to 0 \text{ at rate } 1$$

$$0 \to 1 \text{ at rate } \lambda_1 n_1(x, \xi),$$

For the basic contact process, there exists a critical value $\lambda_c$ whose exact value is not known and which depends on the graph the model is on. If $\lambda_1 > \lambda_c$, then starting with even a single infected individual, there is a positive probability of having infected individuals at all times somewhere in the graph. On the other hand, if $\lambda_1 \leq \lambda_c$, then starting from any finite number of infected individuals all the infected individuals will disappear after a finite time. See Liggett [6] for more on the basic contact process on the square lattice and on trees.

## 3.1 The Space Is the Square Lattice $\mathbf{Z}^d$

We now go back to the multitype contact process. Assume that $\lambda_2 > \lambda_c$ and $\lambda_2 > \lambda_1$ then there is no coexistence of strains 1 and 2 in the sense that

$$\lim_{t \to \infty} P(\xi_t(x) = 1, \xi_t(y) = 2) = 0$$

for any sites $x$ and $y$ in $\mathbf{Z}^d$ and any initial configuration $\xi_0$. In fact, strain 2 always drives out strain 1 in the following sense. Let $A$ be the event that strain 2 will not ever disappear. Then,

$$\lim_{t \to \infty} P(\xi_t(x) = 1 | A) = 0,$$

for any site $x$ in $\mathbf{Z}^d$ and any initial configuration (see Theorem 2 in Cox and Schinazi [1] and also Neuhauser [7]). Hence, assuming that $\lambda_2 > \lambda_1$ (i.e., strain 2 is more virulent than strain 1) this model too predicts that the seasonal flu will be crowded out by the swine strain. The spatial structure seems to have no influence on the outcome. The next section will show that this is not always so and that a different (more crowded) space structure allows coexistence.

## 3.2 The Space Is the Tree $\mathbf{T}_d$

There is a fundamental difference between the basic contact process on the square lattice and the same model on the tree. There are two (instead of one) critical values for the basic contact process on the tree. The definition of $\lambda_c$ is as before. We also define another critical value $\lambda_{cc}$ in the following way. Consider the basic (one type) contact process with birthrate $\lambda_1$. Let $O$ be a fixed site on the tree or square lattice. Start the process with a single infected individual at $O$. The probability that the infection will return to site $O$ infinitely many times is positive if and only if $\lambda_1 > \lambda_{cc}$. It turns out that $\lambda_c < \lambda_{cc}$ on the tree but $\lambda_c = \lambda_{cc}$ on the square lattice.

The fact that the basic contact process has two distinct critical values on the tree has interesting consequences for the multitype contact process on the tree. Let $\lambda_1$ and $\lambda_2$ be in $(\lambda_c, \lambda_{cc})$, and then strains 1 and 2 may coexist on the tree in the following sense. Under suitable initial configurations we have for any sites $x$ and $y$

$$\liminf_{t\to\infty} P(\xi_t(x) = 1, \xi_t(y) = 2) > 0.$$

See Theorem 1 in Cox and Schinazi [1]. Note that coexistence occurs even for $\lambda_1 < \lambda_2$ but both parameters need to be in the rather narrow interval $(\lambda_c, \lambda_{cc})$. This result shows that space structure and geometry may be crucial in allowing coexistence.

## 4 Discussion

Our models show that at least in theory coexistence of two competing strains is unlikely. Coexistence is however possible for the multitype contact process on a tree. The tree can be thought of as a model for high-density populations (in a ball of radius $r$ there are $(d+1)d^{r-1}$ individuals on the tree $\mathbf{T_d}$ but only about $r^d$ on the lattice $\mathbf{Z^d}$). In order to have coexistence, both infection rates cannot be too low or too high but may be unequal. In all other cases, there will be no coexistence on the tree, and there is never coexistence on $\mathbf{Z^d}$ unless $\lambda_1$ is exactly equal to $\lambda_2$, a rather unlikely possibility, see Neuhauser [7]. Interestingly, the behavior of the mean-field ODE model is the same as the behavior of the model on $\mathbf{Z^d}$ but different from the model on the tree. In general, it is expected that the model on the tree to be closer to the mean-field model than to the model on $\mathbf{Z^d}$. This is not so in this example.

## References

1. J.T. Cox and R.B.Schinazi (2009). Survival and coexistence for a multitype contact process. Annals of probability **37**, 853–876
2. C. Fraser et al. (2009).Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. Science, **324**, 1557–1561
3. J.A. Greenbaum et al. (2009). Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. PNAS, **106**, 20365–20370.
4. J. Hofbauer and K. Sigmund (1998). *Evolutionary games and population dynamics.* Cambridge University Press.
5. K.Koelle, S. Cobey, B. Grenfell and M. Pascual (2006) Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in Humans. Science vol. 314, 1898–1903.

6. T.Liggett (1999). *Stochastic interacting systems: contact, voter and exclusion processes,* Springer-Verlag, Berlin.
7. C. Neuhauser (1992). Ergodic theorems for the multitype contact process. Probab. Theory Related Fields, **91**, 467–506.
8. E. van Nimwegen (2006). Influenza escapes immunity along neutral networks. Science vol. 314, 1884–1886.

# Asymptotic Results for Near Critical Bienaymé–Galton–Watson and Catalyst-Reactant Branching Processes

**Amarjit Budhiraja and Dominik Reinhold**

**Abstract** Near critical single-type Bienaymé–Galton–Watson (BGW) processes are considered. Results on convergence of Yaglom distributions of suitably scaled BGW processes to that of the corresponding diffusion approximation are given. Convergences of stationary distributions for Q-processes and models with immigration to the corresponding distributions of the associated diffusion approximations are established. Similar results can be obtained in a multitype setting. To illustrate this, a result on convergence of Yaglom distributions of suitably scaled multitype subcritical BGW processes to that of the associated diffusion model is presented.

In the second part, near critical catalyst-reactant branching processes with controlled immigration are considered. The reactant population evolves according to a branching process whose branching rate is proportional to the total mass of the catalyst. The bulk catalyst evolution is that of a classical continuous-time branching process; in addition, there is a specific form of immigration. Immigration takes place exactly when the catalyst population falls below a certain threshold, in which case the population is instantaneously replenished to the threshold. A diffusion limit theorem for the scaled processes is presented, in which the catalyst limit

This research is partially supported by the National Science Foundation (DMS-1004418, DMS-1016441), the Army Research Office (W911NF-10-1-0158), NSF Emerging Frontiers in Research and Innovation (EFRI) (Grant CBE0736007), and the US-Israel Binational Science Foundation (Grant 2008466).

A. Budhiraja
Department of Statistics and Operations Research, University of North Carolina,
Chapel Hill, NC 27599, USA
e-mail: budhiraj@email.unc.edu

D. Reinhold (✉)
Department of Mathematics and Computer Science,
Clark University, Worcester, MA 01610, USA
e-mail: dreinhold@clarku.edu

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*,
Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_4, 

is described through a reflected diffusion, while the reactant limit is a diffusion with coefficients that are functions of both the reactant and the catalyst. Stochastic averaging under fast catalyst dynamics is considered next. In the case where the catalyst evolves "much faster" than the reactant, a scaling limit, in which the reactant is described through a one-dimensional SDE with coefficients depending on the invariant distribution of the reflected diffusion, is obtained.

**Keywords** Branching processes • Catalyst-reactant branching processes • Quasi stationary distributions • Yaglom distributions • Q-processes • Near critical regime • Chemical reaction networks • Diffusion approximations • Stochastic averaging • Multiscale approximations • Reflected diffusions • Constrained martingale problems • Echeverria criterion • Invariant measure convergence

**AMS subject classifications (2000):** Primary 60J80; secondary 60F05.

## 1 Introduction

This chapter reviews some recent asymptotic results on near critical Bienaymé–Galton–Watson (BGW) branching processes and on catalyst-reactant branching processes with controlled immigration. Proofs of the former results can be found in [3], whereas those of the latter results are in [2, 18].

It is well known that under suitable assumptions, appropriately scaled near critical BGW processes converge in the large population limit to Feller diffusions (see [5, 8]). We are concerned with relationships between the steady-state behavior of the branching processes and that of their approximating diffusions. One, of course, needs to suitably interpret the term "steady state" since, as is well known, with time each branching process will grow to infinity on the set of non-extinction in the supercritical case and become extinct in the critical and subcritical case (see [1]). A natural approach in the subcritical case is to study probability laws conditioned on the event of non-extinction. In the supercritical case, a common approach is to additionally condition on the event of eventual extinction. Such a conditioning leads to the notion of quasi-stationary distributions. Results establishing convergence of quasi-stationary distributions of the scaled BGW processes to that of the limiting Feller diffusions will be presented. Similar results for the closely related setting of BGW models with immigration will also be given. Analogous properties can be established in a multitype setting, and we will illustrate this through a result for multitype subcritical BGW processes.

Next, we consider near critical catalyst-reactant branching processes with a specific form of controlled immigration. The catalyst population evolves according to a classical continuous-time branching process, while the reactant population evolves according to a branching process whose branching rate is proportional to the total mass of the catalyst. Immigration takes place exactly when the catalyst population falls below a certain threshold, in which case the population is instantaneously replenished to the threshold to ensure a certain level of activity. Our main goal here is to establish diffusion approximations for the catalyst and reactant populations in two settings. In the first setting, both populations evolve on "comparable timescales," while in the second setting, the catalyst evolves "much faster" than the reactant, in a sense made precise in Sect. 3. In the first setting, the limit model is described through a coupled system of stochastic differential equations with reflection in the space $[1,\infty)\times\mathbb{R}$. In the second setting, we establish a stochastic averaging result that says that the limit reactant evolution is given through an autonomous one-dimensional SDE with coefficients described in terms of the invariant distribution of a reflected diffusion in $[1,\infty)$; the reflected diffusion can be interpreted as the limiting dynamics of the catalyst process under a suitable scaling.

This chapter is organized as follows: in Sect. 2 we review results on BGW processes that have appeared in [3, 18], and in Sect. 3 we present results on near critical catalyst-reactant branching processes with controlled immigration, proofs of which can be found in [2, 18].

## 2 Asymptotic Results on Near Critical Branching Processes

Consider a population consisting of $k$ types of particles whose evolution is described in terms of a discrete time multitype ($k$ type) Bienaymé–Galton–Watson ($k$-BGW) process—such a process is a Markov chain $\{\mathbf{Z}_p\}_{p\in\mathbb{N}_0}$ on $\mathbb{N}_0^k$, with the vector $\mathbf{Z}_p$ representing the number of particles of each type in generation $p$. We are interested in the longtime behavior of the scaled process $\frac{1}{p}\mathbf{Z}_{\lfloor pt\rfloor}$, $t\geq 0$, when the $k$-BGW process is close to criticality. More precisely, we consider a sequence of BGW processes $\{\mathbf{Z}_p^{(n)}, p\in\mathbb{N}_0\}_{n\in\mathbb{N}}$ such that, as $n$ becomes large, the processes approach criticality (in the sense of Condition 2.1). It is well known (see [5, 8]) that, under suitable conditions, the process $\hat{\mathbf{Z}}_t^{(n)} = \frac{1}{n}\mathbf{Z}_{\lfloor nt\rfloor}^{(n)}$, $t\geq 0$, converges weakly to a diffusion $\boldsymbol{\xi}$. Such a result implies convergence of finite time statistics of $\hat{\mathbf{Z}}^{(n)}$ to those of $\boldsymbol{\xi}$, but does not provide any information on relationships between the time asymptotic behaviors of $\hat{\mathbf{Z}}^{(n)}$ and $\boldsymbol{\xi}$. The main goal of this section is to make such relationships mathematically precise. In particular, we show that, under appropriate assumptions, the time asymptotic distribution of $\hat{\mathbf{Z}}_t^{(n)}$ with suitable conditioning converges to that of $\boldsymbol{\xi}_t$ with a similar conditioning, as $n\to\infty$ (see Theorems 2.3 and 2.6). An analogous result for models with immigration (where no conditioning is required) is also established (Theorem 2.8). The results say that the longtime behavior of a BGW process (under suitable conditioning or with immigration) is well approximated by that of the corresponding diffusion limit $\boldsymbol{\xi}$. Most of the results in this section are for single-type BGW processes, namely for the case $k=1$. Similar results can be obtained in multitype settings and we consider one such result in Theorem 2.12.

When $k = 1$, the transition probabilities of a BGW process $\{Z_p\}$ can be written as

$$P(Z_{p+1} = j | Z_p = i) = \begin{cases} \mu^{*i}(j) & \text{if } i \geq 1, \quad j \geq 0, \\ \delta_{0j} & \text{if } i = 0, \quad j \geq 0, \end{cases} \tag{1}$$

where $\{\mu(l)\}_{l \in \mathbb{N}_0}$ is the offspring distribution of a typical particle and $\{\mu^{*i}(l)\}_{l \in \mathbb{N}_0}$ is the $i$-fold convolution of $\{\mu(l)\}_{l \in \mathbb{N}_0}$. The process starts with $Z_0$ particles; each of the $Z_p$ particles alive at time $p$ lives for one unit of time and then dies, giving rise to $l$ offspring particles with probability $\mu(l)$, $l \in \mathbb{N}_0$. The particles behave independently of each other and of the past.

Depending on the mean $m$ of the offspring distribution, BGW processes can be divided into three cases: subcritical, critical, and supercritical, according to whether $m < 1$, $m = 1$, or $m > 1$, respectively.

In order to describe near critical BGW processes, we will consider a sequence of processes $Z^{(n)}$ with offspring distributions $\mu^{(n)}$. If $Z_0^{(n)} = 1$, then $Z_1^{(n)}$ has the probability-generating function (pgf)

$$F^{(n)}(s) = \sum_{l=0}^{\infty} \mu^{(n)}(l) s^l, \quad s \in [0,1]. \tag{2}$$

We denote the mean of the offspring distribution by $m^{(n)}$ and the variance by $\kappa^{(n)}$. Denote the $p^{th}$ iterate of $F^{(n)}$ by $F_p^{(n)}$, i.e., for $s \in [0,1]$ and $p \geq 0$

$$F_0^{(n)}(s) = s, \quad F_{p+1}^{(n)}(s) = F^{(n)}(F_p^{(n)}(s)).$$

Let $q^{(n)}$ be the extinction probability of $Z^{(n)}$ starting with a single particle, i.e.

$$q^{(n)} = P(Z_p^{(n)} = 0 \text{ for some } p \in \mathbb{N} | Z_0^{(n)} = 1).$$

Denote by $\mathcal{P}(\mathbb{R}_+)$ the set of probability measures on $\mathbb{R}_+ := [0, \infty)$ with the Borel $\sigma$-field.

**Condition 2.1.** (i) For each $n$, $\mu^{(n)}(0) > 0$, $\mu^{(n)}(0) + \mu^{(n)}(1) < q^{(n)}$.

(ii) For each $n$, $m^{(n)} = 1 + \frac{c^{(n)}}{n}, c^{(n)} \in (-n, \infty) \setminus \{0\}$, and as $n \to \infty$, $c^{(n)} \to c \in \mathbb{R} \setminus \{0\}$.

(iii) For each $n$, $\kappa^{(n)} < \infty$, and as $n \to \infty$, $\kappa^{(n)} \to \kappa \in (0, \infty)$.

(iv) As $n \to \infty$, the distribution of $\frac{Z_0^{(n)}}{n}$ converges to some $\mu_0 \in \mathcal{P}(\mathbb{R}_+)$.

(v) The family of functions $\{F^{(n)''}\}_{n \in \mathbb{N}}$ is equicontinuous at 1.

Condition 2.1 (ii) ensures that, as $n \to \infty$, $m^{(n)} \to 1$, and eventually, the processes approach criticality strictly from above or strictly from below. The case where $c < 0$ will be referred to as the (near critical) subcritical case while $c > 0$ corresponds to the supercritical case. Conditions 2.1 (ii)–(v) are used for the diffusion approximation result in Theorem 2.1. Condition 2.1 (v) will also be used in the study of the supercritical case in Theorem 2.3. Let

$$\hat{Z}_t^{(n)} := \frac{1}{n} Z_{\lfloor nt \rfloor}^{(n)}, \quad t \in \mathbb{R}_+; \tag{3}$$

then $\{\hat{Z}_t^{(n)}\}_{t\in\mathbb{R}_+}$ is an $\mathbb{S}^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\}$-valued (time inhomogeneous) Markov process with sample paths in $D(\mathbb{R}_+ : \mathbb{S}^{(n)})$, the space of RCLL (right continuous, left limit) functions from $\mathbb{R}_+$ to $\mathbb{S}^{(n)}$. Throughout, $\mathbb{S}^{(n)}$ is endowed with the discrete topology and given a metric space $S$, $D(\mathbb{R}_+ : S)$ is endowed with the usual Skorohod topology. The space of probability measures on a metric space $S$ will be denoted by $\mathcal{P}(S)$.

We now recall a well-known weak convergence result for $\hat{Z}^{(n)}$ (see [6], [14, Theorem 2.1]), which describes the asymptotic behavior of $\hat{Z}^{(n)}$, as $n \to \infty$, over any fixed finite time horizon. A related multidimensional result will be presented later in this section (see also [8]). Denote by $C^l(\mathbb{R}_+)$ the set of $l$-times differentiable, real-valued functions on $\mathbb{R}_+$. In the following theorem, we do not need part (i) of Condition 2.1, and in part (ii) of the assumption $(-n,\infty) \setminus \{0\}$ and $\mathbb{R} \setminus \{0\}$ can be replaced by $(-n,\infty)$ and $\mathbb{R}$, respectively (see [14, Theorem 2.1]). However, in order to simplify our presentation, we assume Condition 2.1 to hold, even though for some of the results we need only parts of it.

**Theorem 2.1.** *Assume Condition 2.1. Then $\hat{Z}^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the unique (in law) diffusion process $\xi$ with generator*

$$(Lf)(x) = xcf'(x) + \frac{1}{2}x\kappa f''(x), \quad f \in C^2(\mathbb{R}_+), \quad x \in \mathbb{R}_+, \tag{4}$$

*and initial distribution (i.e. probability law of $\xi_0$) equal to $\mu_0$.*

We are concerned with the study of relationships between the “steady-state” behavior of $\hat{Z}^{(n)}$ and that of $\xi$. However, as noted earlier, the term “steady state” needs a careful interpretation. There are two well-studied approaches for formulating time asymptotic questions in the subcritical case. The first is to condition the processes $\hat{Z}^{(n)}$ on non-extinction, where, loosely speaking, the conditioning can either be on non-extinction at the present time or in the distant future. The state process $\hat{Z}^{(n)}$ under these two conditionings has different limiting distributions as $t \to \infty$. The first is called the Yaglom distribution of $\hat{Z}^{(n)}$, while the second is the stationary distribution of the Q-process associated with $\hat{Z}^{(n)}$ (see Sect. I.14 of [1]). The second approach for obtaining a nontrivial time asymptotic behavior is to introduce an immigration component. Namely, in each generation a (random) number of particles that are indistinguishable from the original set of particles are added to the population. The immigration component in particular ensures that the resulting scaled state process, denoted by $\hat{V}^{(n)}$, has a nondegenerate stationary distribution. For the supercritical case, a common approach is to reduce the problem to that of a subcritical setting by conditioning on the event of eventual extinction. The so conditioned state process $\hat{Z}^{(n)}$ has the same law as the state process corresponding to a certain subcritical BGW process. In this section we will show that the time asymptotic distribution of $\hat{Z}_t^{(n)}$ (in both subcritical and supercritical settings), under suitable conditioning, converges to that of $\xi_t$ under a similar conditioning, as $n \to \infty$. For models with immigration, we will prove convergence of stationary distributions.

We begin by describing results for models without immigration. Let $\mathbb{S}$ be a subset of $\mathbb{R}_+^k$, for some $k \in \mathbb{N}$. When $\mathbb{S}$ is endowed with a topology, we will denote by $\mathcal{B}(\mathbb{S})$ the $\sigma$-field generated by the open sets of $\mathbb{S}$. Let $\mathbf{Y} \equiv \{\mathbf{Y}_t\}_{t \in \mathbb{R}_+}$ be an $\mathbb{S}$-valued Markov process such that $\mathbf{0} \in \mathbb{S}$ is an absorbing state. If $\mathbf{Y}_0 = \mathbf{y}$, we write $P(\mathbf{Y}_t \in \cdot)$ as $P_{\mathbf{y}}(\mathbf{Y}_t \in \cdot)$. Similarly, when the distribution of $\mathbf{Y}_0$ is $\mu$, we write $P(\mathbf{Y}_t \in \cdot)$ as $P_\mu(\mathbf{Y}_t \in \cdot)$. Similar notations will be used for conditional expectations.

**Definition 2.1.** (i) A quasi-stationary distribution (qsd) for $\mathbf{Y}$ is a probability distribution $\mu$ on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ such that $P_\mu(\mathbf{Y}_t \in B | t < T_{\mathbf{Y}} < \infty) = \mu(B)$ for all $B \in \mathcal{B}(\mathbb{S})$ and $t \geq 0$, where $T_{\mathbf{Y}} := \inf\{t | \mathbf{Y}_t = \mathbf{0}\}$.

(ii) If for all $\mathbf{y} \in \mathbb{S} \setminus \{\mathbf{0}\}$, as $t \to \infty$, $P_{\mathbf{y}}(\mathbf{Y}_t \in \cdot | t < T_{\mathbf{Y}} < \infty)$ converges weakly to some probability measure $\mu$ on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$, then $\mu$ is called the Yaglom distribution of $\mathbf{Y}$.

The following result is a special case of results in [16] (pp. 77–78); see also [11] and Proposition 2.3.2.1 of [12].

**Theorem 2.2.** *The Yaglom distribution of $\xi$ exists and is exponential with density*

$$f(x) = \frac{2|c|}{\kappa} \exp\left(-\frac{2|c|}{\kappa} x\right), \quad x \geq 0. \tag{5}$$

Our first result, Theorem 2.3 below, says that the Yaglom distribution of $\hat{Z}^{(n)}$ approaches that of $\xi$, as $n \to \infty$.

**Theorem 2.3.** *Assume Condition 2.1. For each n, $\hat{Z}^{(n)}$ has a Yaglom distribution $\nu^{(n)}$. This distribution is also a qsd, and it converges weakly to the Yaglom distribution $\nu$ of $\xi$.*

We now consider the second form of conditioning where one conditions the process on not being extinct in the "distant future." We will see that in this case a somewhat different asymptotic behavior emerges. For this result we restrict ourselves to the subcritical case (i.e., $c < 0$). We begin with the definition of a Q-process (see [1,11]).

Let $\hat{\Omega} = D(\mathbb{R}_+ : \mathbb{R}_+)$ and $\hat{\mathcal{F}}$ be the corresponding Borel $\sigma$-field (with the usual Skorohod topology). Denote by $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ the canonical filtration on $(\hat{\Omega}, \hat{\mathcal{F}})$, i.e., $\mathcal{F}_t = \sigma(\pi_s : s \leq t)$, where $\pi_s(x) = x_s$ for $x \in \hat{\Omega}$. We denote by $\hat{P}_\mu^{(n)}$ the measure induced by $\hat{Z}^{(n)}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ when $Z_0^{(n)}$ has distribution $\mu$ (supported on $\mathbb{N}$). Let $T := \inf\{t | \pi_t = 0\}$.

It is easy to check (see [1], p. 58; also [13] and [3]) that there is a probability measure $P_\mu^{(n)\uparrow}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ such that, as $s \to \infty$, $\hat{P}_\mu^{(n)}(\Theta | T > s) \to P_\mu^{(n)\uparrow}(\Theta)$, for all $\Theta \in \mathcal{F}_t$, $t \in \mathbb{R}_+$. Furthermore, this unique measure on $(\hat{\Omega}, \hat{\mathcal{F}})$ can be characterized as follows. Let $\{Z_k^{(n)\uparrow}\}_{k \in \mathbb{N}_0}$ be a Markov chain with state-space $\mathbb{N}$, transition probabilities

$$P(Z_{l+1}^{(n)} = j | Z_l^{(n)} = i) \frac{j}{im^{(n)}}, \quad i, j \in \mathbb{N}, l \in \mathbb{N}_0,$$

and initial distribution $\mu$, and let $\hat{Z}_t^{(n)\uparrow} := \frac{1}{n} Z_{\lfloor nt \rfloor}^{(n)\uparrow}$, $t \in \mathbb{R}_+$. Then $P_\mu^{(n)\uparrow}$ is the law of $\{\hat{Z}_t^{(n)\uparrow}\}_{t \in \mathbb{R}_+}$. The process $Z^{(n)\uparrow}$ [respectively $\hat{Z}^{(n)\uparrow}$] is called the Q-process associated with $Z^{(n)}$ [respectively $\hat{Z}^{(n)}$]. Q-processes associated with branching processes can be interpreted as branching processes conditioned on being never extinct.

Next, we introduce the Q-process associated with the diffusion $\xi$ from Theorem 2.1. Denote by $P_{\xi,x}$ the measure induced by $\xi$ on $(\hat{\Omega}, \hat{\mathcal{F}})$, where $\xi_0 = x > 0$. The following theorem is contained in [11].

**Theorem 2.4.** *There is a probability measure $P_{\xi,x}^{\uparrow}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$, such that for all $t \in \mathbb{R}_+$ and $\Theta \in \mathcal{F}_t$, $P_{\xi,x}(\Theta | T > s)$ converges to $P_{\xi,x}^{\uparrow}(\Theta)$, as $s \to \infty$. Let $\xi^{\uparrow}$ be the unique weak solution of the SDE*

$$d\xi_t^{\uparrow} = c\xi_t^{\uparrow} dt + \sqrt{\kappa \xi_t^{\uparrow}} dB_t + \kappa dt, \quad \xi_0^{\uparrow} = x,$$

*where $B$ is a standard Brownian motion. Then $P_{\xi,x}^{\uparrow}$ equals the measure induced by $\xi^{\uparrow}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$.*

The process $\xi^{\uparrow}$ is referred to as the Q-process associated with $\xi$. The following result (see [11], Sect. 5.2) says that the process $\xi^{\uparrow}$ has a unique stationary distribution, $\nu^{\uparrow}$, which is given as the convolution of two copies of the exponential distribution $\nu$ with density as in (5).

**Theorem 2.5.** *Assume $c < 0$. As $t \to \infty$, for every initial condition $x$, $\xi_t^{\uparrow}$ converges in distribution to a random variable $\xi_\infty^{\uparrow}$, whose distribution, denoted by $\nu^{\uparrow}$, is the convolution of two copies of the Yaglom distribution $\nu$. In particular, $\nu^{\uparrow}$ has density*

$$f(x) = \left(\frac{2c}{\kappa}\right)^2 x \exp\left(\frac{2c}{\kappa} x\right), \quad x \geq 0. \tag{6}$$

Our next result shows that the time asymptotic behavior of the Q-process associated with $\hat{Z}^{(n)}$ can be well approximated by that of the Q-process associated with the diffusion approximation of $\hat{Z}^{(n)}$.

**Theorem 2.6.** *Assume Condition 2.1 and that $c_n < 0$ for all $n \in \mathbb{N}$. For each $n$, $\hat{Z}_t^{(n)\uparrow}$ converges in distribution, as $t \to \infty$, to a random variable $\hat{Z}_\infty^{(n)\uparrow}$. The distribution $\nu^{(n)\uparrow}$ of $\hat{Z}_\infty^{(n)\uparrow}$ is the unique stationary distribution of the $\mathbb{S}^{(n)}$-valued Markov process $\hat{Z}^{(n)\uparrow}$. As $n \to \infty$, $\nu^{(n)\uparrow}$ converges weakly to $\nu^{\uparrow}$.*

We now describe the results for BGW processes with immigration. Let $F$ and $G$ be pgf's of $\mathbb{N}_0$-valued random variables. A Bienaymé–Galton–Watson branching process with immigration corresponding to $(F,G)$ (referred to as a BPI$(F,G)$ process) is a Markov chain $\{V_n\}$ with state-space $\mathbb{N}_0$ and transition probability function described in terms of the corresponding pgf: Given $V_0 = i \in \mathbb{N}$, the pgf $H(i,\cdot)$ of $V_1$ is $H(i,s) = \sum_{j=0}^{\infty} P(V_1 = j | V_0 = i)s^j = F(s)^i G(s)$, $s \in [0,1]$.

Let $G^{(n)}$ be a sequence of pgf's of $\mathbb{N}_0$-valued random variables, and let $F^{(n)}$ be as in (2). Let $V^{(n)}$ be a sequence of BPI$(F^{(n)}, G^{(n)})$ processes.

**Condition 2.2.** (i) There is a $\iota_0 \in (0,\infty)$ such that for all $n \in \mathbb{N}$ $G^{(n)\prime}(1) = \iota^{(n)} \geq \iota_0$, and as $n \to \infty$, $\iota^{(n)} \to \iota$.
(ii) There is a $\beta_0 \in (0,\infty)$ such that for all $n \in \mathbb{N}$ $G^{(n)\prime\prime}(1) \leq \beta_0$.
(iii) There is a $\tau_0 \in (0,\infty)$ such that for all $n \in \mathbb{N}$ $F^{(n)\prime\prime\prime}(1) \leq \tau_0$.

Let $\hat{V}_t^{(n)} := \frac{1}{n} V_{\lfloor nt \rfloor}^{(n)}$, $t \in \mathbb{R}_+$. The proof of the following theorem is easy to establish using [11] and [14, Theorem 2.1].

**Theorem 2.7.** *Assume Conditions 2.1 and 2.2 and that $c < 0$. Suppose that $\hat{V}_0^{(n)}$ converges in distribution to some $\mu \in \mathcal{P}(\mathbb{R}_+)$. Then $\hat{V}^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the process $\zeta$ which is the unique weak solution of*

$$d\zeta_t = c\zeta_t dt + \sqrt{\kappa \zeta_t} dB_t + \iota dt, \quad t \geq 0,$$

*where $\zeta_0$ has distribution $\mu_0$. The Markov process $\zeta$ has a unique stationary distribution $\eta$, which is a gamma distribution with parameters $2\iota/\kappa$ and $\kappa/(2|c|)$, i.e., $\eta$ has density $g$ given as*

$$g(x) = x^{\frac{2\iota}{\kappa} - 1} \frac{\exp\left(-x\frac{2|c|}{\kappa}\right)}{\left(\frac{\kappa}{2|c|}\right)^{\frac{2\iota}{\kappa}} \Gamma\left(\frac{2\iota}{\kappa}\right)}, \quad x > 0.$$

We are interested in the longtime behavior of the scaled processes $\hat{V}^{(n)}$ as they approach criticality. Our main result in the single-type setting is the following.

**Theorem 2.8.** *Assume Conditions 2.1 and 2.2 and that $c_n < 0$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, $\hat{V}^{(n)}$ has a unique stationary distribution $\eta^{(n)}$ on $\mathbb{S}^{(n)}$, and as $n \to \infty$, $\eta^{(n)}$ converges weakly to $\eta$.*

As noted earlier, results similar to Theorems 2.3, 2.5, and 2.8 can be established for multitype settings as well. We only discuss one case in detail, namely the convergence of the Yaglom distribution in the setting of a subcritical multitype process. We begin with some notation and definitions. Let $\{\mathbf{Z}_j^{(n)}, j \in \mathbb{N}_0\}_{n \in \mathbb{N}}$ be a sequence of $k$-BGW processes with transition mechanism as described below. Let $C := [0,1]^k$, $\mathbf{e}_\alpha := (\delta_{1\alpha}, \ldots, \delta_{k\alpha})'$ be the $\alpha^{th}$ canonical basis vector and $\mathbf{s}^{\mathbf{i}} := \prod_{\alpha=1}^{k} s_\alpha^{i_\alpha}$, for $\mathbf{i} = (i_1, \ldots, i_k)' \in \mathbb{N}_0^k$ and $\mathbf{s} = (s_1, \ldots, s_k)' \in \mathbb{R}_+^k$. Similar to the single-type case, the evolution of $\mathbf{Z}_j^{(n)} = (Z_{j,1}^{(n)}, \cdots Z_{j,k}^{(n)})'$ is described as follows. For any

$\alpha = 1, \ldots, k$, each of the $Z^{(n)}_{j,\alpha}$ type $\alpha$ particles alive at time $j$ (if any) lives for one unit of time and then dies, giving rise to a number of offspring particles, represented by $\mathbf{l} = (l_1, \ldots, l_k)$, $l_\beta$ being the number of type $\beta$ offspring, with probability $\mu^{(n)}_\alpha(\mathbf{l})$. The particles behave independently of each other and of the past. The probability law of $\mathbf{Z}^{(n)}$ is given in terms of the pgf $\mathbf{F}^{(n)}(\mathbf{s}) := (F^{(n)}_{(1)}(\mathbf{s}), \ldots, F^{(n)}_{(k)}(\mathbf{s}))$, $\mathbf{s} \in C$, where

$$F^{(n)}_{(\alpha)}(\mathbf{s}) := \sum_{\mathbf{j} \in \mathbb{N}_0^k} \mu^{(n)}_\alpha(\mathbf{j})\mathbf{s}^{\mathbf{j}}, \quad 1 \le \alpha \le k.$$

Let $m^{(n)}_{\alpha\beta} = E_{\mathbf{e}_\alpha} Z^{(n)}_{1,\beta}$ be the expected number of type $\beta$ offspring from a single particle of type $\alpha$ in one generation. Then the $k \times k$ matrix $\mathbf{M}^{(n)} = (m^{(n)}_{\alpha\beta})_{\alpha,\beta=1,\ldots,k}$ is called the *mean matrix* of $\mathbf{Z}^{(n)}$. Note that $m^{(n)}_{\alpha\beta} = \frac{\partial F^{(n)}_{(\alpha)}}{\partial s_\beta}(\mathbf{1})$, where the partial derivative is understood to be the left-hand derivative. The processes $\mathbf{Z}^{(n)}$ will be assumed to have a *uniformly strictly positive* mean matrix $\mathbf{M}^{(n)}$, by which we mean that there exist $U \in \mathbb{N}$ and $a \in (0, \infty)$ such that for every $n \ge 1$ $((\mathbf{M}^{(n)})^U)_{\alpha,\beta} \ge a$ for all $1 \le \alpha, \beta \le k$. From the Perron–Frobenius Theorem, it then follows that $\mathbf{M}^{(n)}$ has a real, positive maximal eigenvalue $\rho^{(n)}$ with associated positive left and right eigenvectors $\mathbf{v}^{(n)}$ and $\mathbf{u}^{(n)}$, respectively, which, without loss of generality, are normalized so that $\mathbf{u}^{(n)\prime}\mathbf{v}^{(n)} = 1$ and $\mathbf{u}^{(n)\prime}\mathbf{1} = 1$ (see [1]). The maximal eigenvalue $\rho^{(n)}$ plays a similar role in the classification of the $k$-BGW process as the mean played in classifying the (single type) BGW process. The $k$-BGW process is called subcritical, critical, or supercritical, according to whether $\rho^{(n)} < 1$, $\rho^{(n)} = 1$, or $\rho^{(n)} > 1$, respectively. We will consider the subcritical case, namely for all $n \ge 1$ $\rho^{(n)} \in (0,1)$, and study the behavior of quasi-stationary and Yaglom distributions of the scaled process

$$\hat{\mathbf{Z}}^{(n)}_t = \frac{1}{n}\mathbf{Z}^{(n)}_{\lfloor nt \rfloor}, \quad t \ge 0,$$

as $\rho^{(n)} \to 1$. The existence of the Yaglom distribution of $\hat{\mathbf{Z}}^{(n)}$ is assured by the following result.

**Condition 2.3.** For each $n \ge 1$, $E_{\mathbf{1}}(||\mathbf{Z}^{(n)}_1|| \log ||\mathbf{Z}^{(n)}_1||) < \infty$.

**Theorem 2.9.** *Assume Condition 2.3. For each $n \in \mathbb{N}$, $\hat{\mathbf{Z}}^{(n)}$ has a Yaglom distribution $\nu^{(n)}$. This distribution is also a qsd.*

**Condition 2.4.** There exist $b, d \in (0, \infty)$ such that for all $n \in \mathbb{N}$

(i) $\sum_{\alpha\beta\gamma} \partial^2 F^{(n)}_{(\alpha)}(\mathbf{1})/\partial s_\beta \partial s_\gamma \ge b$ and

(ii) $\sum_{\alpha,\beta,\gamma,\delta} \partial^3 F^{(n)}_{(\alpha)}(\mathbf{1})/\partial s_\beta \partial s_\gamma \partial s_\delta \le d$,

where $\alpha, \beta, \gamma, \delta$ in the above sums vary over $\{1, \ldots, k\}$.

Part (i) of the condition can be interpreted as a nondegeneracy condition, and part (ii) says that the third moments of the offspring distributions are uniformly bounded in $n$.

We now introduce a condition, analogous to Condition 2.1 (ii), for the multitype setting.

**Condition 2.5.** For some strictly positive matrix $\mathbf{M}$ and each $n \in \mathbb{N}$, $\mathbf{M}^{(n)} = \mathbf{M} + \frac{\mathbf{C}^{(n)}}{n}$, and $\lim_{n\to\infty} \mathbf{C}^{(n)} = \mathbf{C}$. The maximal eigenvalues $\rho^{(n)}$ of $\mathbf{M}^{(n)}$ are of the form $\rho^{(n)} = 1 + \frac{c^{(n)}}{n}$, with $c^{(n)} \in (-n, 0)$ and $\lim_{n\to\infty} c^{(n)} = c \in (-\infty, 0)$. Moreover, $\mathbf{M}$ has maximal eigenvalue 1 with corresponding eigenvectors $\mathbf{v} = \lim \mathbf{v}^{(n)}$ and $\mathbf{u} = \lim \mathbf{u}^{(n)}$. Finally, $\mathbf{v}'\mathbf{C}\mathbf{u} = c$.

*Example 2.1.* Let $\mathbf{C}^{(n)} = c^{(n)}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $c^{(n)} \in (-n, 0)$ such that $c^{(n)} \to c \in (-\infty, 0)$. Let $\mathbf{M}$ be a strictly positive matrix with maximal eigenvalue equal to 1. Then $\mathbf{M}^{(n)} = \mathbf{M} + \frac{\mathbf{C}^{(n)}}{n}$ satisfies Condition 2.5.

Let

$$\kappa_{i,j}^{(n)}(l) := \sum_{\mathbf{r}\in\mathbb{N}_0^k} (r_i - m_{li}^{(n)})(r_j - m_{lj}^{(n)})\mu_l^{(n)}(\mathbf{r}).$$

We need the following condition on the variance of the offspring distributions.

**Condition 2.6.** As $n \to \infty$, $\kappa_{i,j}^{(n)}(l) \to \kappa_{i,j}(l)$ for all $1 \le i, j, l \le k$ and $Q := \frac{1}{2}\sum_{l=1}^{k} v_l \mathbf{u}'\boldsymbol{\kappa}(l)\mathbf{u} \in (0, \infty)$, where $\boldsymbol{\kappa}(l)$ is the matrix with $(i, j)$th entry $\kappa_{i,j}(l)$.

The following diffusion approximation result can be established along the lines of Theorem 4.3.1 of [8] and Theorem 9.2.1 of [5]. Let $C_c^\infty(\mathbb{R}_+)$ be the space of infinitely differentiable, real-valued functions with compact support on $\mathbb{R}_+$.

**Theorem 2.10.** *Assume Conditions 2.4, 2.5, and 2.6. Suppose that the distribution of $\hat{\mathbf{Z}}_0^{(n)}$ converges to some $\mu_0 \in \mathcal{P}(\mathbb{R}_+^k)$. Let $\mu_1 \in \mathcal{P}(\mathbb{R}_+)$ be given as*

$$\mu_1(A) = \mu_0\{\mathbf{x} \in \mathbb{R}_+^k | \mathbf{x}'\mathbf{u} \in A\}, \quad A \in \mathcal{B}(\mathbb{R}_+). \tag{7}$$

*Let $\zeta^{(n)} = \hat{\mathbf{Z}}^{(n)'}\mathbf{u}^{(n)}$. Then $\zeta^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the unique (in law) diffusion $\zeta$ with initial distribution $\mu_1$ and generator $\tilde{L}$ given as*

$$(\tilde{L}f)(x) = cxf'(x) + Qxf''(x), \quad f \in C_c^\infty(\mathbb{R}_+), \quad x \in \mathbb{R}_+. \tag{8}$$

*Furthermore, for any $t_0 \in (0, \infty)$, the process $\hat{\mathbf{Z}}^{(n,t_0)}$, defined by $\hat{\mathbf{Z}}_t^{(n,t_0)} := \hat{\mathbf{Z}}_{t_0+t}^{(n)}$, $t \ge 0$, converges weakly to $\mathbf{Z}^{(t_0)} := \mathbf{v}\zeta^{(t_0)}$, where $\zeta_t^{(t_0)} := \zeta_{t_0+t}$, $t \ge 0$.*

The process $\mathbf{Z}^{(t_0)}$ is a Markov process with state-space $S_\mathbf{v} = \{\theta\mathbf{v} | \theta \ge 0\}$ and can be formally regarded as the limit of $\hat{\mathbf{Z}}^{(n)}$. Indeed, if the support of $\mu_0$ is contained in $S_\mathbf{v}$, then, noting that $\mathbf{u}'\mathbf{v} = 1$, we see that the law of $\mathbf{v}\zeta_0$ equals $\mu_0$ and that in fact $\hat{\mathbf{Z}}^{(n)}$

converges weakly to $\mathbf{v}\zeta$, where $\zeta$ is as in Theorem 2.10. We will be concerned with the Yaglom distribution of the $S_{\mathbf{v}}$-valued Markov process $\mathbf{Z}^{(t_0)}$ and its relation to the Yaglom distribution of $\hat{\mathbf{Z}}^{(n)}$. For that it will be convenient to regard a probability measure on $S_{\mathbf{v}}$ as one on $\mathbb{R}^k_+$. Denote by $\tilde{\nu}$ the exponential distribution with density $f(x) = -cQ^{-1}\exp(cQ^{-1}x)$, $x \geq 0$. Theorem 2.2 says that the Yaglom distribution of $\zeta^{(t_0)}$ is given by $\tilde{\nu}$. Since $\mathbf{Z}^{(t_0)} = \mathbf{v}\zeta^{(t_0)}$, the Yaglom distribution of $\mathbf{Z}^{(t_0)}$ exists as well and equals the distribution of $\mathbf{v}Y$, where $Y$ has distribution $\tilde{\nu}$. Thus, we have the following:

**Theorem 2.11.** *The Yaglom distribution of $\zeta^{(t_0)}$ exists and equals $\tilde{\nu}$. Furthermore, the Yaglom distribution of $\mathbf{Z}^{(t_0)}$, denoted by $\bar{\nu}$, exists and equals the distribution of $\mathbf{v}Y$, where $Y$ has distribution $\tilde{\nu}$.*

The following result relates the qsd's and Yaglom distributions of $\hat{\mathbf{Z}}^{(n)}$ to that of its "diffusion limit" $\mathbf{Z}^{(t_0)}$. Probability distributions similar to $\bar{\nu}$ have previously been noted in the study of qsd's of multitype BGW processes. In [1] (p. 191), a single critical BGW process $\mathbf{Z}$ (rather than a sequence of near critical BGW processes) is considered and it is shown that $\mathbf{Z}_n/n$ conditioned on non-extinction converges to a random variable that is concentrated on the ray $\{x\mathbf{v_Z} | x \geq 0\}$, where $\mathbf{v_Z}$ is the left eigenvector of the mean matrix of $\mathbf{Z}$ corresponding to the eigenvalue 1. In [17] (see Theorem 3 therein) the case where $\mathbf{Z}$ is near critical and a somewhat differently (component wise) scaled process $\mathbf{Z}^*$ is considered. The asymptotic behavior of $\mathbf{Z}^*_n$ conditioned on non-extinction, as $n \to \infty$ and the offspring distribution approaches criticality, is related to the limiting distributions considered here. In fact, we use an estimate from [17] to prove Theorem 2.12, below. We remark that none of these results concern the setting of diffusion approximation, where time and space are scaled and one starts with a large number of particles.

**Theorem 2.12.** *The Yaglom distribution $\nu^{(n)}$ of $\hat{\mathbf{Z}}^{(n)}$ converges weakly to the Yaglom distribution $\bar{\nu}$ of $\mathbf{Z}^{(t_0)}$.*

## 3 Catalyst-Reactant Branching Processes with Controlled Immigration

The particles in the Bienaymé–Galton–Watson processes considered in the previous section evolved independently of each other. In this section, we consider catalytic branching processes that model the dynamics of catalyst-reactant populations in which the activity level of the reactant depends on the amount of catalyst present. Branching processes in catalytic environment have been studied extensively and are motivated, for instance by biochemical reactions (see [4, 7, 9, 15], and references therein). A typical setting consists of populations of multiple types such that the rate of growth (depletion) of one population type is directly affected by population sizes of other types. The simplest such model consists of a continuous-time countable-state branching process describing the evolution of the catalyst population and a second branching process for which the branching rate is proportional to the total mass of the catalyst population modeling the evolution of reactant particles. Such processes were introduced in [4] in the setting of super-Brownian motions [15]. For classical catalyst-reactant branching processes, the catalyst population dies out with positive probability and subsequent to the catalyst extinction, the reactant population stays unchanged and therefore the population dynamics are modeled until the time the catalyst becomes extinct. In this work, we consider a setting where the catalyst population is maintained above a positive threshold through a specific form of controlled immigration. Branching process models with immigration have also been well studied in literature (see [1, 15] and references therein). However, typical mechanisms that have been considered correspond to adding an independent Poisson component (see, e.g., [10]). Here, instead, we consider a model where immigration takes place only when the population drops below a certain threshold. Roughly speaking, we consider a sequence $\{X^{(n)}\}_{n\in\mathbb{N}}$ of continuous-time branching processes, where $X^{(n)}$ starts with $n$ particles. When the population drops below $n$, it is instantaneously restored to the level $n$.

There are many settings where controlled immigration models of the above form arise naturally. One class of examples arise from predator-prey models in ecology, where one may be concerned with the restoration of populations that are close to extinction by reintroducing species when they fall below a certain threshold. In our work, the motivation for the study of such controlled immigration models comes from problems in chemical reaction networks where one wants to keep the levels of certain types of molecules above a threshold in order to maintain a desired level of production (or inhibition) of other chemical species in the network. Such questions are of interest in the study of control and regulation of chemical reaction networks. A control action where one minimally adjusts the levels of one chemical type to keep it above a fixed threshold is one of the simplest regulatory mechanism and the goal of this research is to study system behavior under such mechanisms with the long-term objective of designing optimal control policies. The specific goal of the current work is to derive simpler approximate and reduced models, through the theory of diffusion approximations and stochastic averaging techniques, that are more tractable for simulation and mathematical treatment than the original branching process models. In order to keep the presentation simple, we consider the setting of one catalyst and one reactant. However similar limit theorems can be obtained for a more general chemical reaction network in which the levels of some of the chemical species are regulated in a suitable manner. Settings where some of the chemical species act as inhibitors rather than catalysts are also of interest and can be studied using similar techniques. These extensions will be pursued elsewhere.

Our main goal is to establish diffusion approximations for such regulated catalyst-reactant systems under suitable scalings. We consider two different scaling regimes; in the first setting the catalyst and reactant evolve on "comparable timescales," while in the second setting the catalyst evolves "much faster" than the reactant. In the former setting, the limit model is described through a coupled system of reflected stochastic differential equations with reflection in the space

$[1,\infty) \times \mathbb{R}$. The precise result (Theorem 3.2) is stated in Sect. 3.1. Such limit theorems are of interest for various analytic and computational reasons. It is simpler to simulate (reflected) diffusions than branching processes, particularly for large network settings. Analytic properties such as hitting time probabilities and steady-state behavior are more easily analyzed for the diffusion models than for their branching process counterparts. In general, such diffusion limits give parsimonious model representations and provide useful qualitative insight to the underlying stochastic phenomena.

For the second scaling regime, where the catalyst evolution is much faster, we establish a stochastic averaging limit theorem. The limit evolution of the reactant population is given through an autonomous one-dimensional SDE with coefficients that depend on the stationary distribution of a reflected diffusion in $[1,\infty)$. Such model reductions are important in that they not only help in better understanding the dynamics of the system but also help in reducing computational costs in simulations. Indeed, since in the model considered here the invariant distribution is explicit, the coefficients in the one-dimensional averaged diffusion model are easily computed and consequently this model is significantly easier to analyze and simulate than the original two-dimensional model. We refer the reader to [9] and references therein for similar results in the setting of (nonregulated) chemical reaction networks. It will be of interest to see if similar model reductions can be obtained for general multidimensional regulated chemical reaction networks.

## 3.1 Diffusion Limit under Comparable Timescales

Consider a sequence of pairs of continuous-time, countable-state Markov branching processes $(X^{(n)}, Y^{(n)})$, where $X^{(n)}$ and $Y^{(n)}$ represent the number of catalyst and reactant particles, respectively. The dynamics are described as follows. Each of the $X_t^{(n)}$ particles alive at time $t$ has an exponentially distributed life time with parameter $\lambda_1^{(n)}$ (mean life time $1/\lambda_1^{(n)}$). When it dies, each such particle gives rise to a number of offspring, according to the offspring distribution $\mu_1^{(n)}(\cdot)$. Additionally, if the catalyst population drops below $n$, it is instantaneously replenished back to the level $n$ (*controlled immigration*). The branching rate of the reactant process $Y^{(n)}$ is of the order of the current total mass of the catalyst population, i.e., $X^{(n)}/n$, and we denote the offspring distribution of $Y^{(n)}$ by $\mu_2^{(n)}(\cdot)$. A precise definition of the pair $(X^{(n)}, Y^{(n)})$ will be given below. We are interested in the study of asymptotic behavior of $(X^{(n)}, Y^{(n)})$, under suitable scaling, as $n \to \infty$.

We now give a precise description of the various processes and the scaling that is considered. Roughly speaking, time is accelerated by a factor of $n$, and mass is scaled down by a factor of $n$. Define RCLL processes

$$\hat{\mathbf{W}}_t^{(n)} := \left(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)}\right) := \left(\frac{X_{nt}^{(n)}}{n}, \frac{Y_{nt}^{(n)}}{n}\right), \quad t \in \mathbb{R}_+. \tag{9}$$

Let $\mathbb{S}_X^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\} \cap [1,\infty)$, $\mathbb{S}_Y^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\}$, $\mathbb{W}^{(n)} := \mathbb{S}_X^{(n)} \times \mathbb{S}_Y^{(n)}$, $\mathbb{W} = [1,\infty) \times \mathbb{R}_+$, and

$$D(\mathbb{R}_+ : S) := \{f : \mathbb{R}_+ \to S | f \text{ is right continuous and has left limits}\}$$

endowed with the usual Skorohod topology. Assume that $(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}) = (x_0^{(n)}, y_0^{(n)}) \in \mathbb{W}^{(n)}$. Then $\{\hat{\mathbf{W}}_t^{(n)}\}_{t \in \mathbb{R}_+}$ is characterized as the $\mathbb{W}^{(n)}$-valued Markov process with sample paths in $D(\mathbb{R}_+ : \mathbb{W}^{(n)})$, starting at $\hat{\mathbf{W}}_0^{(n)} = (x_0^{(n)}, y_0^{(n)})$, and having infinitesimal generator $\hat{\mathcal{A}}^{(n)}$ given as

$$\begin{aligned}\hat{\mathcal{A}}^{(n)} \phi(x,y) = \lambda_1^{(n)} n^2 x \sum_{k=0}^{\infty} \left[\phi\left(\left(1 \vee x + \frac{k-1}{n}, y\right) - \phi(x,y)\right]\mu_1^{(n)}(k) \\ + \lambda_2^{(n)} n^2 xy \sum_{k=0}^{\infty} \left[\phi\left(x, y + \frac{k-1}{n}\right) - \phi(x,y)\right]\mu_2^{(n)}(k),\end{aligned} \tag{10}$$

where $(x,y) \in \mathbb{W}^{(n)}$ and $\phi \in \mathrm{BM}(\mathbb{W})$ with $\mathrm{BM}(\mathbb{W})$ being the space of bounded, measurable, real-valued functions $\mathbb{W}$. From the definition of the generator, we see that, for each $k \geq 0$, given $\hat{\mathbf{W}}_t^{(n)} = (x,y) \in \mathbb{W}^{(n)}$, the process jumps to $(x, y + \frac{k-1}{n})$ with rate $\lambda_2^{(n)} n^2 xy \mu_2^{(n)}(k)$ and to $(x + \frac{k-1}{n}, y)$ with rate $\lambda_1^{(n)} n^2 x \mu_1^{(n)}(k)$, except when $k = 0$ and $x = 1$, in which case the latter jump is to $(x,y)$ with rate $\lambda_1^{(n)} n^2 \mu_1^{(n)}(0)$. This property of the generator at $x = 1$ accounts for the instantaneous replenishment of the (unscaled) catalyst population to level $n$, whenever the catalyst drops below $n$.

For $i = 1,2$, let

$$m_i^{(n)} := \sum_{k=0}^{\infty} k \mu_i^{(n)}(k) \quad \text{and} \quad \alpha_i^{(n)} = \sum_{k=0}^{\infty} (k-1)^2 \mu_i^{(n)}(k).$$

We make the following basic assumption on the parameters of the branching rates and offspring distributions as well as on the initial configurations of the catalyst and reactant populations.

**Condition 3.1.** (i) For $i = 1,2$ and for $n \in \mathbb{N}$, $\alpha_i^{(n)}, \lambda_i^{(n)} \in (0,\infty)$ and $m_i^{(n)} = 1 + \frac{c_i^{(n)}}{n}$, $c_i^{(n)} \in (-n,\infty)$.

(ii) For $i = 1,2$, as $n \to \infty$, $c_i^{(n)} \to c_i \in \mathbb{R}$, $\alpha_i^{(n)} \to \alpha_i \in (0,\infty)$, and $\lambda_i^{(n)} \to \lambda_i \in (0,\infty)$.

(iii) For $i = 1,2$ and for every $\varepsilon \in (0,\infty)$,

$$\lim_{n\to\infty}\sum_{l:l>\varepsilon\sqrt{n}}(l-m_i^{(n)})^2\mu_i^{(n)}(l)=0.$$

(iv) As $n\to\infty$, $(x_0^{(n)},y_0^{(n)})\to(x_0,y_0)\in[1,\infty)\times\mathbb{R}_+$.

Condition 3.1 and the form of the generator in (10) ensure that the scaled catalyst and reactant processes transition on comparable timescales, namely, $\mathcal{O}(n^2)$.

In order to state the limit theorem for $(\hat{X}^{(n)},\hat{Y}^{(n)})$, we need some notation and definitions associated with the one-dimensional Skorohod map with reflection at 1. Let $D_1(\mathbb{R}_+:\mathbb{R}):=\{f\in D(\mathbb{R}_+:\mathbb{R})|f(0)\geq 1\}$, and let $\Gamma:D_1(\mathbb{R}_+:\mathbb{R})\to D(\mathbb{R}_+:[1,\infty))$ be defined as

$$\Gamma(\psi)(t):=(\psi(t)+1)-\inf_{0\leq s\leq t}\{\psi(s)\wedge 1\},\quad \text{for } \psi\in D(\mathbb{R}_+:\mathbb{R}). \tag{11}$$

The function $\Gamma$, known as Skorohod map, can be characterized as follows: If $\psi,\phi,\eta^*\in D(\mathbb{R}_+:\mathbb{R})$ are such that (i) $\psi(0)\geq 1$, (ii) $\phi=\psi+\eta^*$, (iii) $\phi\geq 1$, (iv) $\eta^*$ is nondecreasing, $\int_{[0,\infty)}1_{\{\phi(s)\neq 1\}}d\eta^*(s)=0$, and $\eta^*(0)=0$, then $\phi=\Gamma(\psi)$ and $\eta^*=\phi-\psi$. The process $\eta^*$ can be regarded as the reflection term that is applied to the original trajectory $\psi$ to produce a trajectory $\phi$ that is constrained to $[1,\infty)$. From the definition of the Skorohod map and using the triangle inequality, we get the following Lipschitz property: For $\psi,\tilde{\psi}\in D_1(\mathbb{R}_+:\mathbb{R})$,

$$\sup_{s\leq t}|\Gamma(\psi)(s)-\Gamma(\tilde{\psi})(s)|\leq 2\sup_{s\leq t}|\psi(s)-\tilde{\psi}(s)|. \tag{12}$$

Let

$$\hat{\eta}_t^{(n)}:=\lambda_1^{(n)}n\mu_1^{(n)}(0)\int_0^t 1_{\{\hat{X}_s^{(n)}=1\}}ds. \tag{13}$$

This process will play the role of the reflection term, in the dynamics of the catalyst, arising from the controlled immigration. The diffusion limit of $(\hat{X}^{(n)},\hat{Y}^{(n)})$ will be the process $(X,Y)$ which is characterized in the following proposition through a system of stochastic integral equations.

The diffusion limit of $(\hat{X}^{(n)},\hat{Y}^{(n)})$ will be the process $(X,Y)$, starting at $(x_0,y_0)$, which is given through a system of stochastic integral equations as in the following proposition.

**Proposition 3.1.** *Let $(\bar{\Omega},\bar{\mathcal{F}},\bar{P},\{\bar{\mathcal{F}}_t\})$ be a filtered probability space on which are given independent standard $\{\bar{\mathcal{F}}_t\}$ Brownian motions $B^X$ and $B^Y$. Let $X_0,Y_0$ be square integrable $\bar{\mathcal{F}}_0$ measurable random variables with values in $[1,\infty)$ and $\mathbb{R}_+$, respectively. Then the following system of stochastic integral equations has a unique strong solution:*

$$X_t=\Gamma\left(X_0+\int_0^{\cdot}c_1\lambda_1X_sds+\int_0^{\cdot}\sqrt{\alpha_1\lambda_1X_s}dB_s^X\right)(t), \tag{14}$$

$$Y_t = Y_0 + \int_0^t c_2\lambda_2 X_s Y_s ds + \int_0^t \sqrt{\alpha_2\lambda_2 X_s Y_s} dB_s^Y, \tag{15}$$

$$\eta_t = X_t - X_0 - \int_0^t c_1\lambda_1 X_s ds - \int_0^t \sqrt{\alpha_1\lambda_1 X_s} dB_s^X, \tag{16}$$

*where $\Gamma$ is the Skorohod map defined in* (11).

In the above proposition, by a strong solution of (14)–(16) we mean an $\bar{\mathcal{F}}$-adapted continuous process $(X,Y,\eta)$ with values in $[1,\infty)\times\mathbb{R}_+\times\mathbb{R}_+$ that satisfies (14)–(16). The following is the main result of this subsection.

**Theorem 3.2.** *Suppose Condition 3.1 holds. The process $(\hat{X}^{(n)}, \hat{Y}^{(n)})$ converges weakly in $D(\mathbb{R}_+ : \mathbb{W})$ to the process $(X,Y)$ given in Proposition 3.1 with $(X_0,Y_0) = (x_0,y_0)$.*

## *3.2 Asymptotic Behavior of the Catalyst Population*

Stochastic averaging results in this work rely on understanding the time asymptotic behavior of the catalyst process. Such behavior, of course, is also of independent interest. We begin with the following result on the stationary distribution of $X$, where $X$ is the reflected diffusion from Proposition 3.1, approximating the catalyst dynamics (Theorem 3.2). The proof uses an extension of the Echeverria criterion for stationary distributions of diffusions to the setting of constrained diffusions (see [2, 18] and references therein). We will make the following additional assumption. Recall the constants $c_1^{(n)} \in (-n,\infty)$ and $c_1 \in \mathbb{R}$ introduced in Condition 3.1.

**Condition 3.2.** For all $n \in \mathbb{N}$, $c_1^{(n)} < 0$ and $c_1 < 0$.

**Proposition 3.2.** *Suppose Condition 3.2 holds. The process $X$ defined through (14) has a unique stationary distribution, $\nu_1$, which has density*

$$p(x) := \begin{cases} \frac{\theta}{x}\exp(2\frac{c_1}{\alpha_1}x), & \text{if } x \geq 1 \\ 0, & \text{if } x < 1, \end{cases} \tag{17}$$

*where $\theta := \left(\int_1^\infty (\frac{1}{x}\exp(2\frac{c_1}{\alpha_1}x))dx\right)^{-1}$.*

The following result shows that the time asymptotic behavior of the catalyst population is well approximated by that of its diffusion approximation given through (14). We make the following additional assumption on the moment generating function of the offspring distribution, which will allow us to construct certain "uniform Lyapunov functions" that play a key role in the analysis.

**Condition 3.3.** For some $\bar{\delta} > 0$,

$$\sup_{n\in\mathbb{N}} \sum_{k=0}^{\infty} e^{\bar{\delta}k} \mu_1^{(n)}(k) < \infty. \tag{18}$$

**Theorem 3.2.** *Suppose Conditions 3.1, 3.2, and 3.3 hold. Then, for each $n \in \mathbb{N}$, the process $\hat{X}^{(n)}$ has a unique stationary distribution $\nu_1^{(n)}$, and the family $\{\nu_1^{(n)}\}_{n\in\mathbb{N}}$ is tight. As $n \to \infty$, $\nu_1^{(n)}$ converges weakly to $\nu_1$.*

## *3.3 Diffusion Limit of the Reactant under Fast Catalyst Dynamics*

As noted in Sect. 3.1, the catalyst and reactant populations whose scaled evolution is described through (10) transition on comparable timescales. In situations in which the catalyst evolves "much faster" than the reactant, one can hope to find a simplified model that captures the dynamics of the reactant population in a more economical fashion. One would expect that the reactant population can be approximated by a diffusion whose coefficients depend on the catalyst only through the catalyst's stationary distribution. Indeed, we will show that the (scaled) reactant population can be approximated by the solution of

$$\check{Y}_t = \check{Y}_0 + \int_0^t c_2 \lambda_2 m_X \check{Y}_s ds + \int_0^t \sqrt{\alpha_2 \lambda_2 m_X \check{Y}_s} dB_s, \quad \check{Y}_0 = y_0, \tag{19}$$

where $m_X = \int_1^\infty x \nu_1(dx) = -\frac{\alpha_1 \theta}{2c_1} \exp(2c_1/\alpha_1)$.

Such model reductions (see [9] and references therein for the setting of chemical reaction networks) not only help in better understanding the dynamics of the system but also help in reducing computational costs in simulations. In this section, we will consider such stochastic averaging results in two model settings. First, in Sect. 3.3.1, we consider the simpler setting where the population mass evolutions are described through (reflected) stochastic integral equations and a scaling parameter in the coefficients of the model distinguishes the timescales of the two processes. In Sect. 3.3.2, we will consider a setting which captures the underlying physical dynamics more accurately in the sense that the mass processes are described in terms of continuous-time branching processes, rather than diffusions.

### 3.3.1 Stochastic Averaging in a Diffusion Setting

In this section, we consider the setting where the catalyst and reactant populations evolve according to (reflected) diffusions similar to $X$ and $Y$ from Proposition 3.1, but where the evolution of the catalyst is accelerated by a factor of $a_n$ such that $a_n \uparrow \infty$ as $n \uparrow \infty$ (i.e. drift and diffusion coefficients are scaled by $a_n$). More precisely, we consider a system of catalyst and reactant populations that are given as solutions of the following system of stochastic integral equations: For $t \geq 0$,

$$\check{X}_t^{(n)} = \Gamma \left( \check{X}_0^{(n)} + \int_0^{\cdot} a_n c_1 \lambda_1 \check{X}_s^{(n)} ds + \int_0^{\cdot} \sqrt{a_n \alpha_1 \lambda_1 \check{X}_s^{(n)}} dB_s^X \right)(t)$$

$$\check{Y}_t^{(n)} = \check{Y}_0^{(n)} + \int_0^t c_2 \lambda_2 \check{X}_s^{(n)} \check{Y}_s^{(n)} ds + \int_0^t \sqrt{\alpha_2 \lambda_2 \check{X}_s^{(n)} \check{Y}_s^{(n)}} dB_s^Y,$$

where $(\check{X}_0^{(n)}, \check{Y}_0^{(n)}) = (x_0, y_0)$, $c_1, c_2 \in \mathbb{R}$, $\alpha_i, \lambda_i \in (0, \infty)$, $B^X$ and $B^Y$ are independent standard Brownian motions and $\Gamma$ is the Skorohod map described above Proposition 3.1.

The following result says that if $c_1 < 0$, then the reactant population process $\check{Y}^{(n)}$, which is given through a coupled two-dimensional system, can be well approximated by the one-dimensional diffusion $\check{Y}$ in (19), whose coefficients are given in terms of the stationary distribution of the catalyst process.

**Theorem 3.3.** *Suppose Condition 3.2 holds. The process $\check{Y}^{(n)}$ converges weakly in $C(\mathbb{R}_+ : \mathbb{R}_+)$ to the process $\check{Y}$.*

### 3.3.2 Stochastic Averaging for Scaled Branching Processes

We now consider stochastic averaging for the setting where the catalyst and reactant populations are described through branching processes. Consider catalyst and reactant populations evolving according to the branching processes introduced in Sect. 3.1, but where the catalyst evolution is sped up by a factor of $a_n$ such that $a_n \uparrow \infty$ monotonically as $n \uparrow \infty$. That is, we consider a sequence of catalyst populations $\tilde{X}_t^{(n)} := X_{a_n t}^{(n)}$, $t \geq 0$, where $X^{(n)}$ are the branching processes introduced in Sect. 3.1. The reactant population evolves according to a branching process, $\tilde{Y}^{(n)}$, whose branching rate, as before, is of the order of the current total mass of the catalyst population, $\tilde{X}^{(n)}/n$. The infinitesimal generator $\check{\mathcal{G}}^{(n)}$ of the scaled process

$$\left( \check{X}_t^{(n)}, \check{Y}_t^{(n)} \right) := \left( \frac{1}{n} \tilde{X}_{nt}^{(n)}, \frac{1}{n} \tilde{Y}_{nt}^{(n)} \right), \quad t \geq 0,$$

is given as

$$\check{\mathcal{G}}^{(n)} \phi(x,y) = \lambda_1^{(n)} n^2 a_n x \sum_{k=0}^{\infty} \left[ \phi\left( 1 \vee \left( x + \frac{k-1}{n} \right), y \right) - \phi(x,y) \right] \mu_1^{(n)}(k)$$
$$+ \lambda_2^{(n)} n^2 xy \sum_{k=0}^{\infty} \left[ \phi\left( x, y + \frac{k-1}{n} \right) - \phi(x,y) \right] \mu_2^{(n)}(k), \qquad (20)$$

where $(x,y) \in \mathbb{S}_X^{(n)} \times \mathbb{S}_Y^{(n)}$ and $\phi \in \mathrm{BM}([1,\infty) \times \mathbb{R}_+)$.

We note that a key difference between the generators $\check{\mathcal{G}}^{(n)}$ above and $\hat{\mathcal{A}}^{(n)}$ in (10) is the extra factor of $a_n$ in the first term of (20), which says that, for large $n$, the catalyst dynamics are much faster than that of the reactant.

We will show in Theorem 3.4 that the reactant population process $\check{Y}^{(n)}$ can be well approximated by the one-dimensional diffusion $\check{Y}$ in (19). Once again, the result provides a model reduction that is potentially useful for simulations and also for a general qualitative understanding of reactant dynamics near criticality.

**Theorem 3.4.** *Suppose Conditions 3.1, 3.2, and 3.3 hold. Then, as $n \to \infty$, $\check{Y}^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the process $\check{Y}$.*

**Acknowledgements** We gratefully acknowledge the valuable feedback from the referee which, in particular, led to a simplification of Condition 2.1.

## References

1. K. B. Athreya and P. Ney. *Branching Processes*. Springer-Verlag, 1972.
2. A. Budhiraja and D. Reinhold. Near critical catalyst reactant branching processes with controlled immigration. *Annals of Applied Probability (to appear),* arXiv:1203.6879 [math.PR].
3. A. Budhiraja and D. Reinhold. Some asymptotic results for near critical branching processes. *Communications on Stochastic Analysis*, Vol. 4, No. 1:91–113, 2010.
4. D. A. Dawson and K. Fleischmann. A continuous super-brownian motion in a super-brownian medium. *Journal of Theoretical Probability*, 10, No. 1:213–276, 1997.
5. S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.
6. W. Feller. Diffusion processes in genetics. *Proc. Second Berkeley Symp. on Math. Statist. and Prob.*, pages 227–246, 1951.
7. A. Greven, L. Popovic, and A. Winter. Genealogy of catalytic branching models. *The Annals of Applied Probability*, 19:1232–1272, 2009.
8. A. Joffe and M. Metivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, Vol. 18:pp. 20–65, 1986.
9. H.-W. Kang and T. G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Annals of Applied Probability (to appear),* arXiv:1011.1672 [math.PR].
10. S. Karlin. *A First Course in Stochastic Processes*. Academic Press, 1966.
11. A. Lambert. Quasi-stationary distributions and the continuous branching process conditioned to be never extinct. *Electronic Journal of Probability*, 12:420–446, 2007.
12. A. Lambert. Population dynamics and random genealogies. *Stochastic Models*, 24:45–163, 2008.
13. J. Lamperti and P. Ney. Conditioned branching processes and their limiting diffusions. *Theory of Probbility and its Applications*, 13:128–139, 1968.
14. Z. Li. A limit theorem for discrete Galton-Watson branching processes with immigration. *Journal of Applied Probability*, 43:289–295, 2006.
15. Z. Li and C. Ma. Catalytic discrete state branching models and related limit theorems. *Journal of Theoretical Probability*, 21:936–965, 2008.
16. Z.-H. Li. Asymptotic behaviour of continuous time and state branching processes. *J. Austral. Math. Soc.*, (Series A) 68:68–84, 2000.
17. M. P. Quine. The multitype Galton-Watson process with $\rho$ near 1. *Advances in Applied Probability*, 4:pp.429–452, 1972.
18. D. Reinhold. *Asymptotic Behavior of Near Critical Branching Processes and Modeling of Cell Growth Data*. PhD thesis, University of North Carolina at Chapel Hill, NC, USA, 2011.

# Some Path Large-Deviation Results for a Branching Diffusion

**Robert Hardy and Simon C. Harris**

**Abstract** We give an intuitive proof of a path large-deviations result for a typed branching diffusion as found in Git, J.Harris and S.C.Harris (Ann. App. Probab. 17(2):609-653, 2007). Our approach involves an application of a change of measure technique involving a distinguished infinite line of descent, or *spine*, and we follow the spine set-up of Hardy and Harris (Séminaire de Probabilités XLII:281–330, 2009). Our proof combines simple martingale ideas with applications of Varadhan's lemma and is successful mainly because a "spine decomposition" effectively reduces otherwise difficult calculations on the whole collection of branching diffusion particles down to just a single particle (the spine) whose large-deviations behaviour is well known. A similar approach was used for branching Brownian motion in Hardy and Harris (Stoch. Process. Appl. 116(12):1992–2013, 2006). Importantly, our techniques should be applicable in a much wider class of branching diffusion large-deviations problems.

**Keywords** Branching diffusions • Spatial branching process • Path large deviations • Spine decomposition • Spine change of measure • Additive martingales

**AMS subject classification:** 60J80.

---

R. Hardy
Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK
Currently at VTB Capital, London
e-mail: Robert.Hardy@vtbcapital.com

S.C. Harris (✉)
Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK
e-mail: S.C.Harris@bath.ac.uk

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_5, 

# 1 Overview

Harris and Williams [7] introduced a model of a branching diffusion in which the diffusion and breeding rate of particles is controlled by their type process which moves as an Ornstein–Uhlenbeck process on $\mathbb{R}$, independently of the particle's position, associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{with } \theta > 0 \text{ considered as the } \textit{temperature}. \quad (1)$$

Throughout this chapter we shall refer to an OU process with generator $\frac{\theta}{2}\frac{\partial^2}{\partial y^2} - \mu y \frac{\partial}{\partial y}$ as an $\mathrm{OU}(\theta,\mu)$.

More precisely, the spatial movement of a particle of type $y$ is a driftless Brownian motion with instantaneous variance

$$A(y) := ay^2, \qquad \text{for some fixed } a \geq 0.$$

The breeding of a particle of type $y$ occurs at a rate

$$R(y) := ry^2 + \rho, \qquad \text{where } r, \rho > 0,$$

and binary splitting occurs at the fission times. The model has very different behaviour for low-temperature values (i.e., low $\theta$), but throughout we consider that $\theta > 8r$—the high-temperature regime.

We can suppose that the probabilities of this are $\{P^{x,y} : x,y \in \mathbb{R}\}$ so that $P^{x,y}$ is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of this branching diffusion process initiated from a single particle positioned at the space-type location $(x,y)$. The configuration of this branching diffusion at time $t$ is to be given by the $\mathbb{R}^2$-valued point process $\mathbb{X}_t := \{(X_u(t), Y_u(t)) : u \in N_t\}$ where $N_t$ is the set of individuals alive at time $t$, and without loss of generality, we can assume that the initial ancestor starts out at the space-type origin—henceforth, we use $P$ to mean $P^{0,0}$.

The main aim of this chapter is to prove upper and lower bounds for the probability of finding at least one of the branching particles very far from the space-type origin at large times, in a suitable large-deviations sense. The question of the lower bound was originally motivated by Git et al. [4] in order to determine the exponential growth rates and asymptotic shape of the branching diffusion. We briefly discuss this result in Sect. 2. In fact, our approach naturally gives rise to a stronger result where particles not only arrive at a very large space-type location $(\beta t, \kappa\sqrt{t})$ at a fixed time $\tau$ but are also known to have stayed "near" a specific space-type trajectory throughout the whole time interval $[0,\tau]$. The spine techniques used in this chapter involve a change of measure that makes a single "spine" particle "follow" a given trajectory. Our "spine" methods (for both bounds) should also prove useful to obtain large-deviations results in more general branching diffusions.

Although Git et al. [4] also used a spine change of measure, their original approach was far more model specific and quite different in flavour.

With $\theta > 8r$, let

$$\overline{\lambda} := \sqrt{\frac{\beta^2\theta(\theta-8r)}{a^2\kappa^4+4a\theta\beta^2}}, \qquad \overline{\mu} := \frac{\kappa^2\sqrt{\theta(\theta-8r)}}{2\sqrt{\kappa^4+4\theta\beta^2/a}} \tag{2}$$

and define a space-type trajectory $(x_s, y_s)_{s\in[0,\tau]}$ by

$$\overline{y}_s := \kappa\frac{\sinh\overline{\mu}s}{\sinh\overline{\mu}\tau}, \quad \overline{x}_s := a\overline{\lambda}\int_0^s y_w^2\,\mathrm{d}w, \qquad s\in[0,\tau]. \tag{3}$$

Note that the path endpoints are $y_\tau = \kappa$ and $x_\tau = \beta$. Also define,

$$\Theta(\beta,\kappa) := \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta-8r)(a^2\kappa^4+4\theta\beta^2)}}{4a\theta}. \tag{4}$$

**Theorem 1.1 (The Short-Climb Probability).** *Let $\beta < 0$, $\kappa\in\mathbb{R}$ and $\varepsilon > 0$.*
*(a) If $\tau > 0$ is sufficiently large, then for all $\delta, \delta' > 0$*

$$\liminf_{t\to\infty} t^{-1}\log P\Big(\exists u\in N_\tau : \forall s\in[0,\tau], \big|t^{-1}X_u(s) - \overline{x}_s\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(s) - \overline{y}_s\big| < \delta'\Big)$$
$$\geq -\Theta(\beta,\kappa) - \varepsilon.$$

*(b) If $\delta, \delta' > 0$ are sufficiently small, then for all $\tau > 0$*

$$\limsup_{t\to\infty} t^{-1}\log P\Big(\exists u\in N_\tau : \forall s\in[0,\tau], \big|t^{-1}X_u(s) - \overline{x}_s\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(s) - \overline{y}_s\big| < \delta'\Big)$$
$$\leq \limsup_{t\to\infty} t^{-1}\log P\Big(\exists u\in N_\tau : \big|t^{-1}X_u(\tau) - \beta\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(\tau) - \kappa\big| < \delta'\Big)$$
$$\leq -\Theta(\beta,\kappa) + \varepsilon.$$

Note that the above new upper bound result shows that the trajectory followed really is "optimal" in order to achieve the required large position at time $\tau$. We will actually prove a more general result for *any* (as opposed to sufficiently large) fixed time $\tau$ with a rate of decay $J(\tau)$, where $J(\tau)\downarrow\theta(\beta,\kappa)$. We state this stronger result as Theorem 3.2. We also note that the above lower bound on the probability of following the optimal paths to reach $(\beta t, \kappa\sqrt{t})$ is for a large *fixed* time $\tau$. In contrast, the method used for the analogous result in [4, Theorem 7] dictated that $\tau = \tau(t) \propto \log t$. Our stronger result would enable a corresponding simplification in the proof of the application in [4].

The principle behind the proof of the lower bound is to design new measures $\mathbb{Q}_t$ for the branching diffusion such that one of the particles (the spine) will closely follow a specific space-type path. Our spine approach, which we briefly lay out in Sect. 4 (and which is fully presented in Hardy and Harris [5]), will allow us to explicitly find the Radon–Nikodym derivatives (martingales) of these new measures with respect to the original measure $P$. Then, using the spine decomposition together with Doob's submartingale inequality, we shall show that the growth rate of these martingales under $\mathbb{Q}_t$ is exactly the correct rate for the large-deviations lower bound.

In such branching diffusion settings, we comment that a large-deviations upper bound is usually easier to obtain than the lower bound. Generally, we can overestimate the probability that any particle succeeds in performing a certain "rare" event by the expected number of particles performing that event, which then reduces to a single-particle (large deviation) calculation. In the present context, the upper bound of Theorem 1.1 can also be proved directly using some fundamental "additive" martingales.

The layout of this article is as follows: in the next section we discuss the results of Git et al. [4] in order to give a context to our work. Section 3 contains a heuristic discussion of the large deviations for the model which motivates the choice of the subsequent martingales. A statement of a stronger path large-deviations result (Theorem 3.2) is also found in this section. In Sect. 4, we briefly present the foundations of our spine approach, giving definitions of the underlying space and its filtrations, measures and the fundamental martingales of interest. These strictly positive martingales, $Z_t$, are defined in terms of specific paths as suggested by our heuristic arguments. As Radon–Nikodym derivatives, these martingales can define the new measures $\mathbb{Q}_t$ (under which they become *sub*martingales), and we state a key result (Theorem 4.3) on their growth under the measures $\mathbb{Q}_t$; this growth result leads directly to the proof of the large-deviations lower bounds which we present in Sect. 5. Section 6 contains proofs for the upper bounds of the two main large-deviations results. Section 7 is devoted to proving the martingale growth Theorem 4.3; the proof is not particularly short, but neither is it difficult given the spine technology. It should be noted that this result is the main application of spines in this chapter. We use the so-called *spine decomposition* to simplify a computation involving the branching particle martingale $Z_t$ into a computation involving a single "spine" particle that permits a standard application of Varadhan's lemma.

## 2 The Git et al.'s Almost-Sure Result

We first give some key parameter definitions for this model. Let

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}. \tag{5}$$

Let $\lambda \in \mathbb{R}$, with the following convention which we *always* use for $\lambda$:

$$\lambda_{\min} < \lambda < 0. \tag{6}$$

Also, define

$$\mu_\lambda := \frac{1}{2}\sqrt{\theta(\theta - 8r - 4a\lambda^2)} \qquad \psi_\lambda^{\pm} := \frac{1}{4} \pm \frac{\mu_\lambda}{2\theta} \qquad E_\lambda^{\pm} := \rho + \theta\psi_\lambda^{\pm}, \quad c_\lambda^{\pm} := -E_\lambda^{\pm}/\lambda \tag{7}$$

Note, $\lambda_{\min}$ is the point beyond which $\mu_\lambda$ is no longer a real number. The parameters $E_\lambda^{\pm} \in \mathbb{R}$ are in fact certain key eigenvalues, as will be described in the next section.

Before we move on to prove the above theorem, we summarise the main results from Git et al. [4] and the earlier Harris and Williams [7] so that the reader might understand how Theorem 1.1 fits into the picture.

Work on large-deviations results for this type of branching diffusion began in the paper by Harris and Williams [7], where they considered the behaviour in expectation of the counting function

$$N_t(\gamma) = \sum_{u \in N_t} \mathbf{1}_{(X_u(t) \le -\gamma t)}$$

(not to be confused with $N_t$, the set of individuals alive at time $t$) for each $\gamma \in \mathbb{R}$. In Git *et al.* [4] it was shown that

$$\lim_{t\to\infty} t^{-1} \log N_t(\gamma) = \Delta(\gamma)$$

exists almost surely and is *finite* for all $0 \le \gamma < \tilde{c}(\theta)$, for some constant $\tilde{c}(\theta)$; in the case that $\gamma \ge \tilde{c}(\theta)$ the limit is $-\infty$ since *no* particles will be as far out as the ray $-\gamma t$ at large times. In other words, this result says that we almost surely have exponential growth in numbers of particles following close to rays that are *not too steep*.

For later reference, the almost-sure growth rate is given explicitly by

$$\Delta(\gamma) = \inf_{\lambda \in (\lambda_{\min}, 0)} \{E_\lambda^- + \lambda\gamma\} = \rho + \frac{\theta}{4} - \frac{1}{4}\sqrt{\theta(\theta - 8r)(1 + 4\gamma^2/(\theta a))},$$

and it is also found that

$$\tilde{c}(\theta) := \sup\{\gamma : \Delta(\gamma) > 0\} = \sqrt{2a\Big(r + \rho + \frac{2(2r+\rho)^2}{\theta - 8r}\Big)}.$$

In fact, the work of Git et al. [4] improves this to obtain the almost-sure rate of growth in numbers of particles at certain *spatial and type* positions at large times. They study the following function that counts how many particles occupy a particular region in the type-space domain:

$$N_t(\gamma, \kappa) := \sum_{u \in N(t)} \mathbf{1}\{X_u(t) \le -\gamma t, Y_u(t)^2 \ge \kappa^2 t\}.$$

**Theorem 2.1.** *Under each $P^{x,y}$ law, the limit*

$$D(\gamma, \kappa) := \lim_{t\to\infty} t^{-1} \log N_t(\gamma, \kappa)$$

*exists almost surely and is given by*

$$D(\gamma,\kappa)=\begin{cases}\Delta(\gamma,\kappa) & \text{if } \Delta(\gamma,\kappa)>0,\\ -\infty & \text{otherwise.}\end{cases}$$

*Here,*

$$\begin{aligned}\Delta(\gamma,\kappa) &= \inf_{\lambda\in(\lambda_{min},0)}\left\{E_\lambda^- + \lambda\gamma - \kappa^2\psi_\lambda^+\right\},\\ &= \rho + \frac{\theta-\kappa^2}{4} - \frac{1}{4a\theta}\sqrt{\theta(\theta-8r)(4a\theta\gamma^2+a^2(\theta+\kappa^2)^2)}.\end{aligned} \tag{8}$$

## 2.1 The Almost-Sure Upper Bound

Harris and Williams [7] and Harris [8] showed that there are *two* strictly positive martingales $Z_\lambda^-$ and $Z_\lambda^+$ defined as

$$Z_\lambda^\pm(t) := \sum_{k=1}^{N(t)} v_\lambda^\pm(Y_k(t))e^{\lambda X_k(t) - E_\lambda^\pm t}, \tag{9}$$

where $v_\lambda^-$ and $v_\lambda^+$ are strictly positive eigenfunctions of the self-adjoint operator $\frac{1}{2}\lambda^2 A + R + Q_\theta$, with corresponding eigenvalues $E_\lambda^- < E_\lambda^+$. The explicit form for these eigenfunctions is

$$v_\lambda^\pm(y) = e^{\psi_\lambda^\pm y^2}$$

where $\psi_\lambda^\pm$ and $\mu_\lambda$ are given at (7) and $\psi_\lambda^\pm$ are both positive for all $\lambda\in(\lambda_{\min},0)$.

A useful trick to obtain upper bounds is to overestimate indicator functions by an exponential and optimise over parameters. It is often the case that this will bring in one of the martingales of the model: for $\lambda\in(\lambda_{\min},0)$,

$$\begin{aligned}\sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t)\le -\gamma t, Y_k(t)^2\ge \kappa^2 t\right\} &\le \sum_{k=1}^{N(t)} \exp\left\{\psi_\lambda^+(Y_k(t)^2-\kappa^2 t)\right\}\exp\left\{\lambda(X_k(t)+\gamma t)\right\}\\ &= e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{(E_\lambda^- + \lambda\gamma - \kappa^2\psi_\lambda^+)t}.\end{aligned} \tag{10}$$

(Importantly for this, the parameter $\psi_\lambda^+$ is positive and $\lambda$ is negative; the functions $c_\lambda^-$ and $c_\lambda^+$ are defined as $c_\lambda^\pm := E_\lambda^\pm/(-\lambda)$.)

The expression for $\Delta(\gamma,\kappa)$ as a Legendre conjugate—see (8)—explains why $\Delta(\gamma,\kappa)$ relates to (10) above: by choosing $\lambda$ at the infimum we get

$$N_t(\gamma,\kappa) \le e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{\Delta(\gamma,\kappa)t}. \tag{11}$$

We remember that $N_t(\gamma,\kappa)$ takes only integer values, and a separate theorem by Harris and Git states that

$$\limsup_{t\to\infty} e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) \le 0, \qquad \text{for each } \lambda \in (\lambda_{\min}, 0). \tag{12}$$

Thus if $\Delta(\gamma,\kappa) < 0$ we deduce that almost surely

$$N_t(\gamma,\kappa) = 0, \quad \text{eventually,}$$

whence $\lim_{t\to\infty} t^{-1}\log N_t(\gamma,\kappa) = -\infty$, as required.

On the other hand, if $\Delta(\gamma,\kappa) \ge 0$, (11) and (12) immediately imply that

$$\lim_{t\to\infty} t^{-1}\log N_t(\gamma,\kappa) \le \Delta(\gamma,\kappa).$$

## 2.2 A Two-Phase Mechanism for the Lower Bound

For their proof of the almost-sure lower bound of Theorem 2.1, Git et al. [4] propose an explicit mechanism by which a sufficient number of particles will obtain a position near $(\gamma T, \kappa\sqrt{T})$ in the type-space domain at large times $T$. It is made up of two phases:

*The long tread:* Over a long period $[0,t]$, taking up nearly all of the time, a large number of particles will drift spatially with speed $\gamma\theta/(\theta+\kappa^2)$—as if their type has had a modified occupation measure, as described by Harris and Williams [7];
*The short climb:* Following this, over a short period of time $[t, t+\tau]$ with $\tau$ a fixed time ($\tau \ll t$), each of the particles from this group will have a small probability of further rushing to the large type position $\kappa\sqrt{t}$ whilst additionally gaining $\left\{\gamma\kappa^2/(\theta+\kappa^2)\right\}t$ in spatial position.The combination of these two phases will present us with *sufficiently many* particles at the space-type position $\left(\gamma T, \kappa\sqrt{T}\right)$ at the large time $T = t+\tau$, as Git et al. [4] show in their proof of the lower bound of Theorem 2.1 – we refer the reader to their work for further details. This lower bound requires a substantial amount of technical work, mainly focussed on the short climb in which, in particular, they required $\tau = O(\log t)$. Our Theorem 1.1 includes an alternative *short-climb* lower bound for $\tau$ fixed, and using this would slightly simplify the combination of phases. Our current proof will also provide a cleaner, more intuitive and more generic approach to such path large-deviations results in branching diffusions.

# 3 Large Deviations Heuristics

We now present some heuristic arguments concerning the large-deviations behaviour of the branching diffusion which will serve as the *intuition* behind our later *rigourous* proofs.

Under a measure $\tilde{P}$ let $(\xi_s, \eta_s)$ satisfy

$$\mathrm{d}\eta_s = \sqrt{\theta}\mathrm{d}B_s - \frac{\theta}{2}\eta_s \mathrm{d}s, \quad \text{and} \quad \mathrm{d}\xi_s = \sqrt{a}\eta_s\,\mathrm{d}W_s,$$

for two independent $\tilde{P}$-Brownian motions $B_s$ and $W_s$. Under $\tilde{P}$, $(\xi_s, \eta_s)$ moves like a single particle within the branching diffusion. For a large-deviations analysis we observe that for any $t > 0$,

$$\mathrm{d}\left(\frac{\eta_s}{\sqrt{t}}\right) = \sqrt{\theta}\left(\frac{\mathrm{d}B_s}{\sqrt{t}}\right) - \frac{\theta}{2}\left(\frac{\eta_s}{\sqrt{t}}\right)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}\left(\frac{\xi_s}{t}\right) = \sqrt{a}\left(\frac{\eta_s}{\sqrt{t}}\right)\left(\frac{\mathrm{d}W_s}{\sqrt{t}}\right),$$

and it will be natural to work with the *rescaled* processes $(\xi_s/t, \eta_s/\sqrt{t})$ since in this way we obtain a variance coefficient of $1/\sqrt{t}$ on the driving Brownian motions.

**Definition 3.1.** For each $t > 0$ we define the *rescaled* single-particle motion $(\xi_s^t, \eta_s^t)$ by

$$\xi_s^t := \xi_s/t, \quad \text{and} \quad \eta_s^t := \eta_s/\sqrt{t},$$

We note that under $\tilde{P}$ we have for $s \in [0, \tau]$:

$$\mathrm{d}\eta_s^t = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}B_s - \frac{\theta}{2}\eta_s^t \mathrm{d}s, \quad \text{and} \quad \mathrm{d}\xi_s^t = \frac{\sqrt{a}\eta_s^t}{\sqrt{t}}\mathrm{d}W_s,$$

for two independent $\tilde{P}$-Brownian motions $B_s$ and $W_s$.

Throughout the remainder of this chapter and different from the earlier parts, the variable $t$ will not be a time parameter but will bring about this large-deviations scaling; typically we shall use either $w$ or $s$ to denote the time parameter from the time interval $[0, \tau]$ where $\tau > 0$ is considered as fixed.

Suppose that we are given two paths: a type path $y : [0, \tau] \to \mathbb{R}$ and a spatial path $x : [0, \tau] \to \mathbb{R}$. On a *heuristic* level we can say that the probability of the type diffusion $\eta_s^t$ closely following $y$ and the space diffusion $\xi_s^t$ closely following $x$ is roughly

$$\exp\left(-\frac{t}{2\theta}\int_0^\tau \left(\dot{y}_s + \frac{\theta}{2}y_s\right)^2 \mathrm{d}s - \frac{t}{2}\int_0^\tau \frac{\dot{x}_s^2}{a y_s^2}\mathrm{d}s\right), \tag{13}$$

for large enough $t$. See [2], for example, for the large-deviations theory of Wentzell–Freidlin.

The reader who is familiar with the large-deviations principle for branching Brownian motion (see Hardy and Harris [6] for a spine proof or Lee [14] for a classical proof) might guess that the *probability at least one* of the *rescaled* branching particles $\left(X_u(s)/t, Y_u(s)/\sqrt{t}\right)$ follows the *difficult* space-type path $(x_s, y_s)$ 'closely' over the time interval $[0, \tau]$ is roughly

$$\exp\left\{-\sup_{w\in[0,\tau]}\left[\left(\int_0^w \frac{1}{2\theta}\left(\dot{y}_s + \frac{\theta}{2}y_s\right)^2 + \frac{1}{2}\frac{\dot{x}_s^2}{a y_s^2} - r y_s^2\,\mathrm{d}s\right)t - \rho w\right]\right\},$$

when $t$ is large. This *guess* can be obtained by upper estimating the *probability* by the *expected number of particles* following the path, then using the "one-particle picture" of Sect. 4 and large-deviations theory for one particle before optimising by choosing the most difficult time along the path.

By standard optimisation arguments (Git et al. [4] give some details of how this can be carried out) this implies that the probability of at least one of the rescaled branching particles being near the space-type position $(\beta,\kappa)$ at a fixed time $\tau$ (which is also the event that the non-rescaled particles arrive near $(\beta t,\kappa\sqrt{t})$ of course) should be roughly

$$\exp\left\{-\inf_{x,y}\sup_{w\in[0,\tau]}\left[\left(\int_0^w \frac{1}{2\theta}\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2+\frac{1}{2}\frac{\dot{x}_s^2}{ay_s^2}-ry_s^2\,\mathrm{d}s\right)t-\rho w\right]\right\}, \tag{14}$$

when $t$ is large and where the infimum is taken over all paths $x,y\in C[0,\tau]$ satisfying

$$y(0)=0, y(\tau)=\kappa, x(0)=0, x(\tau)=\beta. \tag{15}$$

This is typical in a large-deviations setting: although there are many possible trajectories that the (rescaled) particles could travel along to get to a position $(\beta,\kappa)$, the *dominant number* will have followed *optimal* paths.

Git et al. [4] state that for any given type path $y$, the optimal space path $x$ for (14) under the constraint $x(\tau)=\beta$ will always be given by

$$x_s=\lambda\int_0^s ay_w^2\,\mathrm{d}w, \qquad \text{for } s\in[0,\tau], \tag{16}$$

for some value $\lambda\in\mathbb{R}$. Briefly, their arguments rely on the fact that in the definition of our model the spatial diffusion $X_u(s)$ of the branching particles can be seen as a time-changed Brownian motion where the time scaling is determined by its type process $Y_u(s)$:

$$X_u(s)=\hat{B}\left(\int_0^s aY_u(w)^2\,\mathrm{d}w\right)$$

for a Brownian motion $\hat{B}(\cdot)$ on $[0,\tau]$. A measure change that introduces a linear drift of $\lambda$ to this Brownian motion will give

$$X_u(s)=\tilde{B}\left(\int_0^s aY_u(w)^2\,\mathrm{d}w\right)+\lambda\int_0^s aY_u(w)^2\,\mathrm{d}w,$$

where $\tilde{B}(\cdot)$ is a Brownian motion under the new measure—this clearly relates to (16). Linear drifts are the optimal path (in a large-deviations sense) for a Brownian motion to be at a given point at a given time, and the constraint $x(\tau)=\beta$ for our problem will determine the value of $\lambda$ in terms of the type path $y$:

$$\lambda = \frac{\beta}{a \int_0^\tau y_s^2 \, \mathrm{d}s}. \tag{17}$$

Thus for the event being considered in Theorem 1.1, the optimal spatial path $x$ is determined *uniquely* by (16) together with (17). Therefore, an equivalent but easier statement of our large-deviations result is that the probability of at least one of the rescaled branching particles being near the space-type position $(\beta, \kappa)$ at a fixed time $\tau$ is roughly

$$\exp\left\{ - \inf_{y} \sup_{w \in [0,\tau]} \left[ \left( \int_0^w \frac{1}{2\theta} \left( \dot{y}_s + \frac{\theta}{2} y_s \right)^2 + \frac{a\lambda^2}{2} y_s^2 - r y_s^2 \, \mathrm{d}s \right) t - \rho w \right] \right\}, \tag{18}$$

when $t$ is large and where the infimum is taken over all paths $y \in C[0,\tau]$ and all $\lambda \in (\lambda_{\min}, 0)$ satisfying

$$y(0) = 0, y(\tau) = \kappa, \quad \lambda = \frac{\beta}{a \int_0^\tau y_s^2 \, \mathrm{d}s}. \tag{19}$$

Git et al. [4] presented alternative heuristic arguments based on birth-death processes to arrive at the expression (18). Using Euler–Lagrange techniques, they showed that the specific path

$$y_s = \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0, \tau] \tag{20}$$

is optimal for this expression, where

$$\mu_\lambda = \frac{\sqrt{\theta(\theta - 8r - 4a\lambda^2)}}{2}, \tag{21}$$

and $\lambda \in (\lambda_{\min}, 0)$ is dependent on the choice of $\tau$ (which we are anyway considering as fixed throughout) and is chosen to satisfy

$$\frac{\beta}{a\lambda} = \kappa^2 \left( \frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2 \sinh^2 \mu_\lambda \tau} \right). \tag{22}$$

We refer the reader to Git et al. [4] for details of these relationships between the parameters, but note that particles staying close to this path will arrive near $y(\tau) = \kappa$ at time $\tau$ in agreement with the heuristics.

As we mentioned just before the statement of Theorem 1.1, our spine techniques will naturally use the path $y_s$ defined at (20) together with $x_s$ defined at (16), since they are the optimal paths (in a large-deviations sense) for accumulating particles near the point $(\beta t, \kappa\sqrt{t})$ at time $\tau$. In fact, our spine proof of Theorem 1.1 will result in a proof of the following stronger result, from which Theorem 1.1(a) would follow as a corollary.

**Theorem 3.2.** *Let $\tau > 0$ be fixed and suppose $\beta < 0$ and $\kappa \in \mathbb{R}$. Define two paths on $[0,\tau]$ by*

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad x_s := a\lambda \int_0^s y_w^2 \, \mathrm{d}w, \quad s \in [0,\tau], \tag{23}$$

*where $\lambda \in (\lambda_{\min}, 0)$ is chosen so that*

$$\beta = a\lambda \int_0^\tau y_w^2 \, \mathrm{d}w = a\lambda \kappa^2 \left( \frac{\coth \mu_\lambda \tau}{2\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau} \right). \tag{24}$$

*Note that the path endpoints are $y_\tau = \kappa$ and $x_\tau = \beta$. Define*

$$J(\tau) := \int_0^\tau \left[ \frac{1}{2\theta} (\dot{y}_s + \frac{\theta}{2} y_s)^2 + \frac{a\lambda^2}{2} y_s^2 - r y_s^2 \right] \mathrm{d}s = \lambda\beta + \kappa^2 \left( \frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \right). \tag{25}$$

*(a) For all $\delta, \delta' > 0$,*

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|X_u(s) - t x_s\big| < \delta t, \big|Y_u(s) - \sqrt{t} y_s\big| < \delta' \sqrt{t}\Big) \geq -J(\tau),$$

*(b) Let $\varepsilon > 0$. For all sufficiently small $\delta, \delta' > 0$,*

$$\limsup_{t\to\infty} t^{-1} \log P\left(\exists u \in N_\tau : \forall s \in [0,\tau], \big|X_u(s) - t x_s\big| < \delta t, \big|Y_u(s) - \sqrt{t} y_s\big| < \delta' \sqrt{t}\right) \leq -J(\tau) + \varepsilon.$$

It can be easily checked that $J(\tau) \downarrow \Theta(\beta,\kappa)$, $\lambda \to \overline{\lambda}$ and $(x_s, y_s) \to (\overline{x}_s, \overline{y}_s)$ as $\tau \to \infty$. Theorem 1.1 can now easily be deduced from Theorem 3.2 by choosing $\tau$ sufficiently large.

Although some additional work would be required to prove as much, the paths $(x_s, y_s)$ above are chosen as they are the "best" ones for particles to follow in order to reach position $(\beta t, \kappa\sqrt{t})$ at (fixed) time $\tau$, as found in the large-deviations heuristics discussed above and in [4].

Importantly, we emphasise that it should be possible to develop the ideas and techniques used in this chapter to obtain proofs of large-deviations principles for many other branching diffusion models, essentially because we can reduce the branching particles down to the spine and in general this gives a technique for deriving large-deviations principles for the branching diffusion from those of the single diffusing particle (the spine) which are already well studied.

## 4 The Spine Approach, Martingales and Measures

In this section we will introduce some of the key concepts used in the proofs of the main results. We shall construct the branching diffusion with a distinguished infinite line of decent, the *spine*, and then perform a change of measure that will make the spine "closely" follow a given path. Estimates on the martingale associated with this change of measure can then give a lower bound for large-deviations events in the branching diffusion. The heuristics of the previous section will serve as an important guide. However, although they have already indicated a specific path at (20), it should be noted that in our proofs we use properties of this path only at a few points—elsewhere the techniques can be applied in general to any path. Therefore the reader may suppose that $y : [0,\tau] \to \mathbb{R}$ is any given and fixed path, and we shall be very careful to highlight those points when we use specific properties of the path defined at (20). Also, to keep notational complexity to a reasonable minimum, we tend not to make the dependencies of the martingales and action functionals on the underlying chosen paths explicit in the notation.

**The Spine Set-Up.** Recall that the original branching diffusion $\mathbb{X}_t := \{(X_u(t), Y_u(t)) : u \in N_t\}$ where $N_t$ is the set of individuals alive at time $t$ has associated probability measures $P^{x,y}$ with natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$. We label all particles according to the Ulam–Harris convention. For example, "$\emptyset$213" represents "the 3rd child of the 1st child of the 2nd child of the initial ancestor." For two labels $v, u \in \Omega$ the notation $v < u$ means that $v$ is an *ancestor* of $u$, $|u|$ is the generation of particle $u$, and so forth.

A *spine* $\xi$ is a *distinguished* infinite line of descent starting with the initial ancestor, where $\xi = \{\xi_0, \xi_1, \xi_2, \ldots\}$ with $\xi_0 = \emptyset$, $\xi_n$ the label of the spine at the $n$-th generation and $u \in \xi$ means that $u = \xi_i$ for some $i \geq 0$. Let $\{(\xi(t), \eta(t))\}_{t\geq 0}$ represent the space-type path of spine, that is, the time position of the spine at time $t$ is $(\xi(t), \eta(t)) := (X_u(t), Y_u(t))$ for $u \in N(t) \cap \xi$. Define $n = \{n_t : t \geq 0\}$ to be the counting process for the number of fissions that have occurred along the path of the spine by time $t$, with the actual fission times along the spine denoted by $\{S_i\}_{\geq 1}$. Note that $\xi_{n_t}$ is the label of the spine at time $t$.

We will make important use of a variety of filtrations for the process with a distinguished spine. Let the *enriched* filtration for the branching diffusion with distinguished spine be $\tilde{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \{\xi_{n_s}\}_{s\leq t})$. Then $\tilde{\mathcal{F}}_t$ knows everything up to time $t$, all particle paths, genealogy and identification of spine, whereas $\mathcal{F}_t$ knows about the paths and genealogy of all particles up to the time $t$, but does not know the identity of the spine. In addition, let $\tilde{\mathcal{G}}_t := \sigma(\{(\xi(s), \eta(s), n_s, \xi_{n_s})\}_{s\leq t})$ and $\mathcal{G}_t := \sigma(\{(\xi(s), \eta(s))\}_{s\leq t})$. Then $\tilde{\mathcal{G}}_t$ knows everything *along* the spine up to time $t$ - the *spine's* motion, the *spine's* genealogy and the *spine's* fission times. It does not know about any information "off" the spine. On the other hand, $\tilde{\mathcal{G}}_t$ only knows about the *spine's* motion but not about the births along the spine, the spine's genealogical information, nor any information "off" the spine.

*The Spine Construction.* Under a measure $\tilde{P}^{x,y}$, the branching diffusion $(\mathbb{X}_s)_{s\geq 0}$ with distinguished spine $\xi$ is constructed as follows:

- The spine process $(\xi_s, \eta_s)$ starts at $(x, y)$ and diffuses as a solution to

$$\mathrm{d}\eta_s = \sqrt{\theta}\mathrm{d}B_s - \frac{\theta}{2}\eta_s \mathrm{d}s, \quad \text{and} \quad \mathrm{d}\xi_s = \sqrt{a}\eta_s \mathrm{d}W_s, \tag{26}$$

where $B_s$ and $W_s$ are standard Brownian motions.

- At rate $R(\eta_s)$ the spine undergoes fission producing two particles.
- With equal probability, one of these two particles is selected to continue the spine $\xi$.
- The other particle initiates, from its birth space-type position, an independent copy of the original $P$ branching diffusion with branching rate $R(\cdot)$.

We note that $\tilde{P}^{x,y}$ is an extension of the original measure $P^{x,y}$, with $P = \tilde{P}|_{\mathcal{F}_\infty}$. In fact, it is easy to see that an alternative way to construct the process under $\tilde{P}$ is to first construct the entire tree of the branching diffusion according to $P$ and, secondly, choose the spine by starting from the initial ancestor then following the spine path forward in time with independent uniform choices made from the particles produced at each fission. In particular, for $u \in N(t)$, we note that $\tilde{P}(u \in \xi | \mathcal{F}_t) = \prod_{v<u} 2^{-1} = 2^{-|u|}$.

We also have the very useful and intuitive *'one-particle picture' (OPP)*. For example, for any measurable function $f$ of single-particle paths on $[0,t]$,

$$\tilde{P}^{x,y}\left(\sum_{u\in N(t)} f(X_u(s), Y_u(s); s \le t)\right) = \tilde{P}^{x,y}\left(e^{\int_0^t \beta(\eta(s))\,ds} f(\xi(s), \eta(s); s \le t)\right)$$

For further details of this spine set-up and various key results, see Hardy and Harris [5]. Also see Lyons et al. [11, 15, 16] and other recent work based on these (examples are Kyprianou [12], Kyprianou and Sani [13], Athreya [1], Olofsson [18], amongst others) for similar spine-based approaches in branching processes.

**Changes of Measure.** We can now able to perform our "spine change of measure." For any $t > 0$ and any given $y : [0,\tau] \to \mathbb{R}$ that is square integrable along with its derivative

$$\exp\left(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w \left(\dot{y}_s + \frac{\theta}{2} y_s\right) \mathrm{d}B_s - \frac{t}{2\theta}\int_0^w \left(\dot{y}_s + \frac{\theta}{2} y_s\right)^2 \mathrm{d}s\right) \times \exp\left(\sqrt{at}\lambda \int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\right),$$

is a strictly-positive $\tilde{P}$-martingale over the time period $w \in [0,\tau]$ (see, e.g., Øksendal [17]). As one part of the change of measure defined below, this martingale will introduce drift terms into the diffusions $\eta_s$ and $\xi_s$ such that $\eta_s^t \sim y_s$ and $\xi_s^t \sim a\lambda y_s^2$ when $t$ is large, and we note a comparison between this martingale and the expression (18) above.

The process $n_w$ which counts the number of fission times on the spine up to time $w$ is a Cox process of rate $R(\eta_s)$ and therefore for $w \in [0,\tau]$,

$$w \mapsto e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w}$$

is also a $\tilde{P}$-martingale. We can use the product of these two martingales to define a new measure:

**Theorem 4.1 (Spine Change of Measure).** *Let $\tau > 0$ be fixed. For $t > 0$, we define a measure $\tilde{\mathbb{Q}}_t$ on $\tilde{\mathcal{F}}_\tau$ where*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_t}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_w} := \tilde{\zeta}_t(w) := \exp\left(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w\left(\dot{y}_s+\frac{\theta}{2}y_s\right)\mathrm{d}B_s - \frac{t}{2\theta}\int_0^w\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right)$$

$$\times\exp\left(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\right)\times e^{-\int_0^w R(\eta_s)\,\mathrm{d}s}2^{n_w}, \qquad (27)$$

*for* $w\in[0,\tau]$. *Under the measure* $\tilde{\mathbb{Q}}_t^{x,y}$ *we can give a pathwise construction of the branching diffusion* $(\mathbb{X}_s)_{s\in[0,\tau]}$ *with distinguished spine* $\xi$*:*

- *The spine process* $(\xi_s,\eta_s)$ *starts at* $(x,y)$ *and diffuses as a solution to*

$$\mathrm{d}(\eta_s-\sqrt{t}y_s) = \sqrt{\theta}\mathrm{d}\tilde{B}_s - \frac{\theta}{2}(\eta_s-\sqrt{t}y_s)\mathrm{d}s \qquad (28)$$

  *and*

$$\mathrm{d}\xi_s = \sqrt{a}\eta_s\,\mathrm{d}\tilde{W}_s + a\lambda\sqrt{t}y_s\eta_s\,\mathrm{d}s, \qquad (29)$$

  *where* $\tilde{B}$ *and* $\tilde{W}$ *are standard Brownian motions under* $\tilde{\mathbb{Q}}_t$*, with*

$$\mathrm{d}\tilde{B}_s = \mathrm{d}B_s - \frac{\sqrt{t}}{\sqrt{\theta}}\left(\dot{y}_s+\frac{\theta}{2}y_s\right)\mathrm{d}s, \qquad \mathrm{d}\tilde{W}_s = \mathrm{d}W_s - \sqrt{at}\lambda y_s\,\mathrm{d}s.$$

- *At the accelerated rate* $2R(\eta_s)$*, the spine undergoes fission producing two particles.*
- *With equal probability, one of these two particles is selected to continue the spine.*
- *The other particle initiates, from its birth space-type position, an independent copy of the original* $P$ *branching diffusion with normal branching rate* $R(\cdot)$*.*

Due to our formulation of the underlying spine foundations in terms of filtrations and sub-filtrations, we can project this new measure $\tilde{\mathbb{Q}}_t$ down onto the branching diffusion particles and define a measure $\mathbb{Q}_t$ on $\mathcal{F}_\tau$ by $\mathbb{Q}_t := \tilde{\mathbb{Q}}_t|_{\mathcal{F}_\tau}$.

**Theorem 4.2.** *Let* $\tau>0$ *be fixed. For each fixed* $t>0$*, define the* $\big((\mathcal{F}_w)_{0\le w\le\tau},P\big)$*-martingale* $Z_t(w)$ *for* $w\in[0,\tau]$ *by*

$$Z_t(w) := \left.\frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}P}\right|_{\mathcal{F}_w} = \tilde{P}\big(\tilde{\zeta}_t(w)|\mathcal{F}_w\big).$$

*Then*

$$Z_t(w) = \sum_{u\in N(w)} f_{t,w}(u)$$

*where*

$$f_{t,w}(u) := e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}\exp\left(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w\left(\dot{y}_s+\frac{\theta}{2}y_s\right)\mathrm{d}B_u(s) - \frac{t}{2\theta}\int_0^w\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right)$$

$$\times \exp\left(\sqrt{at}\lambda \int_0^w y_s \, \mathrm{d}W_u(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2 \, \mathrm{d}s\right) \tag{30}$$

*and $B_u$ and $W_u$ to denote the P-Brownian motions driving the type and spatial processes of u.*

See Hardy and Harris [5], or Engländer and Kyprianou [3], for more details and proofs.

**The Spine Decomposition.** Consider those particles alive at time $\tau$ and group them together according to the time that they first branched off the spine's path. Since, under $\tilde{\mathbb{Q}}_t$ particles "off" the spine behave as if under $P$ and since $Z_t$ is a $P$-martingale, it is easy to see the following "spine decomposition":

$$\tilde{\mathbb{Q}}_t(Z_t(\tau)|\tilde{\mathcal{G}}_\infty) = f_{t,\tau}(\xi) + \sum_{i=1}^{n_t} f_{t,S_i}(\xi)$$

where

$$f_{t,w}(\xi) := e^{-\int_0^w R(\eta(s))\,\mathrm{d}s} \exp\left(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w \left(\dot{y}_s + \frac{\theta}{2}y_s\right)\mathrm{d}B(s) - \frac{t}{2\theta}\int_0^w \left(\dot{y}_s + \frac{\theta}{2}y_s\right)^2 \mathrm{d}s\right)$$

$$\times \exp\left(\sqrt{at}\lambda \int_0^w y_s \, \mathrm{d}W(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2 \, \mathrm{d}s\right) \tag{31}$$

The spine decomposition will prove essential for our key martingale growth estimates.

**The Growth of Martingale $Z_t$.** For our proof of Theorem 1.1 (and its stronger version of Theorem 3.2) it is important to know how quickly $Z_t(\tau)$ grows under the measure $\mathbb{Q}_t$. The following key result is the main application of spines in this article:

**Theorem 4.3.** *For the specific path y defined at (20) and for any $\alpha \in [0,1]$ we have*

$$\limsup_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t\left(Z_t(\tau)^\alpha\right) \le \alpha J(\tau) + \alpha^2 M(\tau),$$

*where we define*

$$J(w) := \int_0^w \left[\frac{1}{2\theta}\left(\dot{y}_s + \frac{\theta}{2}y_s\right)^2 + \frac{a\lambda^2}{2}y_s^2 - r y_s^2\right] \mathrm{d}s, \tag{32}$$

*and*

$$M(w) := \int_0^w \left[\frac{1}{2\theta}\left(\dot{y}_s + \frac{\theta}{2}y_s\right)^2 + \frac{a\lambda^2}{2}y_s^2\right] \mathrm{d}s. \tag{33}$$

We emphasise that without the technology of spines, the proof of this result would be far more difficult. The spine decomposition gives us a methodology for reducing the additive structure of these martingales essentially to a single-particle problem, and since it does this through a conditional-expectation operation rather than with an inequality, it is *exact* and therefore can lead to tight estimates that are useful. Due to its length, we dedicate the whole of Sect. 7 to the spine proof of this above theorem and now proceed to show how this result can be used to obtain the upper bound on $Z_t(\tau)$ that we require for Theorem 1.1.

It is not difficult to verify that for any $\alpha \in [0,1]$, $Z_t(w)^\alpha$ is a submartingale with respect to the measure $\mathbb{Q}_t$. Given Theorem 4.3, we can therefore use Doob's submartingale inequality to prove the following:

**Theorem 4.4.** *Let $\tau > 0$ be fixed. Then for all $\varepsilon > 0$,*

$$\lim_{t\to\infty} \mathbb{Q}_t \left( \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t} \right) \to 1.$$

*Proof.* For a given $\varepsilon > 0$ and for any $\alpha \in [0,1]$, Doob's inequality gives

$$\mathbb{Q}_t\Big( \sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t} \Big) = \mathbb{Q}_t\Big( \sup_{s\in[0,\tau]} Z_t(s)^\alpha > e^{\alpha(J(\tau)+\varepsilon)t} \Big) \leq \frac{\mathbb{Q}_t\big(Z_t(\tau)^\alpha\big)}{e^{\alpha(J(\tau)+\varepsilon)t}}.$$

From Theorem 4.3, we know that for each $\alpha \in [0,1]$ and for all large $t$, we have

$$\mathbb{Q}_t\Big( \sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t} \Big) \leq e^{(\alpha M(\tau)-\varepsilon)\alpha t}.$$

If we also have $\alpha \in (0, \varepsilon/M(\tau))$ then clearly this above is a decaying exponential and so it follows that

$$\lim_{t\to\infty} \mathbb{Q}_t\Big( \sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t} \Big) \to 0.$$

□

For the specific $y$ defined at (20), it can be shown that

$$J(\tau) = \int_0^\tau \left[ \frac{1}{2\theta}\left( \dot{y}_s + \frac{\theta}{2} y_s \right)^2 + \frac{a\lambda^2}{2} y_s^2 - r y_s^2 \right] \mathrm{d}s = \lambda\beta + \kappa^2 \left( \frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \right),$$

where we recall that this $\lambda \in (\lambda_{\min}, 0)$ was specifically determined by (22). In fact, Git et al. [4] explain that this choice of $\lambda$ was optimal in that

$$\lambda\beta + \kappa^2 \left( \frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \right) = \sup_\gamma \left\{ \gamma\beta + \kappa^2 \left( \frac{1}{4} + \frac{\mu_\gamma}{2\theta} \coth \mu_\gamma \tau \right) \right\}. \tag{34}$$

On the other hand we can find a similar representation for the parameter $\Theta(\beta,\kappa)$: if we define

$$\bar{\lambda} := \sqrt{\frac{\beta^2\theta(\theta-8r)}{a^2\kappa^4 + 4a\theta\beta^2}}, \quad \text{so that } \mu_{\bar{\lambda}} = \frac{\kappa^2\sqrt{\theta(\theta-8r)}}{2\sqrt{\kappa^4 + 4\theta\beta^2/a}},$$

then

$$\Theta(\beta,\kappa)=\bar{\lambda}\beta+\kappa^2\psi_{\bar{\lambda}}^{+}=\sup_{\gamma}\left\{\gamma\beta+\kappa^2\left(\frac{1}{4}+\frac{\mu_\gamma}{2\theta}\right)\right\}=\lim_{\tau\to\infty}\sup_{\gamma}\left\{\gamma\beta+\kappa^2\left(\frac{1}{4}+\frac{\mu_\gamma}{2\theta}\coth\mu_\gamma\tau\right)\right\},$$

where we recall that

$$\psi_{\bar{\lambda}}^{+}:=\frac{1}{4}+\frac{\mu_{\bar{\lambda}}}{2\theta}.$$

In this way it can be deduced from (34) that $J(\tau)>\Theta(\beta,\kappa)$ with

$$J(\tau)\downarrow\Theta(\beta,\kappa),\quad \text{as } \tau\to\infty.$$

It is now easy to deduce the following corollary to Theorem 4.4:

**Corollary 4.5.** *Given* $\varepsilon>0$, *for* $\tau>0$ *chosen sufficiently large,*

$$\lim_{t\to\infty}\mathbb{Q}_t\left(\sup_{s\in[0,\tau]}Z_t(s)\le e^{(\Theta(\beta,\kappa)+\varepsilon)t}\right)\to 1.$$

## 5 Proving the Large-Deviations Lower Bound

Barring the proof of Theorem 4.3 which we cover fully in Sect. 7, we now have all the ingredients required to prove the large-deviations lower bound for the short-climb event of Theorem 3.2.

Throughout this proof we are focussing on the specific path

$$y_s:=\kappa\frac{\sinh\mu_\lambda s}{\sinh\mu_\lambda\tau},\quad s\in[0,\tau]$$

where $\lambda\in(\lambda_{\min},0)$ satisfies

$$\frac{\beta}{a\lambda}=\kappa^2\left(\frac{\coth\mu_\lambda\tau}{\mu_\lambda}-\frac{\tau}{2\sinh^2\mu_\lambda\tau}\right).$$

as discussed at (22). We define the event that the space-type location $\big(X_u(s),Y_u(s)\big)$ of a particular particle $u\in N_\tau$ remains near $(a\lambda t\int_0^s y_w^2\,\mathrm{d}w,\sqrt{t}y_s)$ throughout the interval $s\in[0,\tau]$:

$$A_t(u):=\left\{\forall s\in[0,\tau],\left|X_u(s)-a\lambda t\int_0^s y_w^2\,\mathrm{d}w\right|<\delta t,\left|Y_u(s)-\sqrt{t}y_s\right|<\delta'\sqrt{t}\right\},$$

where $\delta,\delta'>0$ are given and fixed. In addition, we define the event that any of the particles performs this event (whilst emphasising the parameter dependence) by

$$A_{t,\tau}^{\delta,\delta'} := \bigcup_{u\in N(\tau)} A_t(u)$$

Noting that this event is $\mathcal{F}_\tau$-measurable since it depends only on the branching particles and does not refer to the spine, it follows that on this event the change of measure is carried out by $Z_t$, as noted in Theorem 4.2. The upper bound that we have derived for $Z_t$ at Corollary 4.5 will serve as a lower bound for $1/Z_t(\tau)$ in this change of measure and will combine with the fact that under the measure $\tilde{\mathbb{Q}}_t$ (for large $t$) we know that the spine will carry out the large-deviations behaviour that we want.

Then for any $\varepsilon > 0$,

$$\begin{aligned} P\Big(A_{t,\tau}^{\delta,\delta'}\Big) &= \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u)\Big) \\ &\geq \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u);\ \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big) \\ &\geq e^{-(J(\tau)+\varepsilon)t}\,\mathbb{Q}_t\Big(\exists u \in N_\tau, A_t(u);\ \sup_{s\in[0,\tau]} Z_t(s) \leq e^{J(\tau)+\varepsilon)t}\Big) \\ &\geq e^{-(J(\tau)+\varepsilon)t}\,\tilde{\mathbb{Q}}_t\Big(A_t(\xi);\ \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big). \end{aligned} \tag{35}$$

Given (28) and (29), standard theory says that under the measure $\tilde{\mathbb{Q}}_t$ (with $t$ large) the rescaled spine $(\xi_s^t, \eta_s^t)$ will tend to stay close to the space-type paths $(a\lambda \int_0^s y_w^2\,\mathrm{d}w, y_s)$ over the whole time interval $[0,\tau]$:

$$\xi_s^t \sim a\lambda \int_0^s y_w^2\,\mathrm{d}w, \quad \text{and} \quad \eta_s^t \sim y_s,$$

by which we mean that for a fixed $\tau > 0$ and any $\delta, \delta' > 0$,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi_s^t - a\lambda \int_0^s y_w^2\,\mathrm{d}w\big| < \delta, \big|\eta_s^t - y_s\big| < \delta', \text{for all } s \in [0,\tau]\Big) \to 1,$$

which can equally be written as

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi_s - a\lambda t \int_0^s y_w^2\,\mathrm{d}w\big| < \delta t, \big|\eta_s - y_s\sqrt{t}\big| < \delta'\sqrt{t}, \text{for all } s \in [0,\tau]\Big) \to 1,$$

equivalently,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\big(A_t(\xi)\big) = 1.$$

At the same time, Theorem 4.4 says,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big) = 1.$$

Recalling that $P(A\cap B)=1-P(A^c\cup B^c)\geq 1-P(A^c)-P(B^c)$, we find

$$\lim_{t\to\infty}\tilde{\mathbb{Q}}_t\Big(A_t(\xi);\ \sup_{s\in[0,\tau]}Z_t(s)\leq e^{(J(\tau)+\varepsilon)t}\Big)=1.$$

Since $\varepsilon>0$ was arbitrary, it follows from (35) that for all fixed $\tau>0$ and $\delta,\delta'>0$,

$$\liminf_{t\to\infty}t^{-1}\log P\Big(A_{t,\tau}^{\delta,\delta'}\Big)\geq -J(\tau),$$

which gives the proof of the lower bound of Theorem 3.2. □

## 6 Proving the Large-Deviations Upper Bounds

We first give a quick and direct martingale proof of the upper bound of Theorem 1.1 that identifies the optimal path to reach $(\beta t,\kappa\sqrt{t})$ at large time $\tau$. This follows along similar lines to the almost-sure upper bound of Sect. 2.1.

Let $\gamma,\kappa>0$. Then for $\theta\in(0,-\lambda_{min})$ and $\phi\in(0,\psi_\theta^+]$ to ensure that expectations remain finite for all time and recalling the martingale $Z_\theta^+$ at Eq. (9),

$$\begin{aligned}
&P\Big(\exists u\in N_\tau: X_u(\tau)\leq-\gamma t, Y_u(\tau)^2\geq\kappa^2 t\Big)\\
&\leq P\Big(\sum_{u\in N_\tau}\mathbf{1}\{X_u(\tau)+\gamma t\leq 0, Y_u(\tau)^2-\kappa^2 t\geq 0\}\Big)\\
&\leq P\Big(\sum_{u\in N_\tau}e^{-\theta(X_u(\tau)+\gamma t)+\phi(Y_u(\tau)^2-\kappa^2 t)}\Big)\\
&= e^{-\theta\gamma t-\kappa^2\phi t+E_\theta^+\tau}P\Big(\sum_{u\in N_\tau}e^{-\theta X_u(\tau)+\phi Y_u(\tau)^2-E_\theta^+\tau}\Big)\\
&\leq e^{-\theta\gamma t-\kappa^2\phi t+E_\theta^+\tau}P\Big(Z_\theta^+(\tau)\Big)=e^{-(\theta\gamma+\kappa^2\phi)t+E_\theta^+\tau}.
\end{aligned}$$

Hence, taking $t$ limits and then optimising over $\theta$ we find

$$\limsup_{t\to\infty}t^{-1}\log P\Big(\exists u\in N_\tau: X_u(\tau)\leq-\gamma t, Y_u(\tau)^2\geq\kappa^2 t\Big)\leq-\sup_{\theta>0}\{\theta\gamma+\kappa^2\phi\}=\Theta(\gamma,\kappa) \tag{36}$$

which is then sufficient to give part (b) of Theorem 1.1.

Whilst this upper bound is good enough for sufficiently large $\tau$, for fixed $\tau$ the expectation upper bound can be sharpened with a bit more work to yield part (b) of Theorem 3.2. Firstly, recalling the one-particle picture and that recalling (30), $Z_t(w)=\sum_{u\in N(w)}f_{t,w}(u)$ we have,

$$P(\exists u \in N_\tau \text{ such that } A_t(u)) \le P\left(\sum_{u \in N(\tau)} \mathbf{1}\{A_t(u)\}\right)$$

$$= \tilde{P}\left(e^{\int_0^\tau R(\eta_s)\,ds}; A_t(\xi)\right) = \tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)}; A_t(\xi)\right) \quad (37)$$

Then if a good lower bound for $f_{t,\tau}(\xi)$ can be found on the event $A_t(\xi)$, a suitable upper bound for the required probability will follow.

Consider $y$ and $\tau$ fixed and let us restrict attention to the event $A_t(\xi)$. It is now relatively straightforward making use of Itõ calculus and the constraints from the event, to consider terms in the expression for the spine term $f_{t,\tau}(\xi)$ to derive the following bounds (we omit the proof):

**Proposition 6.1.** *(a) There exists $K > 0$ depending only on $y$ and $\tau$ such that, for all $w \in [0, \tau]$,*

$$\left| \frac{\sqrt{t}}{\sqrt{\theta}} \int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)\,\mathrm{d}B_u(s) - \frac{t}{\theta} \int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s \right| < \frac{t}{\theta} K \delta'$$

*for almost every path in the event $A_t(\xi)$.*

*(b) There exists a $K' > 0$ depending only on $y$ and $\tau$, such that for all $w \in [0, \tau]$,*

$$\left| \int_0^w \eta_s^2\,ds - t \int_0^w y_s^2 ds \right| \le t\delta'(2K' + \tau)$$

*for almost every path in the event $A_t(\xi)$.*

*(c)*

$$\tilde{\mathbb{Q}}_t\left(e^{-\sqrt{at}\lambda \int_0^\tau y_s\,dW_s}; A_t(\xi)\right) \le e^{-a\lambda^2 t \int_0^\tau y_s^2\,ds + t(\frac{1}{2}a\lambda^2\delta'^2\tau + \delta\lambda)}$$

Then,

$$\tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)}; A_t(\xi)\right) = \tilde{\mathbb{Q}}_t\left(e^{\int_0^\tau R(\xi_s)\,\mathrm{d}s} \times e^{-\frac{\sqrt{t}}{\sqrt{\theta}} \int_0^\tau (\dot{y}_s + \frac{\theta}{2} y_s)\,\mathrm{d}B_s} \times e^{\sqrt{at}\lambda \int_0^\tau y_s\,\mathrm{d}W_s}; A_t(\xi)\right)$$

$$\times e^{\frac{t}{2\theta} \int_0^\tau (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s} \times e^{t\frac{a\lambda^2}{2} \int_0^\tau y_s^2\,\mathrm{d}s}$$

$$\le e^{t\int_0^\tau r y_s^2\,\mathrm{d}s + \rho\tau} \times e^{t\delta'(2K' + \tau)} \times e^{-\frac{t}{\theta}\int_0^\tau (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s} \times e^{t\frac{K}{\theta}\delta'}$$

$$\times e^{-ta\lambda^2 \int_0^\tau y_s^2\,\mathrm{d}s} \times e^{t(\frac{1}{2}a\lambda^2\delta'\tau + \delta\lambda)} \times e^{-\frac{t}{2\theta}\int_0^\tau (\dot{y}_s + \frac{\theta}{2} y_s)^2 \mathrm{d}s + t\frac{a\lambda^2}{2}\int_0^\tau y_s^2 \mathrm{d}s}$$

Then, given $\varepsilon > 0$, we can choose $\delta, \delta' > 0$ sufficiently small such that

$$\tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)}; A_t(\xi)\right) \le e^{t\left\{\int_0^\tau r y_s^2\,\mathrm{d}s + \rho\tau - \frac{1}{2\theta}\int_0^\tau (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s - \frac{a\lambda^2}{2}\int_0^\tau y_s^2 \mathrm{d}s + \varepsilon\right\}} = e^{-tJ(\tau) + \rho\tau + \varepsilon\tau}$$

□

## 7 A Spine Proof of the Martingale Upper-Bound

In this section we use the spine decomposition of the martingale $Z_t$ to prove Theorem 4.3.

It is Jensen's inequality that immediately allows us to concentrate on the spine decomposition since

$$\mathbb{Q}_t\big(Z_t(\tau)^{\alpha}\big) \leq \tilde{\mathbb{Q}}_t\Big(\tilde{\mathbb{Q}}_t\big(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\big)^{\alpha}\Big), \qquad \text{for } \alpha \in [0,1].$$

Similar tricks to compute the moments of martingales with Jensen's inequality and the spine decomposition can be found in Iksanov [9] and in Hardy and Harris [5].

The spine decomposition of $Z_t(\tau)$ is

$$\tilde{\mathbb{Q}}_t\left(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\right) = e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s-\rho\tau} e^{\left[\sqrt{a}\lambda\int_0^\tau y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s\right]+\left[\frac{1}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^\tau\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right]}$$
$$+\sum_{k=1}^{n_\tau} e^{-r\int_0^{S_k}\eta_s^2\,\mathrm{d}s-\rho S_k} e^{\left[\sqrt{a}\lambda\int_0^{S_k}y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^{S_k}y_s^2\,\mathrm{d}s\right]+\left[\frac{1}{\sqrt{\theta}}\int_0^{S_k}\left(\dot{y}_s+\frac{\theta}{2}y_s\right)\mathrm{d}B_s-\frac{1}{2\theta}\int_0^{S_k}\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right]}.$$

We consider the two parts of this spine decomposition separately—the *spine term* and then the *sum term*—and aim to show that they both have exponential growth of the same order.

**Definition 7.1.** We define

$$\textbf{spine term} := e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s-\rho\tau} e^{[\sqrt{a}\lambda\int_0^\tau y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]},$$

and

$$\textbf{sum term} :=$$
$$\sum_{k=1}^{n_\tau} e^{-r\int_0^{S_k}\eta_s^2\,\mathrm{d}s-\rho S_k} e^{[\sqrt{a}\lambda\int_0^{S_k}y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^{S_k}y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]}.$$

In each case we first use some martingale techniques to factor out exponential terms that give us the correct growth rate (and here we are guided by the heuristics) and then use Varadhan's lemma to show that the remaining terms do not contribute any further exponential growth. The spine term is simpler to deal with and is considered first.

### 7.1 Factoring Out the Spine Term

Girsanov's theorem (see Øksendal [17]) states that under the new measure $\tilde{\mathbb{Q}}_t$ we have

$$\mathrm{d}B_s = \mathrm{d}\tilde{B}_s + \frac{\sqrt{t}}{\sqrt{\theta}}\Big(\dot{y}_s + \frac{\theta}{2}y_s\Big)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}W_s = \mathrm{d}\tilde{W}_s + \sqrt{at}\lambda y_s\,\mathrm{d}s, \tag{38}$$

where $\tilde{B}$ and $\tilde{W}$ are BMs under $\tilde{\mathbb{Q}}_t$, and these representations can be substituted into the spine term to give

$$\begin{aligned}\textbf{Spine term} &= e^{t\int_0^\tau \frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s-\rho\tau} \times e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\\ &= e^{tJ(\tau)-\rho\tau} \times e^{rt\int_0^\tau[(\eta_s^t)^2-y_s^2]\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.\end{aligned} \tag{39}$$

Using the standard martingale

$$e^{\alpha\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s-\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s},$$

we can factor out one of the terms of the expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big)\\ &= e^{\alpha tJ(\tau)-\alpha\rho\tau} e^{\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha r\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big).\end{aligned}$$

This final expectation can be dealt with by another change of measure:

$$\begin{aligned}&\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big)\\ &\quad= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s} \times \tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\Big),\\ &\quad= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\end{aligned}$$

where we have used the martingale

$$e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau \dot{y}_s+\frac{\theta}{2}y_s\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}$$

to change the measure from $\tilde{\mathbb{Q}}_t$ to $\tilde{\mathbb{Q}}_t^\alpha$. Another application of the Girsanov theorem implies that under the measure $\tilde{\mathbb{Q}}_t^\alpha$, the rescaled process $\eta_s^t$ satisfies

$$\mathrm{d}\big(\eta_s^t-(1+\alpha)y\big)=\frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\bar{B}_s-\frac{\theta}{2}\big(\eta_s^t-(1+\alpha)y_s\big)\mathrm{d}s \tag{40}$$

where $\bar{B}_s$ is a Brownian motion, which is to say that $\eta^t$ is an $\mathrm{OU}(\frac{\theta}{t},\frac{\theta}{2})$ along the *perturbed* path $(1+\alpha)y$.

Putting this all together we are left with a neat factorisation expressed in terms of the rescaled-type process $\eta_s^t$:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^{\alpha}) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}e^{\alpha^2 tM(\tau)}\times\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^{\tau}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right),\\ &\le e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\times\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^{\tau}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right),\end{aligned} \tag{41}$$

where we remember that $M(\tau):=\int_0^{\tau}\left[\frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\right]\mathrm{d}s$. The term $\alpha\rho\tau$ becomes insignificant in the large-deviations limit (for which $t\to\infty$), and therefore it is convenient to have removed it here.

The martingale techniques have now played their part, and we move on to use Varadhan's lemma to show that the term $\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^{\tau}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$ decays exponentially as $t\to\infty$.

## 7.2 A First Application of Varadhan's Lemma

Under the measure $\tilde{\mathbb{Q}}_t^{\alpha}$ the process $\eta^t$ is an $\mathrm{OU}(\frac{\theta}{t},\frac{\theta}{2})$ along the perturbed path $(1+\alpha)y$ (or equivalently we can say that $[\eta_s^t-(1+\alpha)y_s]$ is an $\mathrm{OU}(\frac{\theta}{t},\frac{\theta}{2})$), and therefore it satisfies a large-deviations principle:

**Theorem 7.2.** *If we use the notation $\eta^t$ to refer to the element (path) in $C[0,\tau]$ defined by*

$$\eta^t(s):=\eta_s^t,\quad \textit{for } s\in[0,\tau]$$

*then there is a large-deviations principle for $\eta^t$ with respect to the measure $\tilde{\mathbb{Q}}_t^{\alpha}$:*

- *Upper bound: If $C$ is a closed subset of $C[0,\tau]$ then*

$$\limsup_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}(\eta_s^t\in C)\le -\inf_{g\in C} I(g,\tau),$$

- *Lower bound: If $V$ is an open subset of $C[0,\tau]$ then*

$$\liminf_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}(\eta_s^t\in V)\ge -\inf_{g\in V} I(g,\tau),$$

*where*

$$I(g,w) := \int_0^w \frac{1}{2\theta}\left[\dot{g}_s + \frac{\theta}{2}g_s - (1+\alpha)\left(\dot{y}_s + \frac{\theta}{2}y_s\right)\right]^2 \mathrm{d}s.$$

*if $g \in C[0,\tau]$ with $g(0)=0$ is square integrable along with its derivative; otherwise, we define $I(g)=\infty$.*

Given the upper bound (41), we now want to understand the behaviour of the expectation term $\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$ for large $t$. Varadhan's lemma is a common way to deal with expectations of this form, and we quote the following from Dembo and Zeitouni [2]:

**Theorem 7.3 (Varadhan).** *Let $(X^t)_{t\geq 0}$ be a family of random variables taking values in the space $\mathcal{X}$, and let $\mu_t$ denote the probability measures associated with $(X^t)_{t\geq 0}$.*

*Suppose that the measures $\mu_t$ satisfy the LDP with a good rate function $I:\mathcal{X}\to[0,\infty]$, and let $\phi:\mathcal{X}\to\mathbb{R}$ be any continuous function. Assume further that the following moment condition holds for some $\gamma>1$:*

$$\limsup_{t\to\infty} t^{-1}\log\mathbb{E}\left[e^{\gamma t\phi(X^t)}\right] < \infty. \tag{42}$$

*Then*

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left[e^{t\phi(X^t)}\right] = \sup_{x\in\mathcal{X}}\left[\phi(x)-I(x)\right].$$

This powerful result will confirm our hopes that the expectation decays as $t\to\infty$.

**Theorem 7.4.** *For each $\alpha>0$ the expectation decays exponentially to 0:*

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right) < 0. \tag{43}$$

*For small $\alpha$ we can give more precise expression of the exponential decay:*

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$$
$$= -\alpha^2\left\{k_1\left[\int_0^\tau r y_s^2\,\mathrm{d}s\right] + k_2\left[\frac{1}{2\theta}\int_0^\tau\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right]\right\} + o(\alpha^2), \quad \text{as } \alpha\to 0,$$

*where $k_1$ and $k_2$ are strictly positive.*

*Proof.* Given the large-deviations principle stated in Theorem 7.2, we shall be equating $\mathcal{X}=C[0,\tau]$, $X^t=\eta^t$ and $\mu_t=\tilde{\mathbb{Q}}_t^{\alpha}$ and have $\phi(\eta^t)=\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s$; the moment condition (42) is satisfied because

$$\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{2\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right) < e^{2\alpha rt\int_0^\tau y_s^2\,\mathrm{d}s}.$$

Varadhan's lemma implies that

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)=\sup_{z\in C_0[0,\tau]}\left\{\left(\int_0^\tau \alpha r\left[y_s^2-z_s^2\right]\mathrm{d}s\right)-I(z,\tau)\right\}. \tag{44}$$

Standard Euler–Lagrange techniques for maximising the right-hand integral lead to the following differential equation for $z$:

$$\ddot{z}_s-\left(\frac{\theta^2}{4}+2\theta\alpha r\right)z_s=(1+\alpha)\ddot{y}_s-\frac{\theta^2}{4}(1+\alpha)y_s, \tag{45}$$

which in general will give the optimal path as a solution in terms of the given path $y$.

With the *specific* path (20) that resulted from the Harris and Git optimisations of the large-deviations heuristics, it is relatively simple to solve (45) and find that the optimal path $z$ is just a constant multiple of the path $y$:

$$z_s=K_\alpha y_s,\qquad \text{where } K_\alpha:=\frac{\mu_\lambda^2-\theta^2/4}{\mu_\lambda^2-\theta^2/4-2\theta\alpha r}(1+\alpha). \tag{46}$$

Substituting for $z$ into (44) we find that

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^{\alpha}\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$$
$$=\alpha(1-K_\alpha^2)\left[\int_0^\tau ry_s^2\,\mathrm{d}s\right]-(K_\alpha-(1+\alpha))^2\left[\frac{1}{2\theta}\int_0^\tau\left(\dot{y}_s+\frac{\theta}{2}y_s\right)^2\mathrm{d}s\right], \tag{47}$$

and the following simple bound on $K_\alpha$ implies that this is a negative quantity

**Lemma 7.5.** *For all $\alpha>0$,*

$$1<K_\alpha<1+\alpha. \tag{48}$$

This small lemma can be proved with simple algebra from the definition of $\mu_\lambda$ given at (21): we can use this to show that $\mu_\lambda^2-\theta^2/4=-2\theta r-a\theta\lambda^2<0$, from which it follows that

$$\frac{1}{1+\alpha}<\frac{\mu_\lambda^2-\theta^2/4}{\mu_\lambda^2-\theta^2/4-2\theta\alpha r}<1.$$

If we make a Taylor expansion about $\alpha=0$:

$$\frac{\mu_\lambda^2-\theta^2/4}{\mu_\lambda^2-\theta^2/4-2\theta\alpha r}=\frac{1}{1-k\alpha}=1+k\alpha+k^2\alpha^2+o(\alpha^2)+\cdots$$

where $k:=\frac{2\theta r}{\mu_\lambda^2-\theta^2/4}$, it follows that for strictly positive constants $k_1$ and $k_2$,

$$\alpha(1-K_\alpha^2)=-k_1\alpha^2+o(\alpha^2),\quad\text{and}\quad (K_\alpha-(1+\alpha))^2=k_2\alpha^2+o(\alpha^2)\quad\text{as }\alpha\to 0,$$

completing the proof □

### 7.3 Dealing with the Sum Term

Focusing on the sum term, we can again substitute for $\mathrm{d}W_s$ and $\mathrm{d}B_s$ with (38) and immediately factor out the term $J(S_k)$ by overestimating

$$\textbf{sum term} = \sum_{k=1}^{n_\tau} e^{tJ(S_k)-\rho S_k} e^{r\int_0^{S_k}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{[\sqrt{at}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}$$

$$\leq e^{t\left(\sup_{0\leq w\leq\tau}J(w)\right)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k}\eta_s^2-y_s^2\,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.$$

For the particular path $y$ that we chose at (20), it was shown by Git et al. [4] that

$$\sup_{0\leq w\leq\tau} J(w) = J(\tau)$$

and therefore we have

$$\textbf{sum term} \leq e^{tJ(\tau)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k}\eta_s^2-y_s^2\,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.$$

The following small result is very useful for dealing with the sum term:

**Proposition 7.6.** *If $\alpha\in(0,1]$ and $u,v>0$ then $(u+v)^\alpha\leq u^\alpha+v^\alpha$.*

This proposition implies that for $0\leq\alpha\leq 1$,

$$\tilde{\mathbb{Q}}_t\left(\textbf{sum term}^\alpha\right) \leq e^{\alpha tJ(\tau)}\,\tilde{\mathbb{Q}}_t\left(\sum_{k=1}^{n_\tau} e^{\alpha rt\int_0^{S_k}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{at}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\right),$$

and we can transform the sum into an integral by standard techniques (see, e.g., Kallenberg [10]), since the fission times on the spine form a Cox process of rate $2(r\eta_w^2+\rho)$, as explained in Theorem 4.1:

$$= \; 2e^{\alpha tJ(\tau)}\,\tilde{\mathbb{Q}}_t\left(\int_0^\tau e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\left[r\eta_w^2+\rho\right]\mathrm{d}w\right);$$

Fubini's theorem can be applied to this, and the transformations that worked on the spine term to give (41) can here too be applied to arrive at

$$= 2\,e^{\alpha tJ(\tau)}\int_0^\tau e^{\alpha^2\int_0^w\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s} e^{\frac{\alpha^2}{2\theta}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w,$$

$$\leq 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\times\int_0^\tau\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w.$$

We want to take advantage of the fact that the terms in the integral look similar to those already dealt with for the spine term. A first step in this direction is to replace the random factor $rt(\eta_w^t)^2$ at the front of the expectation with the deterministic $rty_w^2$, and since the value of $\alpha$ will eventually be chosen and fixed the following estimate is sufficient for our purposes:

**Lemma 7.7.** *For all $\alpha > 0$ and for all large enough $t$,*

$$\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w \le \frac{1}{\alpha}+\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w.$$

*Proof.* Since

$$\frac{\partial}{\partial w}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right) = \tilde{\mathbb{Q}}_t^\alpha\left(\alpha rt[y_w^2-(\eta_w^t)^2]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right), \tag{49}$$

it follows that

$$\begin{aligned}\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{-\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right) &= \left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\\ &\quad -\frac{1}{\alpha}\frac{\partial}{\partial w}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right).\end{aligned}$$

Integration by parts now proves

$$\begin{aligned}\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w &= \int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w\\ &\quad +\frac{1}{\alpha}\left[1-\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\right].\end{aligned}$$

The exponential decay proved in Theorem 7.4 implies $\lim_{t\to\infty}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$ $= 0$, and this completes the proof □

It follows therefore that for all large enough $t$,

$$\tilde{\mathbb{Q}}_t\left(\textbf{sum term}^\alpha\right)\le 2e^{\alpha tJ(\tau)}e^{\alpha^2tM(\tau)}\times\left(\frac{1}{\alpha}+\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w\right).$$

We now make some simple overestimates of the integral. Firstly, it is immediate that

$$\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w \le \left[rt\kappa^2+\rho\right]\int_0^\tau\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w$$

since $\left(\sup_{0\le w\le\tau}y_w^2\right)=\kappa^2$. Then, for each $w\in[0,\tau]$, it is true by definition that

$$e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\le e^{\alpha rt\left(\sup_{0\le w\le\tau}\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)},$$

and therefore

$$\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\le\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right).$$

Since this holds for all $w \in [0, \tau]$ we can deduce

$$\sup_{0 \le w \le \tau} \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} \right) \le \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \left( \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s \right)} \right),$$

which we can use to get

$$\begin{aligned} \int_0^\tau \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} \right) \mathrm{d}w &\le \tau \times \sup_{0 \le w \le \tau} \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} \right), \\ &\le \tau \times \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \left( \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s \right)} \right). \end{aligned}$$

Thus we arrive at a simple upper bound for the sum term: for all $\alpha \in [0, 1]$ and all large $t$,

$$\tilde{\mathbb{Q}}_t \left( \textbf{sum term}^\alpha \right) \le 2\, e^{\alpha t J(\tau)} e^{\alpha^2 t M(\tau)} \left\{ \frac{1}{\alpha} + \left[ rt\kappa^2 + \rho \right] \tau \, \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \left( \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s \right)} \right) \right\}. \tag{50}$$

## 7.4 A Second Application of Varadhan's Lemma

We already applied Varadhan's lemma to the term $\tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} \right)$, and now we show how it can in fact deal with the more complex term $\tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} \right)$ without much more effort.

Once again the observation

$$\tilde{\mathbb{Q}}_t^\alpha \left( e^{2\alpha rt \left( \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s \right)} \right) < \tilde{\mathbb{Q}}_t^\alpha \left( e^{2\alpha rt\tau (\sup_w y_w^2)} \right)$$

shows that the moment condition (42) is satisfied, and therefore from Varadhan's lemma, Theorem 7.3, it follows that

$$\lim_{t \to \infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \left( e^{\alpha rt \left( \sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s \right)} \right) = \sup_{z \in C_0[0,\tau]} \left\{ \left( \sup_{0 \le w \le \tau} \int_0^w \alpha r \left[ y_s^2 - z_s^2 \right] \mathrm{d}s \right) - I(z, \tau) \right\}.$$

For any path $z$, the action functional $I(z, w)$ is nondecreasing in $w$ and therefore

$$\left( \int_0^w \alpha r \left[ y_s^2 - z_s^2 \right] \mathrm{d}s \right) - I(z, \tau) \le \left( \int_0^w \alpha r \left[ y_s^2 - z_s^2 \right] \mathrm{d}s \right) - I(z, w),$$

and taking the supremum over $w \in [0, \tau]$ of both sides we deduce

$$\left( \sup_w \int_0^w \alpha r \left[ y_s^2 - z_s^2 \right] \mathrm{d}s \right) - I(z, \tau) \le \sup_w \left\{ \left( \int_0^w \alpha r \left[ y_s^2 z_s^2 \right] \mathrm{d}s \right) - I(z, w) \right\},$$

We now take the supremum of both sides over the set of paths $z \in C_0[0,\tau]$ and interchange the order to obtain

$$\sup_z\Big\{\Big(\sup_w \int_0^w \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\} \le \sup_z \sup_w \Big\{\Big(\int_0^w \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\Big\}$$
$$= \sup_{0\le w\le\tau} \sup_z \Big\{\Big(\int_0^w \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\Big\}. \tag{51}$$

If we compare the term

$$\sup_z\Big\{\Big(\int_0^w \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\Big\}$$

with (44) from our first application of Varadhan's lemma, it is clear that Euler–Lagrange optimisation techniques will result in exactly the same optimal path for this integral, namely $z_s = K_\alpha y_s$ as at (46). Furthermore, evaluating the left-hand side of (51) shows that we actually have the equality

$$\sup_z\Big\{\Big(\int_0^\tau \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\} = \sup_z\Big\{\Big(\sup_w \int_0^w \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\},$$
$$= \alpha(1-K_\alpha^2)\Big[\int_0^\tau r y_s^2\,\mathrm{d}s\Big] - (K_\alpha - (1+\alpha))^2\Big[\frac{1}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2\,\mathrm{d}s\Big],$$
$$< 0 \ \big(\text{and } = O(\alpha^2) \text{ as } \alpha \to 0\big).$$

Consequently we see that there is no difference in the growth rate between the remaining terms of the spine term and the sum term:

$$\lim_{t\to\infty} t^{-1}\log \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s\right)}\Big) = \lim_{t\to\infty} t^{-1}\log \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big) < 0. \tag{52}$$

## 7.5 *Concluding the Upper-Bound for* $Z_t(\tau)$

We have shown that

$$\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) \le e^{\alpha t J(\tau)} e^{\alpha^2 t M(\tau)} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big),$$

and since we clearly have $\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\leq\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right)$, it follows that

$$\begin{aligned}\tilde{\mathbb{Q}}_t\left(Z_t(\tau)^\alpha\right)&\leq\tilde{\mathbb{Q}}_t\left(\textbf{spine term}^\alpha\right)+\tilde{\mathbb{Q}}_t\left(\textbf{sum term}^\alpha\right)\\&\leq e^{\alpha tJ(\tau)}e^{\alpha^2tM(\tau)}\left\{\left(1+2\left[rt\kappa^2+\rho\right]\tau\right)\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\mathrm{d}s\right)}\right)+\frac{2}{\alpha}\right\}.\end{aligned}\tag{53}$$

Thus

$$\lim_{t\to\infty}t^{-1}\log\tilde{\mathbb{Q}}_t\left(Z_t(\tau)^\alpha\right)\leq\alpha J(\tau)+\alpha^2M(\tau),$$

and the proof of Theorem 4.3 is completed. □

## 8 An Alternative Approach to the Lower Bound

Finally, we suggest an alternative approach to gaining the lower bound of Theorem 3.2. The key tool is still the spine decomposition and this approach should work more generally. In particular, it would not require calculation of the $\tilde{\mathbb{Q}}_t(Z_t(\tau)^\alpha)$ in order to control the size of the martingale $Z_t(s)$, which may be advantageous in some models. However, in this instance, there appears little to choose between the two approaches.

Making use of some simple estimation, the tower property, recalling that $\mathcal{G}_\infty$ contains information only about the spine's spatial trajectory and using the conditional form of Jensen's inequality, we have

$$\begin{aligned}P\left(\exists u\in N_\tau\text{ such that }A_t(u)\right)&=\tilde{\mathbb{Q}}_t\left(\frac{1}{Z_t(\tau)};\exists u\in N_\tau,A_t(u)\right)\\&\geq\tilde{\mathbb{Q}}_t\left(\frac{1}{Z_t(\tau)};A_t(\xi)\right)=\tilde{\mathbb{Q}}_t\left\{\tilde{\mathbb{Q}}_t\left(\frac{1}{Z_t(\tau)}\middle|\mathcal{G}_\infty\right);A_t(\xi)\right\}\\&\geq\tilde{\mathbb{Q}}_t\left\{\frac{1}{\tilde{\mathbb{Q}}_t\left(Z_t(\tau)|\mathcal{G}_\infty\right)};A_t(\xi)\right\}\\&=\tilde{\mathbb{Q}}_t\left\{\frac{1}{f_{t,\tau}(\xi)+\int_0^\tau 2R(\xi_s)f_{t,s}(\xi)\,\mathrm{d}s};A_t(\xi)\right\}.\end{aligned}$$

If a good upper bound for $f_{t,\tau}(\xi)$ and $\int_0^\tau 2R(\xi_s)f_{s,\tau}(\xi)\,\mathrm{d}s$ can be found on the event $A_t(\xi)$, a suitable lower bound for the required probability will follow. This is a similar idea as used in the upper bound approach (see Eq. (37)), except some additional work would be required to control the integral over time (although in this situation, it will be dominated in exponential order by the final value of $f_{t,\tau}(\xi)$ to match the upper bound).

## References


1. Athreya, K.B. *Change of measures for Markov chains and the* $L\log L$ *theorem for branching processes*, Bernoulli **6** (2000), no. 2, 323–338.
2. Dembo, A. and Zeitouni, O. *Large deviations techniques and applications*, Springer, 1998.
3. Engländer, J. and Kyprianou, A.E. *Local extinction versus local exponential growth for spatial branching processes*, Ann. Probab. **32** (2004), no. 1A, 78–99.
4. Git, Y., Harris, J.W. and Harris, S.C *Exponential growth rates in a typed branching diffusion*, Ann. App. Probab., **17** (2007), no. 2, 609-653.
5. Hardy, R., and Harris, S. C. *A spine approach to branching diffusions with applications to* $L^p$*-convergence of martingales*, Séminaire de Probabilités, **XLII**, (2009), 281–330.
6. Hardy, R., and Harris, S. C. *A conceptual approach to a path result for branching Brownian motion*, Stochastic Processes and their Applications, **116** (2006), no. 12, 1992–2013.
7. Harris, S. C. and Williams, D. *Large deviations and martingales for a typed branching diffusion. I*, Astérisque (1996), no. 236, 133–154, Hommage à P. A. Meyer et J. Neveu.
8. Harris, S.C. *Convergence of a 'Gibbs-Boltzmann' random measure for a typed branching diffusion*, Séminaire de Probabilités, **XXXIV**, Lecture Notes in mathematics, vol. 1729, Springer, (2000), 133–154.
9. Iksanov, A.M. *Elementary fixed points of the BRW smoothing transforms with infinite number of summands*, Stochastic Process. Appl. **114** (2004), no. 1, 27–50.
10. Kallenberg, O. *Foundations of modern probability*, Springer-Verlag, 2002.
11. Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y. *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185.
12. Kyprianou, A.E. *Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris's probabilistic analysis*, **40** (2004), no. 1, 53–72.
13. Kyprianou, A.E. and Rahimzadeh Sani, A. *Martingale convergence and the functional equation in the multi-type branching random walk*, Bernoulli **7** (2001), no. 4, 593–604.
14. Lee, T. *Some large-deviation theorems for branching diffusions*, Ann. Probab. **20** (1992), no. 3, 1288–1309.
15. Lyons, R. *A simple path to Biggins' martingale convergence for branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221.
16. Lyons, R., Pemantle, R. and Peres, Y. *Conceptual proofs of* $L\log L$ *criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), no. 3, 1125–1138.
17. Øksendal, B. *Stochastic differential equations*, fifth ed., Springer, 2000.
18. Olofsson, P. *The* $x\log x$ *condition for general branching processes*, J. Appl. Probab. **35** (1998), no. 3, 537–544.

# Longtime Behavior for Mutually Catalytic Branching with Negative Correlations

**Leif Döring and Leonid Mytnik**

**Abstract** In several examples, dualities for interacting diffusion and particle systems permit the study of the longtime behavior of solutions. A particularly difficult model in which many techniques collapse is a two-type model with mutually catalytic interaction introduced by Dawson/Perkins for which they proved under some assumptions a dichotomy between extinction and coexistence directly from the defining equations.

In the present chapter we show how to prove a precise dichotomy for a related model with negatively correlated noises. The proof uses moment bounds on exit times of correlated Brownian motions from the first quadrant and explicit second moment calculations. Since the uniform integrability bound is independent of the branching rate our proof can be extended to infinite branching rate processes.

**Keywords** Longtime Behavior • Branching process • Planar Brownian Motion • Duality

**Mathematics Subject Classification (2000):** Primary 60J80; Secondary 60J85.

---

L. Döring (✉)
LPMA, Université Paris VI, Tours 16/26, 4 Place Jussieu, 75005 Paris, France

LM is partly supported by the Israel Science Foundation and B. and G. Greenberg Research Fund (Ottawa)
e-mail: leif.doering@googlemail.com

L. Mytnik
Faculty of Industrial Engineering and Management Technion Israel Institute of Technology, Haifa 32000, Israel

LD was supported by the Fondation Science Matématiques de Paris
e-mail: leonid@ie.technion.ac.il

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_6, 

# 1 Introduction

A classical task in the theory of interacting particle systems and interacting diffusion models is a precise understanding of the longtime behavior of the system. For many models dichotomies between extinction and survival have been revealed. In most cases proofs are based on clever duality constructions or explicit representations of the particular process. A good example is the *voter model* for which a graphical construction can be applied that reduces extinction problems to hitting problems of random walks (see, e.g., Chap. V of [L05] for many beautiful results).

In this chapter we aim at giving simple proofs for the dichotomy in a related class of mutually interacting diffusion processes with negatively correlated driving noises. Interestingly, the particular structure of the processes permits a second moment calculation that leads to the precise characterization of survival/extinction via recurrence/transience of the underlying migration mechanism. The approach is more direct than the previously used extension of arguments due to [DP98] for which regularity assumption on the underlying migration mechanism was imposed.

## *1.1 Finite Rate Symbiotic Branching Processes*

In 2004, Etheridge and Fleischmann [EF04] introduced a stochastic spatial model of two interacting populations known as the (finite rate) *symbiotic branching model*, parametrized by a parameter $\rho \in [-1,1]$ governing the correlations between the driving noises and a branching parameter $\gamma > 0$ amplifying the strength of the noises. For a discrete spatial version of their stochastic heat equations, we consider the system of interacting diffusions on a countable set $S$ with values in $\mathbb{R}_{\geq 0}$, defined by the coupled stochastic differential equations

$$\mathrm{SBM}_\gamma = \mathrm{SBM}(\rho,\gamma)_{u_0,v_0}: \quad \begin{cases} du_t(i) = \mathcal{A}u_t(i)\,dt + \sqrt{\gamma u_t(i)v_t(i)}\,dB^1_t(i), \\ dv_t(i) = \mathcal{A}v_t(i)\,dt + \sqrt{\gamma u_t(i)v_t(i)}\,dB^2_t(i), \\ u_0(i) \geq 0, \quad i \in S, \\ v_0(i) \geq 0, \quad i \in S, \end{cases}$$

where $\{B^1(i), B^2(i)\}_{i\in S}$ is a $(\mathbb{R}^2)^S$-valued centered Gaussian process with covariance structure

$$\mathbb{E}\big[B^n_t(i)B^m_t(j)\big] = \begin{cases} \rho t & : i = j \text{ and } n \neq m, \\ t & : i = j \text{ and } n = m, \\ 0 & : \text{otherwise}. \end{cases}$$

The migration operator $\mathcal{A}$ is defined as

$$\mathcal{A}w(i) = \sum_{j\in S} a(i,j)w(j),$$

where $\big(a(i,j)\big)_{i,j\in S}$ will always be assumed to be the $Q$-matrix of a symmetric continuous-time $S$-valued Markov process with bounded jump rate, i.e.,

$$\sup_{k\in S} |a(k,k)| < \infty.$$

Some care is needed to define properly the state-space of solutions. Here, we consider solutions in the space $L^\beta \subset (\mathbb{R}_+^2)^S$, usually referred to as Liggett–Spitzer space in the theory of interacting particle systems. Fixing a test-sequence $\beta \in (0,\infty)^S$ such that

$$\sum_{i\in S}\beta(i) < \infty \quad \text{and} \quad \sum_{i\in S}\beta(i)|a(i,k)| < M\beta(k)$$

for all $k \in S$, the state-space of solutions becomes

$$L^\beta := \big\{(u,v) : S \to \mathbb{R}^+ \times \mathbb{R}^+, \langle u,\beta\rangle, \langle v,\beta\rangle < \infty\big\},$$

where $\langle f,g\rangle = \sum_{k\in S} f(k)g(k)$. The existence of the test-sequence $\beta$ is ensured by Lemma IX.1.6 of [L05]. For $u \in \mathbb{R}^S$, let

$$\|u\|_{\beta,1} = \sum_{k\in S} |u(k)|\beta(k).$$

Furthermore, for $w = (u,v) \in L^\beta$, let $\|w\|_\beta = \|u\|_{\beta,1} + \|v\|_{\beta,1}$. Note that $\|\cdot\|_\beta$ defines a topology on $L^\beta$. We will henceforth assume that $L^\beta$ is equipped with this topology.

We shall say that a pair $(u_t^\gamma, v_t^\gamma)$ of continuous $L^\beta$-valued adapted processes is a solution of $\mathrm{SBM}_\gamma$ on the stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ if there is a family $\big\{B^1(i), B^2(i)\big\}_{i\in S}$ of $\mathcal{F}_t$-Brownian motions with the aforementioned correlation structure and the stochastic equation is fulfilled. Hence, the solutions are defined in the weak sense. They are usually first built on finite subsets of $S$ via standard SDE theory and then transferred to $S$ via a weak-limiting procedure. To avoid long repetitions we refer the reader to Sect. 3 of [BDE11] for a summary of existence of weak solutions, uniqueness and duality results for symbiotic branching processes in the case of $\mathcal{A} = \Delta$ being the discrete Laplacian; note that these results can be easily translated for the general case of $\mathcal{A}$ treated here. We also refer the reader to [DM11] for a general review of the subject.

Let us be more specific about the migration operator $\mathcal{A}$. There are a number of typical choices for $\mathcal{A}$ which the reader might keep in mind. First of all it is the case of the discrete Laplacian

$$\mathcal{A}w(i) = \Delta w(i) = \sum_{|k-i|=1} \frac{1}{2d}(w(k) - w(i)),\ i \in \mathbb{Z}^d,$$

which describes nearest neighbor interaction on $S = \mathbb{Z}^d$. There are also the cases of complete graph interaction on a finite set $S$ corresponding to $a(i,j) = |S|^{-1}$, $i \neq j$, and the trivial migration $\mathcal{A}w = 0$ on a single point set $S = \{s\}$ leading to the nonspatial symbiotic branching SDE

$$\begin{cases} du_t = \sqrt{\gamma u_t v_t}\, dB^1_t, \\ dv_t = \sqrt{\gamma u_t v_t}\, dB^2_t, \end{cases} \tag{1}$$

driven by correlated Brownian motions.

So far we did not motivate the reason to study this particular set of stochastic differential equations. Interestingly, symbiotic branching models relate spatial models from different branches of probability theory of the type

$$dw_t(i) = \mathcal{A}w_t(i)\, dt + \sqrt{\gamma f(w_t(i))}\, dB_t(i), \quad i \in S, \tag{2}$$

that are usually referred to as interacting diffusions models. Here are some particular examples:

*Example 1.1.* The *stepping stone model* from mathematical genetics: $f(x) = x(1-x)$ (see for instance [SS80]).

*Example 1.2.* The *parabolic Anderson model (with Brownian potential)* from mathematical physics: $f(x) = x^2$ (see for instance [S92]).

*Example 1.3.* The *super-random walk* from the theory of branching processes: $f(x) = x$ (see for instance Sect. 2.2.4 of [E00] for the continuum analogue).

For the super-random walk, $\gamma$ is the branching rate which in this case is time-space independent. In [DP98], a two-type model based on two super-random walks with time-space dependent branching was introduced. The branching rate for one species at a given site is proportional to the size of the other species at the same site. More precisely, the authors considered for $i \in S$

$$\begin{cases} du_t(i) = \mathcal{A}u_t(i)\, dt + \sqrt{\gamma u_t(i) v_t(i)}\, dB^1_t(i), \\ dv_t(i) = \mathcal{A}v_t(i)\, dt + \sqrt{\gamma u_t(i) v_t(i)}\, dB^2_t(i), \end{cases} \tag{3}$$

where now $\{B^1(i), B^2(i)\}_{i\in S}$ is a family of independent standard Brownian motions. Solutions are called *mutually catalytic branching* processes. The interest in mutually catalytic branching processes originates from the fact that it (resp. its continuum analogue) constitutes a version of two interacting super-processes with random branching environment. Many of the classical tools from the study of superprocesses fail since the branching property breaks down. Nonetheless, a detailed study is possible due to the symmetric nature of the equations. During the last decade, properties of this model were well studied (see for instance [CK00] and [CDG04]).

Let us now see how the above examples relate to the symbiotic branching model $\mathrm{SBM}_\gamma$. For correlation $\rho = 0$, solutions of the symbiotic branching model are obviously solutions of the mutually catalytic branching model. The case $\rho = -1$ with the additional assumption $u_0 + v_0 \equiv 1$ corresponds to the stepping stone model. To see this, observe that in this perfectly negatively correlated case of $B^1(i) = -B^2(i)$, the sum $u+v$ solves a discrete heat equation, and with the further assumption $u_0 + v_0 \equiv 1$, $u+v$ stays constant for all time. Hence, for all $t \geq 0$, $u_t \equiv 1 - v_t$ which shows that $u_t$ is a solution of the stepping stone model with initial condition $u_0$ and $v_t$ is a solution with initial condition $v_0$. Finally, suppose $w$ is a solution of the parabolic Anderson model, then, for $\rho = 1$, the pair $(u,v) := (w,w)$ is a solution of the symbiotic branching model with initial conditions $u_0 = v_0 = w_0$.

This unifying property motivated the study in [BDE11] of the influence of varying $\rho$ on the longtime behavior for $\gamma$ being a fixed constant. Since the stepping stone model and the parabolic Anderson model have very different longtime properties it could be expected to recover some of those properties in disjoint regions of the parameter range. Restricting to $\mathcal{A} = \Delta$ on $\mathbb{Z}^d$ and $d = 1,2$, the longtime behavior of laws and moments has been analyzed. An important observation of [BDE11], which holds equally for quite general $\mathcal{A}$ generating a recurrent Markov process, is the following moment transition for solutions $\mathrm{SBM}_\gamma$ started in homogeneous initial conditions $u_0 = v_0 \equiv 1$:

$$\rho < 0 \quad \Longleftrightarrow \quad \text{There is } \varepsilon > 0 \text{ such that } \mathbb{E}\big[u_t^\gamma(k)^{2+\varepsilon}\big] \text{ is bounded in } t \text{ for all } k \in S. \tag{4}$$

A property of this kind could of course be expected: for $\rho = -1$ solutions correspond to the stepping stone model which is bounded by 1 and hence has bounded moments of all order. The other extremal case $\rho = 1$ corresponds to the parabolic Anderson model that has exponentially increasing second moments (see for instance Theorem 1.6 of [GdH07]).

## *1.2 Infinite Rate Symbiotic Branching Processes*

Looking more closely at the proofs of [BDE11], one observes that (4) can be strengthened as follows:

$$\rho < 0 \quad \Longrightarrow \quad \text{There is } \varepsilon > 0 \text{ such that } \mathbb{E}\big[u_t^\gamma(k)^{2+\varepsilon}\big] \text{ is bounded in } t \textbf{ and } \gamma \text{ for all } k \in S, \tag{5}$$

and this holds for any $\mathcal{A}$ satisfying the assumptions mentioned in the introduction (recall that $(u^\gamma, v^\gamma)$ is a solution to SBM$_\gamma$). This observation shall be combined in the following with a recent development for mutually catalytic branching processes which we now briefly outline.

In [KM10b], for $\rho = 0$, existence of the *infinite rate mutually catalytic branching processes*, appearing as limits in

$$\left(u^\gamma, v^\gamma\right) \stackrel{\gamma \to \infty}{\Longrightarrow} \left(u^\infty, v^\infty\right) \tag{6}$$

has been shown. This process and its properties have been further studied in [KM10a] and [KO10]. In [DM11] the *infinite rate symbiotic branching processes* have been constructed for the whole range of parameters $\rho \in (-1, 1)$. Below we state the result from [DM11] that introduces the infinite rate symbiotic branching process via the limiting procedure (6). Before doing this, let us shortly discuss why this procedure can lead to an exciting process.

To understand the effect of sending $\gamma$ to infinity, one might take a closer look at the nonspatial symbiotic branching SDE (1). Due to the symmetric structure and the Dubins–Schwartz theorem, solutions $(u^\gamma, v^\gamma)$ can be regarded as time-changed correlated Brownian motions with time-change $\gamma \int_0^t u_s^\gamma v_s^\gamma \, ds$. The boundary of the first quadrant

$$E = [0, \infty) \times [0, \infty) \backslash (0, \infty) \times (0, \infty)$$

is absorbing as the noise is multiplicative with diffusion coefficients proportional to the product of both coordinates. Hence, the pair of time-changed Brownian motions stops once hitting $E$. Increasing $\gamma$ only has the effect that the process follows the Brownian paths with higher speed so that $\gamma = \infty$ corresponds to immediately picking a point in $E$ according to the exit-point measure of the two-dimensional Brownian motion at the boundary of the first quadrant and the process stays at this point forever.

Incorporating space, a second effect appears: both types are distributed in space according to the heat flow corresponding to $\mathcal{A}$. This smoothing effect immediately tries to lift a zero coordinate if it was pushed by the exit measure to zero. Interestingly, none of these two effects dominates and a nontrivial limiting process is obtained via the limiting procedure (6).

Before stating the existence theorem, we need to refine the state-space:

$$L^{\beta,E} := L^\beta \cap E^S.$$

The space $L^{\beta,E}$ consists of sequences of pairs of points in $E$ (i.e. one coordinate is zero, one coordinate is nonnegative) with restricted growth condition. The space $L^{\beta,E}$ is equipped with the topology induced from the topology on $L^\beta$ introduced in Sect. 1.1. From the motivation given above for the infinite branching rate limiting behavior of the one-dimensional version (1), the occurrence of the state-space $L^{\beta,E}$ is not surprising: the infinite rate symbiotic branching process, abbreviated as $\mathrm{SBM}_\infty$, takes values in $E$ at each fixed site $k \in S$. Since in this chapter we will be dealing with finite total mass processes, let us also define

$$\begin{aligned} L^1 &:= \big\{(u,v) : S \to \mathbb{R}^+ \times \mathbb{R}^+, \langle u,1\rangle, \langle v,1\rangle < \infty\big\}, \\ L^{1,E} &:= L^1 \cap E^S. \end{aligned}$$

Additionally, we will denote by $D_{L^{\beta,E}}$ the space of RCLL functions on $L^{\beta,E}$. Before we state the result from [DM11] on existence and uniqueness of the infinite rate symbiotic branching processes, we need to introduce some additional notation. For $\rho \in (-1,1)$, any $(x_1,x_2) \in L^\beta$ and any compactly supported $(y_1,y_2) \in L^{1,E}$, set

$$\begin{aligned} \langle\langle x_1,x_2,y_1,y_2\rangle\rangle_\rho = \sum_{k\in S} \Big[ &-\sqrt{1-\rho}\big(x_1(k)+x_2(k)\big)\big(y_1(k)+y_2(k)\big) \\ &+ i\sqrt{1+\rho}\big(x_1(k)-x_2(k)\big)\big(y_1(k)-y_2(k)\big)\Big], \end{aligned}$$

and define $F(x_1,x_2,y_1,y_2) \equiv \exp(\langle\langle x_1,x_2,y_1,y_2\rangle\rangle_\rho)$. Then we have the following result.

**Theorem 1.1.** *Suppose* $\rho \in (-1,1)$ *and* $\{(u^\gamma,v^\gamma)\}_{\gamma>0}$ *is a family of solutions to* $\mathrm{SBM}_\gamma$ *with initial conditions* $(u_0^\gamma,v_0^\gamma) = (u_0,v_0) \in L^{\beta,E}$ *that do not depend on* $\gamma$. *For any sequence* $\gamma_n$ *tending to infinity, we have the convergence in law*

$$(u^{\gamma_n},v^{\gamma_n}) \Longrightarrow (u^\infty,v^\infty), \quad n\to\infty,$$

*in* $D([0,\infty),L^\beta)$ *equipped with the Meyer–Zheng "pseudo-path" topology. The limiting process* $(u^\infty,v^\infty)$ *is almost surely in* $D_{L^{\beta,E}}$ *and it is the unique solution in* $D_{L^{\beta,E}}$ *to the following martingale problem:*

$$\begin{cases} \text{For any compactly supported } (y_1,y_2) \in L^{1,E}, \\ F(u_t^\infty,v_t^\infty,y_1,y_2) - F(u_0,v_0,y_1,y_2) - \int_0^t \langle\langle \mathcal{A}u_s^\infty,\mathcal{A}v_s^\infty,y_1,y_2\rangle\rangle_\rho F(u_s^\infty,v_s^\infty,y_1,y_2)\,ds \\ \textit{is a martingale null at zero.} \end{cases}$$

Since the theorem is a direct combination of Theorems 3.4 and 3.6 from [DM11] its proof is omitted and we will make just a few comments. The convergence of $(u^\gamma,v^\gamma)$ to $(u^\infty,v^\infty)$ is not stated in the Skorohod topology on $D_{L^\beta}$ but instead in the Meyer–Zheng pseudo-path topology. In fact, the weak convergence is impossible in the Skorohod topology since $(u^\gamma,v^\gamma)$ are continuous processes whereas the limiting process $(u^\infty,v^\infty)$ only has RCLL paths (non-continuity does not directly become apparent but follows from a jump-type diffusion characterization), and in the Skorohod topology, the subspace of continuous functions is closed. For the Meyer–Zheng pseudo-path topology this, in fact, is possible (for more details, see [MZ84] and for other interesting developments, see also [J97]).

For the next theorem, we assume that $\rho < 0$ and, in this case, some useful properties of $(u^\infty, v^\infty)$ with initial conditions in $L^{1,E}$ are established.

**Theorem 1.2.** *Suppose $\rho < 0$ and let $(u^\infty, v^\infty)$ be the infinite rate process from Theorem 1.1 with initial conditions $(u_0^\infty, v_0^\infty) \in L^{1,E}$.*

(a) *The total mass processes $\langle u_t^\infty, 1\rangle$ and $\langle v_t^\infty, 1\rangle$ are well-defined square-integrable martingales for $t \geq 0$.*
(b) *There is an $\varepsilon > 0$ such that*

$$\sup_{t>0} \mathbb{E}\left[\langle u_t^\infty, 1\rangle^{2+\varepsilon}\right] < \infty \quad \text{and} \quad \sup_{t>0} \mathbb{E}\left[\langle v_t^\infty, 1\rangle^{2+\varepsilon}\right] < \infty.$$

(c) *If $\xi_t^1, \xi_t^2$ are independent continuous-time Markov processes with generator $\mathcal{A}$, then*

$$\mathbb{E}\left[u_t^\infty(a) v_t^\infty(b)\right] = \mathbb{E}^{a,b}\left[u_0(\xi_t^1) v_0(\xi_t^2) \mathbf{1}_{\{\xi_s^1 \neq \xi_s^2, \forall s \leq t\}}\right]$$

*for any $t \geq 0$ and $a, b \in S$.*

The second moment expression in part c) of the theorem emphasizes the fact that $(u_t^\infty(a), v_t^\infty(a)) \in E$ for any $a \in S$, $t \geq 0$. This becomes clear since for $a = b$, $\xi_0^1 = \xi_0^2$ which gives $\mathbb{E}[u_t^\infty(a) v_t^\infty(a)] = 0$ and this, in turn, implies that $u_t^\infty(a) v_t^\infty(a) = 0$ almost surely, for any $t \geq 0$.

## 2 The Longtime Behavior for Negative Correlations

We now turn to the results of this chapter that address the longtime behavior of finite and infinite rate symbiotic branching processes. The question we address is classical for many particle systems on $[0,\infty)^S$: suppose a system has initial condition $w_0$ that is either infinite (e.g. $w_0 \equiv 1$) or summable (e.g. $w_0 = \mathbf{1}_{\{k\}}$ for some $k \in S$), what can be said about limits of $w_t$ as $t$ tends to infinity? In many situations it turns out that equivalence holds between almost sure extinction for the (finite) total mass process $\langle w_t, 1\rangle$ when $w_0 = \mathbf{1}_k$ and weak convergence of $w_t$ from constant initial states to the absorbing states of the system. Using different duality techniques, this can be shown, for instance, for SBM$_\gamma$, the stepping stone model, the parabolic Anderson model, and for the voter process. It is not known yet whether this property holds for SBM$_\infty$.

Extinction/survival dichotomies are typically of the following type depending only on the recurrence/transience of the migration mechanism:

$$\lim_{t\to\infty} \langle 1, w_t\rangle \overset{a.s.}{=} 0 \text{ for all summable initial conditions}$$

$$\iff \quad \mathbb{P}^i(\xi_t = j \text{ for some } t \geq 0) = 1 \quad \forall i, j \in S, \tag{7}$$

where $\xi_t$ is a continuous-time Markov process with generator $\mathcal{A}$.

## 2.1 Some Known Results

Before stating our main theorem on the recurrence/transience dichotomy we recall some known results.

For $\gamma < \infty$, by writing the stochastic equations for $SBM_\gamma$ in mild form (see Lemma 3.1 below), it can be shown that if $(u_0^\gamma, v_0^\gamma) \in L^1$, then $\langle u_t^\gamma, 1\rangle$ and $\langle v_t^\gamma, 1\rangle$ are well-defined martingales which by nonnegativity of solutions are nonnegative. Also for $\gamma = \infty$, if $\rho < 0$ and additionally $(u_0^\infty, v_0^\infty) \in L^{1,E}$, the total mass processes $\langle u_t^\infty, 1\rangle$ and $\langle v_t^\infty, 1\rangle$ are nonnegative martingales by Theorem 1.2a). Hence, for $\gamma \in (0,\infty]$, by the martingale convergence theorem, $\langle u_t^\gamma, 1\rangle$ and $\langle v_t^\gamma, 1\rangle$ converge almost surely, as $t \to \infty$, and we denote by $\tilde{u}_\infty^\gamma, \tilde{v}_\infty^\gamma \in [0,\infty)$ their almost sure limits. Therefore,

$$\lim_{t\to\infty} \langle u_t^\gamma, 1\rangle \langle v_t^\gamma, 1\rangle = \tilde{u}_\infty^\gamma \tilde{v}_\infty^\gamma \in [0,\infty)$$

irrespectively of $\gamma$ being finite or infinite. In our two-type model this convergence is used to adapt the notion of existence/survival from one-type models.

**Definition 2.1.** Let $\gamma \in (0,\infty]$, then we say coexistence of types is possible for $SBM_\gamma$ if there are summable initial conditions $(u_0^\gamma, v_0^\gamma)$ such that $\tilde{u}_\infty^\gamma \tilde{v}_\infty^\gamma > 0$ with positive probability. Otherwise, we say that coexistence is impossible.

We define the Green function of $\mathcal{A}$ by

$$g_t(j,k) = \int_0^t p_s(j,k)\,ds,$$

where $p_s(j,k) = \mathbb{P}(\xi_s = k \,|\, \xi_0 = j)$ and $\xi$ is a continuous-time Markov process with generator $\mathcal{A}$. The next theorem was the starting point for the longtime analysis of two-type models that we consider in this chapter.

**Theorem 2.1 (Theorem 1.2 of [DP98]).** *Suppose $\gamma < \infty$ and $\rho = 0$, then the following dichotomy holds:*

(a) *Transient case: If $\sup_{k\in S} g_\infty(k,k) < \infty$, then coexistence of types is possible.*
(b) *Recurrent case: Assuming additionally the uniformity condition*

$$g_T(j,j) \geq C \sup_{k\in S} g_T(k,k), \quad \forall j \in S, T \geq T_0(j), \tag{8}$$

*on $\mathcal{A}$, a criterion for impossibility of coexistence of types is the following:*

$$\mathbb{P}^j(\xi_t = i \text{ for some } t \geq 0) = 1 \quad \forall j,i \in S.$$

The additional assumption (8) is fulfilled, for instance, for $\mathcal{A} = \Delta$ on $S = \mathbb{Z}^d$ so that coexistence in this case occurs if and only if $d \geq 3$.

*Remark 2.1.* The result of Theorem 2.1 has already been partially generalized to $\rho \neq 0$. For $\mathcal{A} = \Delta$ on $S = \mathbb{Z}^d$, the recurrent case (i.e., $d = 1,2$) was dealt with in the proof of Proposition 2.1 of [BDE11]. The proof tacitly uses the condition (8).

In the sequel, we will present a different approach that only works for $\rho < 0$ and proves the full coexistence/non-coexistence dichotomy for general $\mathcal{A}$. The additional assumption (8) on $\mathcal{A}$ is not necessary in this case.

Now let us switch to infinite rate processes. For $\rho = 0$ the longtime behavior was analyzed in [KM10a] based on the approach of [DP98] for $\gamma < \infty$. The proofs required more caution than the proofs for $\gamma < \infty$ as all second moment arguments had to be avoided since the infinite rate mutually catalytic branching process does not possess finite second moments. Such spatial systems are rare but interesting as their scaling properties change. The classical recurrence/transience dichotomy for finite variance models can break down in such a way that the criticality between survival and extinction is shifted to higher orders of the Green function (see, for instance, [DF85]). To extend Theorem 1.2 of [DP98], the log-Green function is needed:

$$g_{\infty,\log}(j,k) = \int_0^\infty p_s(j,k)(1 + |\log(p_s(j,k))|)\, ds.$$

Note that the log-Green function is infinite if the Green-function is infinite so that in the recurrent regime both are infinite.

**Theorem 2.2 (Theorem 1, Theorem 2 of [KM10a]).** *Suppose that* $\rho = 0$, $(u_0^\infty, v_0^\infty) \in L^{1,E}$ *and let* $\mathcal{A}$ *be such that*

$$\sup_{k\in S} |a(k,k)| < 1 \quad \textit{and} \quad \inf_{k\in S} |a(k,k)| > 0.$$

*(a) If* $g_{\infty,\log}(k,l)$ *is "small enough," then* $\langle u_\infty^\infty, 1\rangle\langle v_\infty^\infty, 1\rangle > 0$ *with positive probability for localized initial conditions, i.e., there are* $k,l \in S$ *such that*

$$\left(u_0^\infty, v_0^\infty\right)(i) = \begin{cases} (1,0) & : i = k, \\ (0,1) & : i = l, \\ (0,0) & : \textit{otherwise}. \end{cases}$$

*In particular, for log-Green function "small enough," coexistence of types occurs.*

*(b) The "recurrent" regime holds as in part (b) of Theorem 2.1.*

Unfortunately, the description of the recurrence/transience dichotomy is even less precise than in Theorem 2.1. It remains open what happens in the case when $g_\infty(k,l) < \infty$ and $g_{\infty,\log}(k,l) = \infty$. Again, the simple random walk serves as an example for which the results can be clarified. Suppose $\mathcal{A} = \Delta$, then coexistence is impossible if $d = 1,2$. For $d \geq 3$ one can use the local central limit theorem to show that $g_{\infty,\log}(k,j) \approx \|k-j\|^{2-d}$ as $\|k-j\| \to \infty$. Hence, the assumption of (a) is fulfilled if initially the two populations are sufficiently far apart.

## 2.2 Main Result

The main result of this chapter is a precise recurrence/transience dichotomy for $\rho < 0$ in both cases $\gamma < \infty$ and $\gamma = \infty$. The possibility of a dichotomy in terms of the Green function for $\gamma = \infty$ stems from the fact that for $\rho < 0$ the process has finite variance in contrast to infinite variance for $\rho = 0$.

**Theorem 2.3.** *Suppose $\rho < 0$ and $\gamma$ is either finite or infinite, then*

$$\textit{coexistence of types is impossible} \quad \Longleftrightarrow \quad \mathbb{P}^j(\xi_t = i \textit{ for some } t \geq 0) = 1 \quad \forall j, i \in S.$$

Interestingly, the negative correlations that seem to worsen the problem as symmetry gets lost simplify the problem a lot so that transparent proofs for this precise dichotomy are possible.

## 2.3 An Open Problem: Longtime Behavior for Positive Correlations

The reason we skip the longtime behavior for positive correlations is simple: the longtime behavior even for the finite branching rate processes $\mathrm{SBM}_\gamma$ is unknown for $\rho > 0$. Here we briefly describe what might be expected for $\mathcal{A} = \Delta$ and $S = \mathbb{Z}^d$ but we do not have proofs.

The particular case $\rho = 1$ corresponds to the classical parabolic Anderson problem (see Example 2). Having an explicit representation of the solution as Feynman–Kac functional this problem can be analyzed more easily. It is known that, started at localized initial conditions $w_0 = \mathbf{1}_{\{k\}}$, almost sure extinction $\lim_{t\to\infty}\langle w_t, 1\rangle = 0$ occurs if $d = 1, 2$ (the recurrent case). In [S92], see also [GdH07], it was proved that for $d \geq 3$ (the transient case) there are critical values $\gamma^1(d) > \gamma^2(d) > 0$ such that for $\gamma > \gamma^1(d)$ extinction occurs almost surely and for $\gamma < \gamma^2(d)$, survival occurs with positive probability. This is one example where a “more noise kills” effect can be proved rigorously.

For $\mathrm{SBM}_\gamma$ and $d \geq 3$, we conjecture the following: there should be a strictly decreasing critical curve $\gamma(\cdot, d) : (0, 1] \to \mathbb{R}^+$ coming down from infinity at zero and converging towards the critical threshold for the parabolic Anderson model at 1 such that coexistence of types for $\mathrm{SBM}_\gamma$ is impossible if $\gamma > \gamma(\rho, d)$ and coexistence is possible if $\gamma < \gamma(\rho, d)$. If the conjecture for $\gamma < \infty$ holds, it is furthermore natural to conjecture that for $\gamma = \infty$ coexistence of types is always impossible if $\rho > 0$.

# 3 Proofs

## *3.1 Some Properties of* $\mathrm{SBM}_\gamma$

To prepare for the proofs of the longtime behavior, we start with some lemmas for the finite branching rate processes $\mathrm{SBM}_\gamma$. In order to avoid confusions with the notion $\langle f,g\rangle$, we denote the quadratic variation of square-integrable martingales by $[\cdot,\cdot]$.

**Lemma 3.1.** *Suppose that* $\rho\in(-1,1)$, $\gamma\in(0,\infty)$ *and* $(u_0^\gamma,v_0^\gamma)\in L^1$, *then* $\langle u_t^\gamma,1\rangle$ *and* $\langle v_t^\gamma,1\rangle$ *are nonnegative martingales with cross variations*

$$[\langle u_\cdot^\gamma,1\rangle,\langle u_\cdot^\gamma,1\rangle]_t=[\langle v_\cdot^\gamma,1\rangle,\langle v_\cdot^\gamma,1\rangle]_t=\gamma\int_0^t\langle u_s^\gamma,v_s^\gamma\rangle\,ds$$

*and*

$$[\langle u_\cdot^\gamma,1\rangle,\langle v_\cdot^\gamma,1\rangle]_t=\rho\gamma\int_0^t\langle u_s^\gamma,v_s^\gamma\rangle\,ds.$$

*Proof.* We only sketch a proof as it is rather straightforward. The proof is based on the stochastic variation of constant representation for stochastic heat equations. For bounded test functions $\phi$, one can derive as in the proof of Theorem 2.2 of [DP98] that

$$\langle u_t^\gamma,\phi\rangle=\langle u_0^\gamma,P_t\phi\rangle+\sum_{k\in S}\int_0^t P_{t-s}\phi(k)\sqrt{\gamma u_s^\gamma(k)v_s^\gamma(k)}\,dB_s^1(k),$$

$$\langle v_t^\gamma,\phi\rangle=\langle v_0^\gamma,P_t\phi\rangle+\sum_{k\in S}\int_0^t P_{t-s}\phi(k)\sqrt{\gamma u_s^\gamma(k)v_s^\gamma(k)}\,dB_s^2(k),$$

where $P_tf(k)=\sum_{j\in S}p_t(k,j)f(j)$ and the Brownian motions are as in the definition of $\mathrm{SBM}_\gamma$. The infinite sums can be shown to converge in $L^2(\mathbb{P})$. Setting $\phi\equiv1$ shows that $\langle u_t^\gamma,1\rangle$ and $\langle v_t^\gamma,1\rangle$ are infinite sums of martingales and for any finite subset $\Gamma$ of $S$ the correlation structure of the stochastic integrals can be easily calculated. As the martingale property is conserved under $L^2(\mathbb{P})$-convergence and the cross variations converge, we obtain that the total mass processes are square-integrable martingales with cross variation

$$\begin{aligned}&\big[\langle u_\cdot^\gamma,1\rangle,\langle v_\cdot^\gamma,1\rangle\big]_t\\&=\lim_{|\Gamma|\to\infty}\left[\sum_{k\in\Gamma}\int_0^\cdot\sqrt{\gamma u_s^\gamma(k)v_s^\gamma(k)}\,dB_s^1(k),\sum_{j\in\Gamma}\int_0^\cdot\sqrt{\gamma u_s^\gamma(j)v_s^\gamma(j)}\,dB_s^2(j)\right]_t\\&=\lim_{|\Gamma|\to\infty}\sum_{j,k\in\Gamma}\int_0^t\sqrt{\gamma u_s^\gamma(k)v_s^\gamma(k)}\sqrt{\gamma u_s^\gamma(j)v_s^\gamma(j)}\,d\left[B_\cdot^1(k),B_\cdot^2(j)\right]_s\end{aligned}$$

which by the presumed correlation structure equals

$$\rho \lim_{|\Gamma|\to\infty} \sum_{k\in\Gamma} \int_0^t \gamma u_s^\gamma(k) v_s^\gamma(k)\, ds = \rho\gamma \int_0^t \langle u_s^\gamma, v_s^\gamma\rangle\, ds.$$

The derivation of the quadratic variations of $\langle u_t^\gamma, 1\rangle$ and $\langle v_t^\gamma, 1\rangle$ is similar, by dealing with the same sets of driving Brownian motions instead. □

The next lemma gives a refinement of Theorem 2.5 of [BDE11] uniformly in $\gamma$.

**Lemma 3.2.** *For any $\rho < 0$, there is an $\varepsilon > 0$ such that*

$$\sup_{\gamma<\infty, T>0} \mathbb{E}\big[\sup_{t\le T} \langle u_t^\gamma, 1\rangle^{2+\varepsilon}\big] < \infty,$$

$$\sup_{\gamma<\infty, T>0} \mathbb{E}\big[\sup_{t\le T} \langle v_t^\gamma, 1\rangle^{2+\varepsilon}\big] < \infty.$$

*Proof.* The cross-variation structure found for the martingales $\langle u_t^\gamma, 1\rangle, \langle v_t^\gamma, 1\rangle$ in the previous lemma allows us to obtain an upper bound for arbitrary (not only integer) moments by representing $(\langle u_t^\gamma, 1\rangle, \langle v_t^\gamma, 1\rangle)$ as a pair of time-changed correlated Brownian motions. A version of the Dubins–Schwartz theorem shows that

$$(W_t^1, W_t^2) := \big(\langle u_{A_t}^\gamma, 1\rangle, \langle v_{A_t}^\gamma, 1\rangle\big) \tag{9}$$

is a pair of correlated Brownian motions started in $(W_0^1, W_0^2) = (\langle u_0^\gamma, 1\rangle, \langle v_0^\gamma, 1\rangle)$ with covariance $\mathbb{E}[W_t^1 W_t^2] = \rho t$. Here, $A_t$ is the generalized inverse of $[\langle v_\cdot^\gamma, 1\rangle, \langle v_\cdot^\gamma, 1\rangle]_t = [\langle u_\cdot^\gamma, 1\rangle, \langle u_\cdot^\gamma, 1\rangle]_t$. Furthermore, as the total masses are nonnegative, the time-change $A_t$ is bounded by $\tau$, the first hitting time of $(W_t^1, W_t^2)$ at the boundary of the first quadrant (otherwise one of the total masses would become negative). Applying the Burkholder–Davis–Gundy inequality, this shows that

$$\sup_{\gamma<\infty, T>0} \mathbb{E}\big[\sup_{t\le T} \langle u_t^\gamma, 1\rangle^p\big] \le cE\Big[\tau^{p/2} \,\Big|\, W_0^1 = \langle u_0^\gamma, 1\rangle, W_0^2 = \langle v_0^\gamma, 1\rangle\Big],$$

which is finite if $p = 2 + \varepsilon$ and $\varepsilon$ is sufficiently small. For independent Brownian motions the number of finite moments of the first exit time $\tau$ has been determined by Spitzer [S58] (see also Burkholder [B77]); for the simple transformation to correlated Brownian motions, we refer to Theorem 2.5 of [BDE11]. □

The final tool is a second moment formula for $\mathrm{SBM}_\gamma$ for which we introduce the abbreviation $L_t = \int_0^t \mathbf{1}_{\{\xi_s^1 = \xi_s^2\}}\, ds$ for the collision time of two independent continuous-time Markov processes $\xi^1, \xi^2$ with generator $\mathcal{A}$.

**Lemma 3.3.** *If $L_t$ denotes the collision time, then the second moment formula*

$$\mathbb{E}[u_t^\gamma(a) v_t^\gamma(b)] = \mathbb{E}^{a,b}\big[u_0^\gamma(\xi_t^1) v_0^\gamma(\xi_t^2) e^{\rho\gamma L_t}\big], \quad a, b \in S, \tag{10}$$

*holds.*

*Proof.* There are several ways to see this expression. For instance, this follows as a particularly simple application of the moment duality for $\mathrm{SBM}_\gamma$ (derived in Proposition 9 of [EF04]; see also the explanation in Lemma 3.3 of [BDE11]). □

Formula (10) shows the significant difference of negative, null or positive correlations: only in the case of negative correlations one can hope to have finite second moment for the $\gamma = \infty$ limiting process.

With these preparations, we can proceed with the proof of Theorem 1.2.

## 3.2 Proof of Theorem 1.2

Let $\gamma_n \to \infty$, then we know from Theorem 1.1 that

$$(u^{\gamma_n}, v^{\gamma_n}) \Rightarrow (u^\infty, v^\infty)$$

in $D_{L^\beta}$ in the Meyer–Zheng pseudo-path topology.

*Proof of (a), (b):* First, we would like to show that the limiting total mass processes $\langle u^\infty_\cdot, 1\rangle$ and $\langle v^\infty_\cdot, 1\rangle$ are martingales. One should be a little bit careful, since convergence of $(u^{\gamma_n}, v^{\gamma_n})$ is in $D_{L^\beta}$ and not in $D_{L^1}$, and hence not necessarily $(\langle u^{\gamma_n}_\cdot, 1\rangle, \langle v^{\gamma_n}_\cdot, 1\rangle)$ converges to $(\langle u^\infty_\cdot, 1\rangle, \langle v^\infty_\cdot, 1\rangle)$. However, without loss of generality we may and will assume that for our chosen subsequence $\{(u^{\gamma_n}, v^{\gamma_n})\}_{n\geq 1}$, at least,

$$\left(u^{\gamma_n}_\cdot, v^{\gamma_n}_\cdot, \langle u^{\gamma_n}_\cdot, 1\rangle, \langle v^{\gamma_n}_\cdot, 1\rangle\right) \Rightarrow \left(u^\infty_\cdot, v^\infty_\cdot, \bar{u}^\infty_\cdot, \bar{v}^\infty_\cdot\right) \tag{11}$$

in $D_{L^\beta} \times D_{\mathbb{R}} \times D_{\mathbb{R}}$ in the Meyer–Zheng pseudo-path topology. Let us show that indeed

$$\bar{u}^\infty_t = \langle u^\infty_t, 1\rangle, \quad \bar{v}^\infty_t = \langle v^\infty_t, 1\rangle, \ \forall t \geq 0. \tag{12}$$

By Theorem 5 of [MZ84] the convergence in the Meyer–Zheng topology implies convergence (along a further subsequence which, again, we denote by $\gamma_n$ for convenience) of one-dimensional distributions on a set of full Lebesgue measure, say $\mathbb{T}$. Fix any $t \in \mathbb{T}$. Since $u_0$ and $v_0$ are summable, for any $\varepsilon > 0$, one can fix a sufficiently large compact $\Gamma \subset S$ such that

$$\mathbb{E}\left[\langle u^{\gamma_n}_t, \mathbf{1}_{\Gamma^c}\rangle + \langle v^{\gamma_n}_t, \mathbf{1}_{\Gamma^c}\rangle\right] = \mathbb{E}\left[\langle u^{\gamma_n}_0, P_t\mathbf{1}_{\Gamma^c}\rangle + \langle v^{\gamma_n}_0, P_t\mathbf{1}_{\Gamma^c}\rangle\right] \leq \varepsilon, \quad \forall \gamma_n > 0.$$

For the equality we have used the fact $\mathbb{E}[\langle u^\gamma_t, \phi\rangle] = \langle u^\gamma_0, P_t\phi\rangle$ which follows from the stochastic variation of constant representation utilized in the proof of Lemma 3.1. This implies that, in fact, $(u^{\gamma_n}_t, u^{\gamma_n}_t) \Rightarrow (u^\infty_t, u^\infty_t)$ in $M_F(S) \times M_F(S)$—the product space of finite measures on $S$ equipped with the weak topology. This immediately implies

$$(\langle u^{\gamma_n}_t, 1\rangle, \langle v^{\gamma_n}_t, 1\rangle) \Rightarrow (\langle u^\infty_t, 1\rangle, \langle v^\infty_t, 1\rangle), \ \forall t \in \mathbb{T}, \tag{13}$$

and combined with (11) this gives

$$(\bar{u}_t^\infty, \bar{v}_t^\infty) = (\langle u_t^\infty, 1\rangle, \langle v_t^\infty, 1\rangle), \ \ \forall t \in \mathbb{T}.$$

Since both processes, on the right- and on the left-hand sides, are right-continuous, we get that in fact the equality holds for all $t$, and hence (12) follows.

By the above Theorem 11 of [MZ84] and Lemma 3.2 we immediately get that the limiting total mass processes $\langle u_t^\infty, 1\rangle, \langle v_t^\infty, 1\rangle$, $t \geq 0$, are martingales. We will get the square integrability of these martingales (and thus finish the proof of a)) when we prove b). For b) we use Fatou's lemma combined with (13) and Lemma 3.2 to obtain that

$$\sup_{t\in\mathbb{T}} \mathbb{E}\left[\langle u_t^\infty, 1\rangle^{2+\varepsilon} + \langle v_t^\infty, 1\rangle^{2+\varepsilon}\right] < \infty.$$

Recalling that $t \mapsto (\langle u_t^\infty, 1\rangle, \langle v_t^\infty, 1\rangle)$ is right-continuous, one can take the supremum over $t \in \mathbb{R}^+$, and b) follows.

*Proof of (c):* The representation is first deduced for $t \in \mathbb{T}$. By the choice of $\mathbb{T}$ and (11), we get

$$u_t^{\gamma_n}(a) v_t^{\gamma_n}(b) \Rightarrow u_t^\infty(a) v_t^\infty(b). \tag{14}$$

To get convergence of the first moments of $\{u_t^{\gamma_n}(a) v_t^{\gamma_n}(b)\}_{n\geq 1}$, it is enough to check the uniform integrability of $\{u_t^{\gamma_n}(a) v_t^{\gamma_n}(b)\}_{n\geq 1}$. However this follows from Lemma 3.2 and Hölder's inequality:

$$\sup_{n\geq 1} \mathbb{E}\left[\left(u_t^{\gamma_n}(a) v_t^{\gamma_n}(b)\right)^{\frac{2+\varepsilon}{2}}\right] \leq \sup_{n\geq 1} \sqrt{\mathbb{E}\left[\langle u_t^{\gamma_n}, 1\rangle^{2+\varepsilon}\right] \mathbb{E}\left[\langle v_t^{\gamma_n}, 1\rangle^{2+\varepsilon}\right]} < \infty,$$

if $\varepsilon$ is chosen sufficiently small. The uniform integrability, Lemma 3.3 and the dominated convergence then imply

$$\begin{aligned}
\mathbb{E}[u_t^\infty(a) v_t^\infty(b)] &= \lim_{n\to\infty} \mathbb{E}[u_t^{\gamma_n}(a) v_t^{\gamma_n}(b)] \\
&= \lim_{n\to\infty} \mathbb{E}^{a,b}\left[u_0^{\gamma_n}(\xi_t^1) v_0^{\gamma_n}(\xi_t^2) e^{\rho\gamma_n L_t}\right] \\
&= \mathbb{E}^{a,b}\left[u_0^\infty(\xi_t^1) v_0^\infty(\xi_t^2) \mathbf{1}_{\{L_t=0\}}\right] + \mathbb{E}^{a,b}\left[u_0^\infty(\xi_t^1) v_0^\infty(\xi_t^2) e^{\lim_{n\to\infty}\gamma_n \rho L_t} \mathbf{1}_{\{L_t>0\}}\right] \\
&= \mathbb{E}^{a,b}\left[u_0^\infty(\xi_t^1) v_0^\infty(\xi_t^2) \mathbf{1}_{\{L_t=0\}}\right],
\end{aligned}$$

where we also used that by definition $u_0^{\gamma_n} = u_0^\infty$ and $v_0^{\gamma_n} = v_0^\infty$.

## 3.3 *Proof of Theorem 2.3*

With the preparations of Sect. 3.1 we can now prove the extinction/coextinction dichotomy.

### 3.3.1 Proof of Theorem 2.3, $\gamma < \infty$

Recall that due to the martingale convergence theorem the product $M_t^\gamma := \langle u_t^\gamma, 1\rangle\langle v_t^\gamma, 1\rangle$ of the two nonnegative martingales converges almost surely to $\tilde{u}_\infty^\gamma \tilde{v}_\infty^\gamma$, and our task is to determine when the limit equals zero almost surely. As the limit is nonnegative, the most straightforward approach is to deduce a formula for $\mathbb{E}\left[M_\infty^\gamma\right]$ and to determine when this is strictly positive or null. For this to work, the assumption $\rho < 0$ is crucial.

Luckily, using the results from Sect. 3.1, the convergence of $\mathbb{E}[M_t^\gamma]$ to $\mathbb{E}[M_\infty^\gamma]$ comes almost for granted. The convergence of $M_t^\gamma$ holds almost surely, so it suffices to show uniform integrability in $t$. Choosing $\varepsilon$ small enough, we obtain the uniform integrability from Hölder's inequality and Lemma 3.2:

$$\mathbb{E}\left[(M_t^\gamma)^{\frac{2+\varepsilon}{2}}\right] \le \sqrt{\mathbb{E}[\langle u_t^\gamma, 1\rangle^{2+\varepsilon}]\mathbb{E}[\langle v_t^\gamma, 1\rangle^{2+\varepsilon}]} \le C.$$

Combined with nonnegativity of $\tilde{u}_\infty^\gamma \tilde{v}_\infty^\gamma$ this implies that

$$\tilde{u}_\infty^\gamma \tilde{v}_\infty^\gamma = 0 \quad a.s. \quad \Longleftrightarrow \quad \lim_{t\to\infty} \mathbb{E}[M_t^\gamma] = 0.$$

The first moment of $M_t^\gamma$ can be calculated from the moment duality for SBM$_\gamma$ in such a way that the criterion of the dichotomy directly drops out; to finish the proof, we aim at using Lemma 3.3 to show that

$$\lim_{t\to\infty} \mathbb{E}[M_t^\gamma] = 0 \text{ for all } (u_0^\gamma, v_0^\gamma) \in L^1 \quad \Longleftrightarrow \quad \mathbb{P}^{i,j}(\xi_t^1 = \xi_t^2 \text{ for some } t \ge 0) = 1 \quad \forall i, j \in S.$$

Taking into account Lemma 3.3, we get, for any $(u_0^\gamma, v_0^\gamma) \in L^1$ with $\langle u_0^\gamma, 1\rangle\langle v_0^\gamma, 1\rangle > 0$,

$$\begin{aligned}
\mathbb{E}[M_t^\gamma] &= \sum_{a,b\in S} \mathbb{E}[u_t^\gamma(a) v_t^\gamma(b)] \\
&= \sum_{a,b\in S} \mathbb{E}^{a,b}\left[u_0^\gamma(\xi_t^1) v_0^\gamma(\xi_t^2) e^{\gamma\rho L_t}\right] \\
&= \sum_{a,b\in S} \sum_{i,j\in S} u_0^\gamma(i) v_0^\gamma(j) \mathbb{E}^{a,b}\left[\mathbf{1}_i(\xi_t^1)\mathbf{1}_j(\xi_t^2) e^{\gamma\rho L_t}\right]
\end{aligned}$$

$$= \sum_{i,j \in S} \sum_{a,b \in S} u_0^\gamma(i) v_0^\gamma(j) \mathbb{E}^{i,j}\left[\mathbf{1}_a(\xi_t^1)\mathbf{1}_b(\xi_t^2) e^{\gamma\rho L_t}\right]$$
$$= \sum_{i,j \in S} u_0^\gamma(i) v_0^\gamma(j) \mathbb{E}^{i,j}\left[e^{\gamma\rho L_t}\right].$$

For the fourth equation we used reversibility of the paths of $\xi^1, \xi^2$ since we assumed $\mathcal{A}$ to be symmetric with bounded jump rate. As $\rho$ is negative, by the dominated convergence, and taking into account $\sum_{i,j} u_0^\gamma(i) v_0^\gamma(j) \in (0,\infty)$, we see that the right-hand side converges to zero if and only if $\lim_{t\to\infty} \mathbb{E}^{i,j}\left[e^{\gamma\rho L_t}\right] = 0$ for all $i, j \in S$. The proof can now be finished via the dominated convergence again and the Markov property:

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}^{i,j}\left[e^{\gamma\rho L_t}\right] = 0 \quad \forall i,j \in S \quad &\iff \quad \mathbb{P}^{i,j}(L_t \to \infty) = 1 \quad \forall i,j \in S \\ &\iff \quad \mathbb{P}^{i,j}(\xi_t^1 = \xi_t^2 \text{ for some } t \geq 0) = 1 \\ &\qquad\quad \forall i,j \in S.\end{aligned}$$

### 3.3.2 Proof of Theorem 2.3, $\gamma = \infty$

The proof of the dichotomy on the level of infinite branching rate can now be deduced similarly to the $\gamma < \infty$ case for which we define

$$M_t^\infty := \langle u_t^\infty, 1\rangle \langle v_t^\infty, 1\rangle$$

for $t \geq 0$.

**Lemma 3.4.** *Suppose $\rho < 0$ and for the initial conditions $(u_0^\infty, v_0^\infty) \in L^{1,E}$, then*

$$\mathbb{E}[M_t^\infty] = \sum_{i,j \in S, i \neq j} u_0^\infty(i) v_0^\infty(j) \mathbb{P}^{i,j}(\xi_s^1 \neq \xi_s^2, \forall s \leq t) \tag{15}$$

*for $t \geq 0$.*

*Proof.* We use Theorem 1.2(c), Fubini's theorem and the time-reversion trick used in the proof of Theorem 2.3 for $\gamma < \infty$:

$$\mathbb{E}[M_t^\infty] = \sum_{a,b \in S} \mathbb{E}[u_t^\infty(a) v_t^\infty(b)]$$
$$= \sum_{a,b \in S} \sum_{i,j \in S} \mathbb{E}^{a,b}\left[u_0^\infty(i)\mathbf{1}_i(\xi_t^1) v_0^\infty(j) \mathbf{1}_j(\xi_t^2) \mathbf{1}_{L_t = 0}\right]$$

$$= \sum_{i,j\in S}\sum_{a,b\in S} u_0^\infty(i)v_0^\infty(j)\mathbb{E}^{i,j}\left[\mathbf{1}_a(\xi_t^1)\mathbf{1}_b(\xi_t^2)\mathbf{1}_{L_t=0}\right]$$
$$= \sum_{i,j\in S} u_0^\infty(i)v_0^\infty(j)\mathbb{P}^{i,j}\left[\xi_s^1 \neq \xi_s^2, \forall s\leq t\right].$$

By assumption $(u_0^\infty, v_0^\infty)\in E^S$ so that the terms with $i=j$ vanish as then $u_0^\infty(i)v_0^\infty(i)=0$. □

Now we have to verify necessary and sufficient conditions for the limit

$$\lim_{t\to\infty}\langle u_t^\infty,1\rangle\langle v_t^\infty,1\rangle = \tilde{u}_\infty^\infty \tilde{v}_\infty^\infty$$

being equal to zero almost surely. This can be done similarly as in the case $\gamma<\infty$.

**Lemma 3.5.** *Suppose $\rho<0$ and for the initial conditions $(u_0^\infty, v_0^\infty)\in L^{1,E}$, then*

$$\tilde{u}_\infty^\infty \tilde{v}_\infty^\infty = 0 \quad a.s. \qquad \Longleftrightarrow \qquad \lim_{t\to\infty}\mathbb{E}[M_t^\infty]=0.$$

*Proof.* Almost sure convergence of $M_t^\infty$ to $\tilde{u}_\infty^\infty \tilde{v}_\infty^\infty$ is due to the martingale convergence theorem, so that it suffices to show that $M_t^\infty$ is uniformly integrable in $t$. By Hölder's inequality, we reduce the question to the total mass processes:

$$\mathbb{E}\left[(M_t^\infty)^{\frac{2+\varepsilon}{2}}\right] \leq \sqrt{\mathbb{E}\left[\langle u_t^\infty,1\rangle^{2+\varepsilon}\right]\mathbb{E}\left[\langle v_t^\infty,1\rangle^{2+\varepsilon}\right]}, \tag{16}$$

and the result follows from Theorem 1.2(c) if $\varepsilon$ is chosen small enough. □

Now we can proceed with the proof of Theorem 2.3 for $\gamma=\infty$ exactly as for $\gamma<\infty$.

*Proof of Theorem 2.3, $\gamma=\infty$.* The theorem follows directly from Lemma 3.4 and Lemma 3.5. Using the dominated convergence, justified by $\sum_{i,j\in S}u_0^\infty(i)v_0^\infty(j)<\infty$, and monotonicity of measures we obtain

$$\lim_{t\to\infty}\sum_{i,j\in S, i\neq j} u_0^\infty(i)v_0^\infty(j)\mathbb{P}^{i,j}(\xi_s^i\neq\xi_s^2,\forall s\leq t)=0, \quad \forall (u_0^\infty,v_0^\infty)\in L^{1,E}$$
$$\Longleftrightarrow \quad \mathbb{P}^{i,j}(\xi_t^1=\xi_t^2 \text{ for some } t\geq 0)=1 \quad \forall j,i\in S, i\neq j.$$

This finishes the proof of Theorem 2.3 also for infinite branching rate. □

**Acknowledgements** The authors thank an anonymous referee for a very careful reading of the manuscript.

# References


BDE11. J. Blath, L. Döring, A. Etheridge "On the Moments and the Wavespeed of the Symbiotic Branching Model." The Annals of Probability, 2011, Vol. 39, No. 1, pp. 252–290.

B77. D. L. Burkholder. Exit times of Brownian motion, harmonic majorization, and Hardy spaces. *Advances in Math.*, 26(2):182–205, 1977.

CDG04. J.T. Cox, D. Dawson, A. Greven "Mutually Catalytic Super branching Random Walks: Large Finite Systems and Renormalization Analysis." AMS Memoirs, 2004, Vol. 171.

CK00. J.T. Cox, A. Klenke "Recurrence and Ergodicity of Interacting Particle Systems." Probab. Theory Relat. Fields, 2000, Vol. 116, pp. 239–255.

DF85. D. Dawson, K. Fleischmann "Critical dimension for a model of branching in a random medium." Z. Wahrsch. Verw. Gebiete, 70(3), pp. 315–334, 1985.

DM11. L. Döring, L. Mytnik "Mutually Catalytic Branching Processes on the lattice and Voter Processes with Strength of Opinion." arXiv:1109.6106v1.

DP98. D. Dawson, E. Perkins "Long-Time Behavior and Coexistence in a Mutually Catalytic branching model." The Annals of Probability, 1998, Vol. 26, No.3. pp. 1088–1138.

E00. A. Etheridge "An Introduction to Superprocesses." University Lecture Series, Amer. Math. Soc., 2000.

EF04. A. Etheridge, K. Fleischmann "Compact interface property for symbiotic branching." Stochastic Processes and their Applications, 2004, Vol. 114, pp. 127–160.

GdH07. A. Greven, F. den Hollander "Phase transitions for the longtime behavior of interacting diffusions." The Annals of Probability, 2007, Vol. 35, No.4., pp. 1250–1306.

J97. A. Jakubowski "A non-Skorohod topology on the Skorohod space." Electronic Journal of Probability, 1997, Vol. 2, No.4., pp. 1–21.

KM10a. A. Klenke, L. Mytnik "Infinite Rate Mutually Catalytic Branching in Infinitely Many Colonies: The Longtime Behaviour." to appear in Annals of Probability.

KM10b. A. Klenke, L. Mytnik "Infinite Rate Mutually Catalytic Branching in Infinitely Many Colonies: Construction, Characterization and Convergence." to appear in Probab. Theory Relat. Fields.

KO10. A. Klenke, M. Oeler "A Trotter Type Approach to Infinite Rate Mutually Catalytic Branching." The Annals of Probability, 2010, Vol. 38, 479–497.

L05. T. Liggett "Interacting Particle Systems." Reprint of the 1985 original. Classics in Mathematics. Springer-Verlag, Berlin, 2005.

MZ84. P.A. Meyer, W.A. Zheng "Tightness criteria for laws of semimartingales." Ann. Inst. H. Poincaré Probab. Statist. 20 (1984), no. 4, 353–372.

S92. T. Shiga "Ergodic Theorems and Exponential Decay of Sample Paths for Certain Interacting Diffusion Systems." Osaka J. Math., 1992, Vol. 29, p. 789–807.

SS80. T. Shiga, A. Shimizu "Infinite dimensional stochastic differential equations and their applications." J. Math. Kyoto Univ., 1980, Vol. 20, No.3., pp. 395–416.

S58. F. Spitzer "Some theorems concerning 2-dimensional Brownian motion." Trans. Amer. Math. Soc., 1958, No. 87, pp. 187–197.

# Super-Brownian Motion: $L^p$-Convergence of Martingales Through the Pathwise Spine Decomposition

**A.E. Kyprianou and A. Murillo-Salas**

**Abstract** Evans [7] described the semigroup of a superprocess with quadratic branching mechanism under a martingale change of measure in terms of the semigroup of an immortal particle and the semigroup of the superprocess prior to the change of measure. This result, commonly referred to as the spine decomposition, alludes to a pathwise decomposition in which independent copies of the original process "immigrate" along the path of the immortal particle. For branching particle diffusions, the analogue of this decomposition has already been demonstrated in the pathwise sense; see, for example, [10, 11]. The purpose of this short note is to exemplify a new *pathwise* spine decomposition for supercritical super-Brownian motion with general branching mechanism (cf. [13]) by studying $L^p$-convergence of naturally underlying additive martingales in the spirit of analogous arguments for branching particle diffusions due to Harris and Hardy [10]. Amongst other ingredients, the Dynkin–Kuznetsov $\mathbb{N}$-measure plays a pivotal role in the analysis.

**Keywords** Super-Brownian motion • Additive martingales • $\mathbb{N}$-measure • Spine decomposition • $L^p$-convergence

**MSC subject classifications (2010):** 60J68, 60F25.

---

On the occasion of the 60th birthday of Sergei Kuznetsov

A.E. Kyprianou (✉)
Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom
e-mail: a.kyprianou@bath.ac.uk

A. Murillo-Salas
Departamento de Matemáticas, Universidad de Guanajuato, Jalisco S/N Mineral de Valenciana, Guanajuato, Gto. C.P. 36240, México
e-mail: amurillos@ugto.mx

J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8_7, 

# 1 Introduction

Suppose that $X = \{X_t : t \geq 0\}$ is a (one-dimensional) $\psi$-super-Brownian motion with general branching mechanism $\psi$ taking the form

$$\psi(\lambda) = -\alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (\mathrm{e}^{-\lambda x} - 1 + \lambda x)\nu(\mathrm{d}x), \tag{1}$$

for $\lambda \geq 0$ where $\alpha = -\psi'(0+) \in (0,\infty)$, $\beta \geq 0$ and $\nu$ is a measure concentrated on $(0,\infty)$ which satisfies $\int_{(0,\infty)}(x \wedge x^2)\nu(\mathrm{d}x) < \infty$. Let $\mathcal{M}_F(\mathbb{R})$ be the space of finite measures on $\mathbb{R}$ and note that $X$ is a $\mathcal{M}_F(\mathbb{R})$-valued Markov process under $\mathbb{P}_\mu$ for each $\mu \in \mathcal{M}_F(\mathbb{R})$, where $\mathbb{P}_\mu$ is the law of $X$ with initial configuration $\mu$. We shall use standard inner product notation, for $f \in C_b^+(\mathbb{R})$, the space of positive, uniformly bounded, continuous functions on $\mathbb{R}$, and $\mu \in \mathcal{M}_F(\mathbb{R})$,

$$\langle f, \mu \rangle = \int_{\mathbb{R}} f(x)\mu(\mathrm{d}x).$$

Accordingly we shall write $||\mu|| = \langle 1, \mu \rangle$. Recall that the total mass of the process $X$, $\{||X_t|| : t \geq 0\}$ is a continuous-state branching process with branching mechanism $\psi$. Such processes may exhibit explosive behaviour; however, under the conditions assumed above, $||X||$ remains finite at all times. We insist moreover that $\psi(\infty) = \infty$ which means that with positive probability the event $\lim_{t\uparrow\infty} ||X_t|| = 0$ will occur. Equivalently this means that the total mass process does not have monotone increasing paths; see, for example, the summary in Chap. 10 of Kyprianou [12]. The existence of these superprocesses is guaranteed by [1, 3, 4].

The following standard result from the theory of superprocesses describes the evolution of $X$ as a Markov process. For all $f \in C_b^+(\mathbb{R})$ and $\mu \in \mathcal{M}_F(\mathbb{R})$,

$$-\log \mathbb{E}_\mu(\mathrm{e}^{-\langle f, X_t \rangle}) = \int_{\mathbb{R}} u_f(x,t)\mu(\mathrm{d}x),\ \mu \in \mathcal{M}_F(\mathbb{R}), t \geq 0, \tag{2}$$

where $u_f(x,t)$ is the unique positive solution to the evolution equation for $x \in \mathbb{R}$ and $t > 0$

$$\frac{\partial}{\partial t} u_f(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2} u_f(x,t) - \psi(u_f(x,t)), \tag{3}$$

with initial condition $u_f(x,0) = f(x)$. The reader is referred to Theorem 1.1 of Dynkin [2], Proposition 2.3 of Fitzsimmons [8] and Proposition 2.2 of Watanabe [15] for further details; see also Dynkin [3, 4] and Engländer and Pinsky [6] for a general overview.

Associated to the process $X$ is the following martingale $Z(\lambda) = \{Z_t(\lambda), t \geq 0\}$, where

$$Z_t(\lambda) := \mathrm{e}^{\lambda c_\lambda t}\langle \mathrm{e}^{\lambda \cdot}, X_t \rangle, t \geq 0, \tag{4}$$

where $c_\lambda = \psi'(0+)/\lambda - \lambda/2$ and $\lambda \in \mathbb{R}$ (cf. [13] Lemma 2.2). To see why this is a martingale, note the following steps. Define for each $x \in \mathbb{R}$, $g \in C_b^+(\mathbb{R})$ and $\theta, t \geq 0$, $u_g^\theta(x,t) = -\log \mathbb{E}_{\delta_x}(\mathrm{e}^{-\theta\langle g, X_t\rangle})$. With limits understood as $\theta \downarrow 0$, we have $u_g(x,t)|_{\theta=0} = 0$; moreover $v_g(x,t) := \mathbb{E}_{\delta_x}(\langle g, X_t\rangle) = \partial u_g^\theta(x,t)/\partial\theta|_{\theta=0}$. Differentiating in $\theta$ in (3) shows that $v_g$ solves the equation

$$\frac{\partial}{\partial t} v_g(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2} v_g(x,t) - \psi'(0+) v_g(x,t), \tag{5}$$

with $v_g(x,0) = g(x)$. Classical Feynman–Kac theory tells us that (5) has a unique solution which is necessarily equal to $\Pi_x(\mathrm{e}^{-\psi'(0+)t} g(\xi_t))$ where $\{\xi_t : t \geq 0\}$ is a Brownian motion issued from $x \in \mathbb{R}$ under the measure $\Pi_x$. The above procedure also works for $g(x) = \mathrm{e}^{\lambda x}$ in which case we easily conclude that for all $x \in \mathbb{R}$ and $t \geq 0$, $\mathrm{e}^{\lambda c_\lambda t}\mathbb{E}_{\delta_x}(\langle \mathrm{e}^{\lambda \cdot}, X_t\rangle) = \mathrm{e}^{\lambda x}$. Finally, the martingale property follows using the previous equality together with the Markov branching property associated with $X$. Note that as a positive martingale, it is automatic that the limit

$$Z_\infty(\lambda) := \lim_{t\uparrow\infty} Z_t(\lambda)$$

exists $\mathbb{P}_\mu$ almost surely for all $\mu \in \mathcal{M}_F(\mathbb{R})$ such that $\langle \mathrm{e}^{\lambda \cdot}, \mu\rangle < \infty$.

The purpose of this note is to demonstrate the robustness of a new path decomposition of our $\psi$-super-Brownian motion by studying the $L^p$-convergence of the martingales $Z(\lambda)$. Specifically we shall prove the following theorem.

**Theorem 1.1.** *Assume that $p \in (1,2]$, $\int_{(0,\infty)} r^p \nu(\mathrm{d}r) < \infty$ and $p\lambda^2 < -2\psi'(0+)$. Then $Z_t(\lambda)$ converges to $Z_\infty(\lambda)$ in $L^p(\mathbb{P}_\mu)$, for all $\mu \in \mathcal{M}_F(\mathbb{R})$ such that $\langle \mathrm{e}^{\lambda p \cdot}, \mu\rangle$ are finite.*

The method of proof we use is quite similar to the one used in Harris and Hardy [10] for branching Brownian motion, where a pathwise spine decomposition functions as the key instrument of the proof. Roughly speaking, in that setting, the spine decomposition says that under a change of measure, the law of the branching Brownian motion has the same law as an immortal Brownian diffusion (with drift) along the path of which independent copies of the original branching Brownian motion immigrate at times which form a Poisson process. Until recently, such a spine decomposition for superdiffusions was only available in the literature in a weak form, meaning that it takes the form of a semigroup decomposition. See the original paper of Evans [7] as well as, for example, amongst others, Engländer and Kyprianou [9]. Recently, however, Kyprianou et al. [13] give a pathwise spine decomposition which provides a natural analogue to the pathwise spine decomposition for branching Brownian motion. Amongst other ingredients, the Dynkin–Kuznetsov $\mathbb{N}$-measure plays a pivotal role in describing the immigration off an immortal particle. We give a description of this new spine decomposition in the next section and thereafter we proceed to the proof of Theorem 1.1 in Sect. 3.

## 2 Spine Decomposition

For each $\lambda \in \mathbb{R}$ and $\mu \in \mathcal{M}_F(\mathbb{R})$ satisfying $\langle e^{\lambda \cdot}, \mu \rangle < \infty$, we introduce the following martingale change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}_\mu^\lambda}{\mathrm{d}\mathbb{P}_\mu}\right|_{\mathcal{F}_t} = \frac{Z_t(\lambda)}{\langle e^{\lambda \cdot}, \mu \rangle}, t \geq 0, \tag{6}$$

where $\mathcal{F}_t := \sigma(X_s, s \leq t)$. The preceding change of measure induces the *spine decomposition* of $X$ alluded to above. To describe it in detail, we need some more ingredients.

According to Dynkin and Kuznetsov [5], there exists a collection of measures $\{\mathbb{N}_x, x \in \mathbb{R}\}$, defined on the same probability space as $X$, such that

$$\mathbb{N}_x\left(1 - \mathrm{e}^{-\langle f, X_t \rangle}\right) = u_f(x,t),\ x \in \mathbb{R}, t \geq 0. \tag{7}$$

Roughly speaking, the branching property tells us that for each $n \in \mathbb{N}$, the measures $\mathbb{P}_{\delta_x}$ can be written as the $n$-fold convolution of $\mathbb{P}_{\frac{1}{n}\delta_x}$ which indicates that, on the trajectory space of the superprocess, $\mathbb{P}_x$ is infinitely divisible. Hence, the role of $\mathbb{N}_x$ in (7) is analogous to that of the Lévy measure for positive real-valued random variables.

From identity (7) and Eq. (2), it is straightforward to deduce that

$$\mathbb{N}_x(\langle f, X_t \rangle) = \mathbb{E}_{\delta_x}(\langle f, X_t \rangle), \tag{8}$$

whenever $f \in C_b^+(\mathbb{R})$.

For each $x \in \mathbb{R}$, let $\Pi_x$ be the law of a Brownian motion $\xi := \{\xi_t : t \geq 0\}$ issued from $x$. If $\Pi_x^\lambda$ is the law under which $\xi$ is a Brownian motion with drift $\lambda \in \mathbb{R}$ and issued from $x \in \mathbb{R}$, then for each $t \geq 0$,

$$\left.\frac{\mathrm{d}\Pi_x^\lambda}{\mathrm{d}\Pi_x}\right|_{\mathcal{G}_t} = \mathrm{e}^{\lambda(\xi_t - x) - \frac{1}{2}\lambda^2 t}, t \geq 0, \tag{9}$$

where $\mathcal{G}_t := \sigma(\xi_s, s \leq t)$. For convenience, we shall also introduce the measure

$$\Pi_\mu^\lambda(\cdot) := \frac{1}{\langle \mathrm{e}^{\lambda \cdot}, \mu \rangle} \int \mathrm{e}^{\lambda x} \mu(\mathrm{d}x) \Pi_x^\lambda(\cdot), \tag{10}$$

for all $\lambda \in \mathbb{R}$. In other words, $\Pi_\mu^\lambda$ has the law of a Brownian motion with drift at rate $\lambda$ with an initial position which has been independently randomised in a way that is determined by $\mu$.

Now fix $\mu \in \mathcal{M}_F(\mathbb{R})$ and $x \in \mathbb{R}$ and let us define a measure-valued process $\Lambda := \{\Lambda_t, t \geq 0\}$ as follows:

(i) Take a copy of the process $\xi = \{\xi_t, t \geq 0\}$ under $\Pi_x^\lambda$; we shall refer to this process as the *spine*.
(ii) Suppose that $\mathbf{n}$ is a Poisson point process such that, for $t \geq 0$, given the spine $\xi$, $\mathbf{n}$ issues superprocess $X^{\mathbf{n},t}$ at space-time position $(\xi_t, t)$ with rate $\mathrm{d}t \times 2\beta \mathrm{d}\mathbb{N}_{\xi_t}$.
(iii) Suppose that $\mathbf{m}$ is a Poisson point process such that, independently of $\mathbf{n}$, given the spine $\xi$, $\mathbf{m}$ issues a superprocess $X^{\mathbf{m},t}$ at space-time point $(\xi_t, t)$ with initial mass $r$ at rate $\mathrm{d}t \times r\nu(\mathrm{d}r) \times \mathrm{d}\mathbb{P}_{r\delta_{\xi_t}}$.

Note in particular that, when $\beta > 0$, the rate of immigration under the process $\mathbf{n}$ is infinite and moreover, each process that immigrates is issued with zero initial mass. One may therefore think of $\mathbf{n}$ as a process of continuous immigration. In contrast, when $\nu$ is a non-zero measure, processes that immigrate under $\mathbf{m}$ have strictly positive initial mass and therefore contribute to path discontinuities of $||X||$.

Now, for each $t \geq 0$, we define

$$\Lambda_t = X'_t + X_t^{(\mathbf{n})} + X_t^{(\mathbf{m})}, \tag{11}$$

where $\{X'_t : t \geq 0\}$ is an independent copy of $(X, \mathbb{P}_\mu)$,

$$X_t^{(\mathbf{n})} = \sum_{s \leq t : \mathbf{n}} X_{t-s}^{\mathbf{n},s}, t \geq 0 \quad \text{and} \quad X_t^{(\mathbf{m})} = \sum_{s \leq t : \mathbf{m}} X_{t-s}^{\mathbf{m},s}, t \geq 0.$$

In the last two equalities, we understand the first sum to be over times for which $\mathbf{n}$ experiences points and the second sum is understood similarly. Note that since the processes $X^{(\mathbf{n})}$ and $X^{(\mathbf{m})}$ are initially zero valued, it is clear that since $X'_0 = \mu$, then $\Lambda_0 = \mu$. In that case, we use the notation $\tilde{\mathbb{P}}^\lambda_{\mu,x}$ to denote the law of the pair $(\Lambda, \xi)$. Note also that the pair $(\Lambda, \xi)$ is a time-homogenous Markov process. We are interested in the case that the initial position of the spine $\xi$ is randomised using the measure $\mu$ via (10). In that case, we shall write

$$\tilde{\mathbb{P}}^\lambda_\mu(\cdot) = \frac{1}{\langle \mathrm{e}^{\lambda \cdot}, \mu \rangle} \int_{\mathbb{R}} \mathrm{e}^{\lambda x} \mu(\mathrm{d}x) \tilde{\mathbb{P}}^\lambda_{\mu,x}(\cdot)$$

for short. The next theorem identifies the process $\Lambda$ as the *pathwise* spine decomposition of $(X, \mathbb{P}^\lambda_\mu)$, and in particular, it shows that as a process on its own, $\Lambda$ is Markovian.

**Theorem 2.1 (Theorem 5.1, [13]).** *For all $\mu \in \mathcal{M}_F(\mathbb{R})$ such that $\langle \mathrm{e}^{\lambda \cdot}, \mu \rangle < \infty$, $(X, \mathbb{P}^\lambda_\mu)$ and $(\Lambda, \tilde{\mathbb{P}}^\lambda_\mu)$ are equal in law.*

## 3 Proof of Theorem 1.1

First, note that when $\mu \in \mathcal{M}_F(\mathbb{R})$, the assumption $\langle e^{\lambda p \cdot}, \mu\rangle < \infty$ implies $\langle e^{\lambda \cdot}, \mu\rangle < \infty$. From the last section, we have the following spine decomposition of the martingale (4),

$$Z_t^{\lambda}(\lambda) = Z_t'(\lambda) + \sum_{s \le t:\mathbf{n}} \mathrm{e}^{\lambda c_\lambda s} Z_{t-s}^{\mathbf{n},s}(\lambda) + \sum_{s \le t:\mathbf{m}} \mathrm{e}^{\lambda c_\lambda s} Z_{t-s}^{\mathbf{m},s}(\lambda), \tag{12}$$

where $Z'(\lambda)$ is an independent copy of $Z(\lambda)$ under $\mathbb{P}_\mu$,

$$Z_{t-s}^{\mathbf{n},s} := \mathrm{e}^{\lambda c_\lambda (t-s)} \langle \mathrm{e}^{\lambda \cdot}, X_{t-s}^{\mathbf{n},s}\rangle,$$

and

$$Z_{t-s}^{\mathbf{m},s} := \mathrm{e}^{\lambda c_\lambda (t-s)} \langle \mathrm{e}^{\lambda \cdot}, X_{t-s}^{\mathbf{m},s}\rangle.$$

Since $\{Z_t(\lambda), t \ge 0\}$ is a martingale and we assume that $p \in (1,2]$, then Doob's submartingale inequality tells us that $Z(\lambda)$ converges in $L^p(\mathbb{P}_\mu)$ as soon as we can show that $\sup_{t\ge 0} \mathbb{E}_\mu(Z_t(\lambda)^p) < \infty$. To this end, and with the above pathwise spine decomposition in hand, we may now proceed to address the analogue of the proof for branching Brownian motion given in [10].

First note that, for all $p \in (1,2]$,

$$\mathbb{E}_\mu(Z_t(\lambda)^p) = \langle e^{\lambda \cdot}, \mu\rangle \mathbb{E}_\mu^{\lambda}(Z_t(\lambda)^q) = \langle e^{\lambda \cdot}, \mu\rangle \tilde{\mathbb{E}}_\mu^{\lambda}(Z_t^{\lambda}(\lambda)^q), \text{ for all } t \ge 0, \tag{13}$$

where $q = p - 1$. Now let $m = \{m_t : t \ge 0\}$ where for $t \ge 0$, $m_t = ||X_0^{\mathbf{m},t}||$. In particular, note that the process $\{m_t : t \ge 0\}$ is a Poisson point process on $(0,\infty)^2$, independent of $\xi$, with intensity $\mathrm{d}t \times r\nu(\mathrm{d}r)$. By Jensen's inequality, we have that, for all $q \in (0,1]$

$$\begin{aligned}
&\tilde{\mathbb{E}}_\mu^{\lambda}\left(Z_t^{\lambda}(\lambda)^q \mid \xi, m\right) \\
&\le \left[\tilde{\mathbb{E}}_\mu^{\lambda}\left(Z_t^{\lambda}(\lambda) \mid \xi, m\right)\right]^q \\
&\le \langle \mathrm{e}^{\lambda \cdot}, \mu\rangle^q + \left[\tilde{\mathbb{E}}_\mu^{\lambda}\left(\sum_{s \le t:\mathbf{n}} \mathrm{e}^{\lambda c_\lambda s} Z_{t-s}^{\mathbf{n},s}(\lambda) \middle| \xi\right)\right]^q + \left[\tilde{\mathbb{E}}_\mu^{\lambda}\left(\sum_{s \le t:\mathbf{m}} \mathrm{e}^{\lambda c_\lambda s} Z_{t-s}^{\mathbf{m},s}(\lambda) \middle| \xi, m\right)\right]^q,
\end{aligned} \tag{14}$$

to get the last inequality we have used the fact that $(\sum_i u_i)^q \le \sum_i u_i^q$ with $u_i \ge 0$. On the one hand, recalling from (8) that $\mathbb{N}_{\xi_s}[Z_{t-s}(\lambda)] = \mathbb{E}_{\xi_s}[Z_{t-s}(\lambda)]$, we obtain

$$\tilde{\mathbb{E}}^{\lambda}_{\mu}\left(\sum_{s\le t:\mathbf{n}} \mathrm{e}^{\lambda c_\lambda s} Z^{\mathbf{n},s}_{t-s}(\lambda)\middle|\xi\right) = \int_0^t \mathrm{e}^{\lambda c_\lambda s}\mathbb{N}_{\xi_s}[Z_{t-s}(\lambda)]\mathrm{d}s$$
$$= \int_0^t \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s. \tag{15}$$

On the other hand, we have that

$$\tilde{\mathbb{E}}^{\lambda}_{\mu}\left(\sum_{s\le t:\mathbf{m}} \mathrm{e}^{\lambda c_\lambda s} Z^{\mathbf{m},s}_{t-s}(\lambda)\middle|\xi, m\right) = \sum_{s\le t:\mathbf{m}} \mathrm{e}^{\lambda c_\lambda s}\mathbb{E}_{m_s\delta_{\xi_s}}[Z^{\mathbf{m},s}_{t-s}(\lambda)]$$
$$= \sum_{s\le t:\mathbf{m}} m_s \mathrm{e}^{\lambda(\xi_s + c_\lambda s)} \tag{16}$$

Then, putting (15) and (16) into (14), making use again of the inequality $(\sum_i u_i)^q \le \sum_i u_i^q$ where $u_i \ge 0$ for all $i$, we obtain

$$\tilde{\mathbb{E}}^{\lambda}_{\mu}\left(Z^{\lambda}_t(\lambda)^q \mid \xi, m\right) \le \langle \mathrm{e}^{\lambda\cdot}, \mu\rangle^q + \left(\int_0^t \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s\right)^q + \left(\sum_{s\le t:\mathbf{m}} m_s \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\right)^q$$
$$\le \langle \mathrm{e}^{\lambda\cdot}, \mu\rangle^q + \left(\int_0^\infty \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s\right)^q + \sum_{s\ge 0:\mathbf{m}} m_s^q \mathrm{e}^{q\lambda(\xi_s + c_\lambda s)}. \tag{17}$$

Taking expectations again in (17) gives us that, for all $t \ge 0$,

$$\tilde{\mathbb{E}}^{\lambda}_{\mu}(Z^{\lambda}_t(\lambda)^q) \le \langle \mathrm{e}^{\lambda\cdot}, \mu\rangle^q + \Pi^{\lambda}_{\mu}\left[\left(\int_0^\infty \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s\right)^q\right] + \tilde{\mathbb{E}}^{\lambda}_{\mu}\left(\sum_{s\ge 0:\mathbf{m}} m_s^q \mathrm{e}^{q\lambda(\xi_s + c_\lambda s)}\right). \tag{18}$$

We know that, under $\Pi^{\lambda}_{\mu}$, the process $\xi$ is a Brownian motion with drift $\lambda$. Thus, with respect to the same measure, $\xi_s + c_\lambda s$ is a Brownian motion with drift $\lambda + c_\lambda$ which is strictly negative for $\lambda \in (0, \sqrt{-2\psi'(0+)})$. Note that this latter condition holds in particular under assumption that $p\lambda^2 < -2\psi'(0+)$ and $p > 1$. From Sect. 2 of Maulik and Zwart [14] we can conclude that

$$\Pi^{\lambda}_0\left(\int_0^\infty \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s\right) < \infty,$$

which in turn implies that, for all $q \in (0, 1]$,

$$\Pi^{\lambda}_0\left[\left(\int_0^\infty \mathrm{e}^{\lambda(\xi_s + c_\lambda s)}\mathrm{d}s\right)^q\right] < \infty,$$

and hence

$$\Pi_\mu^\lambda\left[\left(\int_0^\infty e^{\lambda(\xi_s+c_\lambda s)}ds\right)^q\right] = \frac{1}{\langle e^{\lambda\cdot},\mu\rangle}\int e^{\lambda x}\mu(dx)\Pi_0^\lambda\left[\left(\int_0^\infty e^{\lambda(x+\xi_s+c_\lambda s)}ds\right)^q\right]$$
$$= \frac{\langle e^{\lambda p\cdot},\mu\rangle}{\langle e^{\lambda\cdot},\mu\rangle}\Pi_0^\lambda\left[\left(\int_0^\infty e^{\lambda(\xi_s+c_\lambda s)}ds\right)^q\right] < \infty. \tag{19}$$

It remains to prove that the last term in (18) is finite. This can be done by computing the expectation directly. We obtain

$$\tilde{\mathbb{E}}_\mu^\lambda\left(\sum_{s\geq 0:\mathbf{m}} m_s^q e^{q\lambda(\xi_s+c_\lambda s)}\right) = \int_0^\infty ds\int_0^\infty r\nu(dr)r^q\Pi_\mu^\lambda\left(e^{q\lambda(\xi_s+c_\lambda s)}\right)$$
$$= \int_0^\infty ds\int_0^\infty r^p\nu(dr)\frac{1}{\langle e^{\lambda\cdot},\mu\rangle}\int e^{\lambda x}\mu(dx)\Pi_0^\lambda\left(e^{q\lambda(x+\xi_s+c_\lambda s)}\right)$$
$$= \frac{\langle e^{\lambda p\cdot},\mu\rangle}{\langle e^{\lambda\cdot},\mu\rangle}\int_0^\infty r^p\nu(dr)\int_0^\infty \Pi_0^\lambda\left(e^{q\lambda(\xi_s+c_\lambda s)}\right)ds.$$

Note that

$$\Pi_0^\lambda\left(e^{q\lambda(\xi_s+c_\lambda s)}\right) = \exp\{qs\lambda^2+s(q\lambda)^2/2+qs\psi'(0+)-qs\lambda^2/2\}$$
$$= \exp\{qs(p\lambda^2/2+\psi'(0+))\}$$

for all $s\geq 0$. Moreover, this expectation has a negative exponent as soon as $p\lambda^2 < -2\psi'(0+)$. Together with the assumption $\int_0^\infty r^p\nu(dr)<\infty$ we conclude that

$$\tilde{\mathbb{E}}_\mu^\lambda\left(\sum_{s\geq 0:\mathbf{m}} m_s^q e^{q\lambda(\xi_s+c_\lambda s)}\right) < \infty. \tag{20}$$

Finally, from (18) to (20) we get that

$$\sup_{t\geq 0}\tilde{\mathbb{E}}_\mu^\lambda\left(Z_t^\lambda(\lambda)^q\right) < \infty,$$

which, in combination with (13), completes the proof. □

*Remark 3.1.* Following the reasoning in the above proof, one can also adapt the argument to show a slightly different version of Theorem 1.1. Specifically, assume that $p\in(1,2]$, $\int_{[1,\infty)} r^p\nu(dr)<\infty$ and $\lambda^2 < -\psi'(0+)$. Then $Z_t(\lambda)$ converges to $Z_\infty(\lambda)$ in $L^p(\mathbb{P}_\mu)$, for all $\mu\in\mathcal{M}_F(\mathbb{R})$ such that $\langle e^{2\lambda\cdot},\mu\rangle$ are finite.

The way to do this is to split the second Poisson sum on the right-hand side of the first inequality in (17) according to whether the value of $m_s$ exceeds unity or not. One then uses a simple estimate $a^q\leq 1+a$, for $a\geq 0$, to deal with the part of the

Poisson sum that contains the small values of $m_s$. The part of the Poisson sum that contains the large values of $m_s$ is dealt with the same way as before. The details are left to the reader.

**Acknowledgements** The second author would like to thank the University of Bath, where most of this research was done. He also acknowledges the financial support of CONACYT-Mexico grant number 129076. Both authors would like to thank an anonymous referee for their comments on an earlier draft of this chapter.

## References

1. E.B. Dynkin (1991): Branching particle systems and superprocesses, *Ann. Probab.* **19(3)**, 1157–1194.
2. E.B. Dynkin (1991): A probabilistic approach to one class of non-linear differential equations. *Probab. Th. Rel. Fields*. **89**, 89–115.
3. E.B. Dynkin (1993): Superprocesses and Partial Differential Equations. *Ann. Probab.* **21**, 1185–1262.
4. E.B. Dynkin (2002): *Diffusions, Superdiffusions and Partial Differential Equations.* AMS, Providencem R.I.
5. E.B. Dynkin and S.E. Kuznetsov (2004): $\mathbb{N}$-measures for branching Markov exit systems and their applications to differential equations. *Probab. Th. Rel. Fields*. **130**, 135–150.
6. Engländer, J. and Pinsky, R. (1999): On the construction and support properties of measure-valued diffusions on $D \subseteq R^d$ with spatially dependent branching. *Ann. Probab.* **27**, 684–730.
7. S.N. Evans (1993): Two representations of a conditioned superprocess. *Proc. Royal. Soc. Edin.* **123A**, 959–971.
8. P.J. Fitzsimmons (1988): Construction and regularity of measure-valued Markov branching processes. *Israeli J. Math.* **64**, 337–361.
9. J. Engländer and A.E. Kyprianou (2004): Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.* **32**, 78–99.
10. R. Hardy and S.C. Harris (2009): A spine approach to branching diffusions with applications to $\mathcal{L}^p$-convergence of martingales. *Séminaire de Probabilités*, **XLII**, 281–330.
11. Kyprianou, A.E. (2004): Travelling wave solutions to the K-P-P equation: alternitives to Simon Harris' probabilistic analysis. *Ann. Inst. H. Poincaré*. **40**, 53–72.
12. Kyprianou, A.E. (2006): *Introductory lectures on fluctuations of Lévy processes with applications.* Springer.
13. A.E. Kyprianou, R.-L. Liu, A. Murillo-Salas and Y.-X. Ren. (2011): Supercritical super-Brownian motion with a general branching mechanism and travelling waves. *To appear in Ann. Inst. H. Poincaré.*
14. K. Maulik and B. Zwart (2006): Tail asymptotics for exponential functionals of Lévy processes. *Stoch. Proc. Appl.* **116**, 156–177.
15. S. Watanabe (1968): A limit theorem of branching processes and continuous-state branching processes. *J. Math. Kyoto Univ.* **8**, 141–167.

# Index



J. Englander and B. Rider (eds.), *Advances in Superprocesses and Nonlinear PDEs*, Springer Proceedings in Mathematics & Statistics 38,
DOI 10.1007/978-1-4614-6240-8, 