ARTICLE TEMPLATE

Total-Current Population-dependent Branching Processes: Analysis via Stochastic Approximation

Khushboo Agarwal[a] and Veeraruna Kavitha[a]

[a]Industrial Engineering and Operations Reserach, IIT Bombay, 400076, Mumbai, India

ARTICLE HISTORY
Compiled September 21, 2023

ABSTRACT
We consider continuous-time two-type population size-dependent Markov Branching Processes. The offspring distribution can depend on the current (alive) and total (dead and alive) populations. Under finite second-moment conditions, using stochastic approximation technique, we show that the time-asymptotic proportion of the populations either converges to the equilibrium points or infinitely often enters every neighbourhood and exits some neighbourhood of a saddle point of an appropriate ordinary differential equation with a certain probability. We also prove a finite time approximation result for the stochastic trajectory. Further, we analyse branching process with attack and acquisition which captures the competition in online viral markets.

KEYWORDS
ODE approximation; multi-type branching process; total population dependent offspring; attack and acquisition; limit proportion; proportion dependency

AMS CLASSIFICATION
60J80; 60J85; 92D25; 91D30; 62L20

1. Introduction

One considers the study of growth patterns and limit proportions to analyse vector-valued Markov chains that are predominantly transient, like two-type branching processes (BPs) under the super-critical regime (for example, [1, 2]). The limit proportions are important objects and various authors, like in [3–6], have analysed them in the context of BPs. The literature mainly considers offspring that depends only on the current (alive) population; such models are essential in several biological applications (for example, [7, 8]). However, none of the said papers study the dependency on total population, i.e., the dead plus alive population. Recently, authors in [9, 10] introduced total-population dependent BPs; in particular, in [9], we study the content propagation over online social networks (OSNs) where the expected number of new shares decline with the total shares so far; in [10], the authors analyse the negative impacts of growing population over non-renewable resource consumption. However, the two said papers analyse the BPs which shift from the super-to-sub critical regime based on total-population sizes, while we are interested in throughout-super-critical BPs (introduced in the immediate next). To the best of our knowledge, no other work considers such total-population

The work of first author is partially supported by Prime Minister's Research Fellowship, India.
CONTACT Khushboo Agarwal. Email: agarwal.khushboo@iitb.ac.in

dependency; in the second part of this work [11], we further consider multiple death types, and apply the results to design warning mechanisms for fake-post detection over OSNs.

In this paper, we precisely investigate the time-asymptotic proportion of population types for a general class of continuous-time two-type population size-dependent Markov BPs. The offspring depends on the current as well as the total populations, and can also be negative to model attack (removal of offspring of another type). We analyse such BPs, when the expected number of offspring produced by any individual is strictly greater than one, for all population sizes, henceforth referred to as throughout-super-critical BPs. We will refer to the proportion of the current population size (of one of the types) as the proportion and the time-asymptotic proportion as the limit proportion.

Limit proportions are crucial objects for many applications. For example, as already mentioned, the second part [11] and [12] design warning mechanisms robust against fake posts propagation, where the control depends on the proportion of posts marked as fake. In [13], we study the relative visibility of advertisement posts defined in terms of the limit proportion of unread copies of posts shared by competing content providers. The limit proportions in prey-predator BP of [14] denote the proportions in which preys and predators co-survive (if at all).

To analyse proportions, it is sufficient to study the embedded chain of the underlying BP. This study is derived using stochastic approximation (SA) techniques (e.g., [15]); we have previously used such an amalgam of SA-based methods in BPs in [9, 12, 13]. In this paper, we *identify a new notion of limiting behavior which we named as hovering around a saddle point, where the stochastic trajectory moves closer to the point and exits a neighbourhood of it infinitely often.* We proved that hovering around saddle points and convergence to attractors and saddle points of an autonomous, non-smooth ordinary differential equation (ODE) almost surely describes the limiting behaviour of proportion. We further prove that the limit set of a single-dimensional ODE suffices to describe the attractor and saddle sets. We also prove that the ODE solution approximates certain normalized trajectories of the current and total population sizes over any finite time window.

Towards the end, we also introduce and analyse a new variant of two-type BP, named BP with attack, where each population can produce offspring of their type, attack other population's offspring and acquire the attacked individuals; the double-sided attack and acquisition feature makes BP with attack very different from the prey-predator BP of [14]. We model the viral competing markets using BP with attack and draw useful insights about the markets. Lastly, we numerically illustrate that the derived approximate trajectories are indeed good approximations to the stochastic BP trajectory.

Related work

There is a vast literature related to BPs, however here we discuss few relevant strands.

Irreducible population-dependent BP with discrete and continuous-time framework are considered in [5, 16] respectively; they do not consider total population-dependent offspring; further, the population-dependent mean matrix converges to a constant mean matrix, but we support proportion-dependent mean matrix even at the limit. In [17], authors consider continuous-time, but population-independent, irreducible BPs.

In [14], the prey-predator BP is analysed in discrete-time setting and co-survival conditions are identified, but the limit proportion is not derived; they also do not consider population-dependency. There are some more papers on prey-predator BP (like [18, 19]) but none of them also consider population-dependency and do not derive limit proportion. In Section 6, we consider a continuous-time population-dependent BP with double-sided attack and acquisition. One can also analyse the continuous-time population-dependent variant of prey-predator BP using our results, as illustrated in arXiv version [20] of this paper.

The Pólya urn literature is closely related to BP literature, as it is shown in [21] that the Pólya urn models can be embedded into a continuous-time population-independent BP. Thus, the asymptotic analysis of the continuous-time BPs can be derived using the corresponding analysis of the Pólya urn models. In fact, SA based approach has been used in the Pólya urn literature to investigate limit proportions of the balls of a specific colour (see, for example, [21–24]). However, the urn-based literature majorly deals with non-extinction scenarios and considers dependency on the current number of balls (not total) in the urn. While, in BPs, extinction occurs with non-zero probability, even in the super-critical regime. In [23], which is an exception, the possibility of extinction is considered, but they do not consider population-dependency. Further, to the best of our knowledge, no finite time approximation trajectories exist for Pólya urn-based models.

In [22], the authors analyse the urn model with the removal of balls of other colours (not the chosen one). This is similar to the negative offspring in BP with attack, where deletion of offspring (attack) from a population type and addition of the same to the other type (acquisition) occurs, in addition to the production of offspring of own type. However, [22] assume a unique attractor for ODE and a constant number of additions (offspring) to the urn. We again have a significant generalization with a random number of offspring and where the random trajectory of the BP with attack can converge to or hover around one of the attractors/saddle points of ODE.

In all, our paper significantly generalizes the models not only in the BP literature but also in the Pólya urn literature by including (total and current) population dependency and negative offspring. In the arXiv version [20] of this paper, we further consider a variety of existing and new BPs to illustrate the generality of our result.

In [13], we introduce BP with attack and provide limit proportion for the case with population-independent and symmetric offspring. We significantly generalize this by considering total population-dependency and symmetric/asymmetric offspring. We analyse a particular case of proportion-dependent BP (offspring depend on the proportion of the populations) along with other co-authors in [12]. Our results cover the model in [12] and in fact one can study more generalized models, which is considered in the second part ([11]).

Organization: The problem is described in Section 2. In section 3, we discuss the SA based details. The main result and its proof is provided in Section 4. The ODE analysis is derived in Section 5, while BP with attack and its application are in Section 6. Section 7 discusses numerical examples for finite time approximation.

Notations: For convenience, we refer the random variable and the corresponding sequence by the same symbol when the context is clear, for example, $\boldsymbol{\Upsilon}_n$. We abbreviate infinitely often as i.o. and almost surely as a.s. We also use acronyms like BP, SA and ODE defined in the introduction. For any function f and time τ, let $f(\tau^-) := \lim_{t\uparrow\tau} f(t)$ and $f(\tau^+) := \lim_{t\downarrow\tau} f(t)$.

2. Problem description

We are considering multi-type BPs, see for example, [1]. In particular, we consider two types of populations, denoted by x and y. Let c_0^x, c_0^y be their respective initial sizes. Let $C^x(t)$ and $C^y(t)$ be the *current population* sizes, i.e., the number of alive individuals of x and y-type populations respectively at time t. Define $A^x(t), A^y(t)$ as the *total population* sizes, i.e., the sum of the number of alive and dead individuals of respective population types at time t. Define $\Phi(t) := (C^x(t), C^y(t), A^x(t), A^y(t))$ as the tuple of population sizes and set $(A^x(0), A^y(0)) = (c_0^x, c_0^y)$.

The lifetime of any individual of any type is exponentially distributed with parameter $0 < \lambda < \infty$, i.e., *we consider Markovian BPs.* The time instance at which an individual completes its lifetime is referred to as its "death" time. Consider any $n \geq 1$. Let τ_n be the death time of the n-th individual (of any type) dying among the alive population; let $\tau_0 := 0$. Let $C_n^x := \lim_{t\uparrow\tau_n} C^x(t)$ be the current-population size of x-type population, just before the death-time τ_n. Similarly, define $\Phi_n := (C_n^x, C_n^y, A_n^x, A_n^y)$ and let $S_n^c := C_n^x + C_n^y$ be the sum current population just before τ_n.

The state space of the underlying process is $\mathcal{U} := \{\phi = (c^x, c^y, a^x, a^y) \in (\mathbb{Z}^+)^4 : c^x \leq$

$a^x, c^y \leq a^y\}$. Once the population gets extinct, no births are possible as in classical multi-type BPs, therefore, any state ϕ with $s^c := c^x + c^y = 0$ is an absorbing state. Let $\nu_e := \inf\{n : S^c_n = 0\}$ represent the epoch at which the extinction occurs, with the usual convention that $\nu_e = \infty$, when $S^c_n > 0$ for all n. The embedded process after extinction is extended by defining $\Phi_n := \Phi_{\nu_e}$ and $\tau_n := \tau_{\nu_e}$, for all $n \geq \nu_e$, when $\nu_e < \infty$. Observe here that no two individuals can die at the same time, as for each n, $P(\tau_{n+1} - \tau_n > 0) = 1$, since $(\tau_{n+1} - \tau_n)$ is exponentially distributed.

We assume that offspring are produced only at the death time by an individual as in [1, 5]; basically, no offspring will be produced in between two consecutive death times. Let $\Gamma_{ij}(\phi)$ denote the (random) number of j-type offspring produced by an i-type individual at its death time when the population state is given by ϕ, for $i, j \in \{x, y\}$. If an i-type parent dies at τ_n, the system for any $t \in [\tau_n, \tau_{n+1})$ (in case $n = \nu_e$, then for all $t \geq \tau_n$), can be described as:

$$\begin{aligned} &C^i(t) = C^i_n + \Gamma_{ii}(\Phi_n) - 1, \quad A^i(t) = A^i_n + \Gamma_{ii}(\Phi_n), \\ &C^j(t) = C^j_n + \Gamma_{ij}(\Phi_n), \text{ and } A^j(t) = A^j_n + \Gamma_{ij}(\Phi_n), \text{ for any } i, j \in \{x, y\} \text{ with } j \neq i. \end{aligned} \tag{1}$$

Basically, the sizes of i and j-type populations change by $\Gamma_{ii}(\Phi_n)$ and $\Gamma_{ij}(\Phi_n)$ respectively[2], and the current size (not the total size) of i-type reduces by 1 due to death. Thus, the underlying process is a continuous-time Markov jump process.

Also, observe[3] that the probability of an x-type parent dying at time τ_n is $C^x_n/(C^x_n + C^y_n) =: B^c_n$, when conditioned on $\sigma\{\Phi_n\}$, the sigma-algebra generated by Φ_n; here, B^c_n *is the proportion of x-type population among the current population.* Further, the conditional probability that a y-type parent dies at time τ_n is $1 - \mathrm{B}^c_n$.

In (1), the offspring distribution depends on Φ_n, therefore, we have a total-current population-size dependent BP. In literature, several current-population dependent BPs are analysed, for example, see [7, 8]. However, to the best of our knowledge, only [9, 10] study total-population dependent BPs; as mentioned before, the second part of this work [11] also considers total-current population-size dependency along with multiple types of deaths. For each ϕ, let $\Gamma_{ii}(\phi) \geq 0$ a.s. for each $i \in \{x, y\}$, i.e., each individual produces non-negative offspring of its type. Since $\Gamma_{ij}(\phi) \in \mathbb{Z}$ for $i \neq j$, therefore, any individual can produce either positive or negative (valued) offspring of the other population, depending upon the population-sizes. Negative offspring are used to model attack.

Many applications can be modelled using the dynamics as in (1). We now discuss briefly one application, namely the viral markets on online social networks (OSNs), which requires several features of the above total-current population-dependent BP.

Example - Viral markets

In online social networks, content providers (CPs) share a variety of content. Each content is shared in the form of a post with an initial set of users, called seed users. The seed users receive the post on their timeline, an inverse stack of posts specific to each user. When a seed user reads the post, it may forward the post to its friends/followers depending upon how much it likes the post. This sharing procedure is followed by the recipients and the process continues. The content either gets extinct in the initial phase or gets viral (the copies of the post grow significantly with time).

Now, there is an important feature about content propagation over OSNs — after reading the post, the users most likely lose interest in it forever. Thus, reading the post is analogous to death, while the number of shares by a user is analogous to offspring, when posts from different CPs are modelled as different population-types. Further, unread and total (read + unread) copies are analogous to the current and total population respectively. Since the network is

[2] For each i, j, the distribution of $\Gamma_{ij}(\Phi_n)$ depends on the population size Φ_n, and not on the value of the epoch, τ_n.

[3] At any time, the residual lifetime of x and y-tpe individuals are exponentially distributed with parameters λC^x_n and λC^y_n respectively, by memoryless property. Thus, the first individual to die is x-type w.p. $C^x_n/(C^x_n + C^y_n)$.

closed and some users may share with previous recipients who would not be interested in the post again, therefore, the effective shares (which actually represent the offspring) depend on total copies.

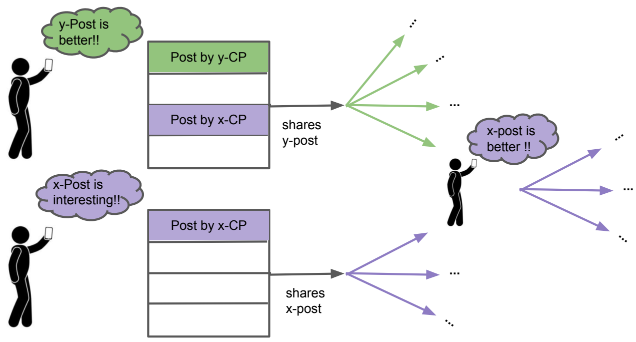

Figure 1. Viral competing markets

On OSNs, contents often compete with each other (e.g., advertisements of similar products). When a new competing post (say y-type) is shared on the user's timeline of an OSN, the user might find y-post more attractive than an older residing x-post on its timeline (see Figure 1). As a result, such a user would not share x-post. This aspect leads to viral competing markets, where we say y-post has attacked and acquired the opportunities of x-post. Such attacks can be modelled as negative offspring.

The contents can also compliment each other. Then, the opportunities of the older y-post may increase - upon seeing the new x-post, the user might get interested in y-post and thus, share both. Such shares can be modelled as positive offspring for both content (population) types. In fact, some advertisement posts may seem complimentary to some users, while competitive to others. Then, the shares/offspring can either be positive or negative.

One can analyse such viral markets using the BP described in (1); more modelling details and analysis of such markets is provided in sections 6 and 7. For now, we proceed towards providing the analysis of the BP (1), and firstly, begin with some preliminary analysis.

2.1. Preliminary analysis

In classical BPs with population-independent offspring ([1]), it is shown that almost surely, the population either grows exponentially at a certain rate time-asymptotically, or gets extinct. *When the BP is super-critical, the probability that the population grows exponentially is non-zero.* For example, for a single x-type BP, [1, Theorem 1, Section 7, Chapter III] shows that $\lim_{t\to\infty} C^x(t)e^{-\lambda(E[\Gamma_{xx}]-1)t}$ exists a.s. and is strictly positive with positive probability when the mean $E[\Gamma_{xx}] > 1$, which hence is the condition for super-criticality. One says that such systems exhibit dichotomy — here, the process either gets extinct or explodes. We show that the total-current population dependent BP exhibits similar dichotomy, when a uniform lower bound $\underline{\Gamma}$ on the population-dependent offspring is "super-critical", i.e., when $E[\underline{\Gamma}] > 1$:

A.1 There exist two integrable random variables $\overline{\Gamma}$ and $\underline{\Gamma}$ which bound the random offspring as: $0 \le \underline{\Gamma} \le \Gamma_{ix}(\phi) + \Gamma_{iy}(\phi) \le \overline{\Gamma}$ a.s., for each ϕ. Also, $E[\overline{\Gamma}^2] < \infty$ and $E[\underline{\Gamma}] > 1$.

Proposition 2.1. [Dichotomy] *Let assumption* (**A.1**) *hold and define* $\underline{m} =: E[\underline{\Gamma}]$*. Then:*

$$P\left(\left\{\liminf_n S_n^c e^{-\lambda(\underline{m}-1)\tau_n} > 0\right\} \cup \left\{\lim_{n\to\infty} S_n^c = 0\right\}\right) = 1.$$

Proof. See Appendix (A). □

In the above result, we could not comment upon the exact rate of explosion, nonetheless, it is clear that the sum current population exhibits dichotomy at limit – either gets extinct or explodes; in the latter case, the explosion is lower bounded by an exponentially growing trajectory with rate $\lambda(\underline{m} - 1)$.

Define $m_{ij}(\phi) := E[\Gamma_{ij}(\phi)]$ for $i, j \in \{x, y\}$ as the mean functions and let $M(\phi) := [m_{ij}(\phi)]_{i,j\in\{x,y\}}$ be the corresponding mean matrix. Then, the assumption (**A.1**) implies that:

$$m_{ix}(\phi) + m_{iy}(\phi) \geq E[\underline{\Gamma}] > 1 \text{ for each } \phi \text{ and } i \in \{x, y\}. \tag{2}$$

One can view the above condition as an equivalent to the super-criticality for our general framework which encompasses irreducible BPs, negative offspring, proportion-dependency and more. For this reason, with slight abuse of notation, we shall call the total-current population-dependent BP to be in throughout super-critical regime if it satisfies (2).

We now proceeds towards the main aim of the paper, the time-asymptotic analysis of the limit proportion ($\lim_{n\to\infty} B_n^c$) of the BPs described via (1). We begin by stating and discussing an important assumption for the limits of the mean functions $m_{ij}(\phi)$:

A.2 There is a function $m_{ij}^{\infty} : [0, 1] \to \mathbb{R}$ (referred to as limit mean function) such that

$$|m_{ij}(\phi) - m_{ij}^{\infty}(\beta^c)| \leq \frac{1}{(s^c)^{\alpha}} \text{ as } s^c = c^x + c^y \to \infty. \tag{3}$$

for each $i, j \in \{x, y\}$ and for some $\alpha \geq 1$, where $\beta^c := \frac{c^x}{c^x + c^y}$.

The limit mean functions are generally assumed to be constant (see [5, 7, 16]). Here we relaxed this restriction. Moreover, the limit mean functions need not be continuous. Later, in Section 6, we show that this assumption holds true in our viral market example.

3. ODE-based stochastic approximation for BPs

When one considers a process which explodes with time, like a typical BP, it is a common practice to scale the process appropriately such that the scaled process converges to a finite limit; this enables the asymptotic study of the rate of explosion, proportions of various components of the process, etc. In Proposition 2.1, it is shown that the sum current population can explode; thus, analysis of BP also requires appropriate scaling. We propose one such scaling in (4) given below, further, the scaling is chosen such that the scaled process can be modelled and analysed using the stochastic approximation (SA) technique (e.g., [15]). Furthermore, as we are primarily interested in studying the limit proportion, $\lim_{t\to\infty} B^c(t)$, it is sufficient to analyse the embedded process, i.e., the discrete-time chain defined at death-times τ_n.

Scaled process:

Analogous to S_n^c, define the sum total population, $S_n^a := A_n^x + A_n^y$. Further, define the following ratios, for $n \geq 1$, towards constructing the scaled process:

$$\mathbf{\Upsilon}_n := (\Psi_n^c, \Theta_n^c, \Psi_n^a, \Theta_n^a), \text{ where } \Psi_n^c := \frac{S_n^c}{n}, \Theta_n^c := \frac{C_n^x}{n}, \Psi_n^a := \frac{S_n^a}{n} \text{ and } \Theta_n^a := \frac{A_n^x}{n}, \tag{4}$$

with $\mathbf{\Upsilon}_0 := (c_0^x + c_0^y, c_0^x, c_0^x + c_0^y, c_0^x)$. Observe that the proportion B_n^c also equals Θ_n^c/Ψ_n^c.

Iterative form:

We now claim that the scaled process in (4) can be written in a form that can be analysed using SA-tools, and proceed towards the same. To this end, we begin with some definitions.

For $n \geq 1$, let $\Gamma_{ij,n}$ be the j-type offspring produced by i-type parent at n-th epoch and let $\Gamma_{i,k} := \Gamma_{ix,k} + \Gamma_{iy,k}$ for $i \in \{x, y\}$. Define

$$H_n := \begin{cases} 1 & \text{if } x\text{-type individual dies at } n\text{-th epoch,} \\ 0 & \text{otherwise,} \end{cases}$$

and $\overline{H}_n := 1 - H_n$. Further, define the function $\mathbf{L}_n := (L_n^{\psi,c}, L_n^{\theta,c}, L_n^{\psi,a}, L_n^{\theta,a})^t$ as follows:

$$\begin{aligned} L_n^{\psi,c} &:= \left\{ H_n \left(\Gamma_{x,n}(\Phi_{n-1}) - 1 \right) + \overline{H}_n \left(\Gamma_{y,n}(\Phi_{n-1}) - 1 \right) \right\} 1_{\Psi_{n-1}^c > 0} - \Psi_{n-1}^c, \\ L_n^{\theta,c} &:= \left\{ H_n \left(\Gamma_{xx,n}(\Phi_{n-1}) - 1 \right) + \overline{H}_n \Gamma_{yx,n}(\Phi_{n-1}) \right\} 1_{\Psi_{n-1}^c > 0} - \Theta_{n-1}^c, \\ L_n^{\psi,a} &:= \left\{ H_n \Gamma_{x,n}(\Phi_{n-1}) + \overline{H}_n \Gamma_{y,n}(\Phi_{n-1}) \right\} 1_{\Psi_{n-1}^c > 0} - \Psi_{n-1}^a, \text{ and} \\ L_n^{\theta,a} &:= \left\{ H_n \Gamma_{xx,n}(\Phi_{n-1}) + \overline{H}_n \Gamma_{yx,n}(\Phi_{n-1}) \right\} 1_{\Psi_{n-1}^c > 0} - \Theta_{n-1}^a. \end{aligned} \tag{5}$$

Using the above definitions, the ratios in $\boldsymbol{\Upsilon}_n$ can be expressed as the following iterative form:

$$\boldsymbol{\Upsilon}_n = \boldsymbol{\Upsilon}_{n-1} + \frac{1}{n} \mathbf{L}_n. \tag{6}$$

The above equation resembles the typical SA-based iterative schemes, where the step-size equals $1/n$ (for example, as in [15]). Taking inspiration from SA-based literature [15], we next suggest an ODE that can approximate the above scheme, which in turn will be instrumental in analysing the limits of the BP.

ODE:

Define $t_n := \sum_{k=1}^{n} \frac{1}{k}$ and let $\eta(t) := \max\{n : t_n \leq t\}$. Let $\Upsilon := (\psi^c, \theta^c, \psi^a, \theta^a)$ be a realisation of $\boldsymbol{\Upsilon}$. Then, the population sizes can be re-written in terms of ratios as:

$$\phi = \phi(\Upsilon, t) := \left(\theta^c \eta(t),\ (\psi^c - \theta^c)\eta(t),\ \theta^a \eta(t),\ (\psi^a - \theta^a)\eta(t) \right). \tag{7}$$

From (5) and (7), the conditional expectation of L_n with respect to the sigma algebra, $\mathcal{F}_n := \sigma\{\Phi_k : 1 \leq k < n\}$, equals:

$$\begin{aligned} &\boldsymbol{\varrho}(\boldsymbol{\Upsilon}_n, t_n) := E[\mathbf{L}_n | \mathcal{F}_n] \text{ for } \boldsymbol{\varrho} = (\rho_\psi^c, \rho_\theta^c, \rho_\psi^a, \rho_\theta^a) \text{ where} \\ &\rho_\psi^c(\Upsilon, t) := \left\{ \beta^c \left(m_{xx}(\phi) + m_{xy}(\phi) \right) + (1 - \beta^c) \left(m_{yy}(\phi) + m_{yx}(\phi) \right) - 1 \right\} 1_{\{\psi^c > 0\}} - \psi^c, \\ &\rho_\theta^c(\Upsilon, t) := \left\{ \beta^c \left(m_{xx}(\phi) - 1 \right) + (1 - \beta^c) m_{yx}(\phi) \right\} 1_{\{\psi^c > 0\}} - \theta^c, \\ &\rho_\psi^a(\Upsilon, t) := \left\{ \beta^c \left(m_{xx}(\phi) + m_{xy}(\phi) \right) + (1 - \beta^c) \left(m_{yy}(\phi) + m_{yx}(\phi) \right) \right\} 1_{\{\psi^c > 0\}} - \psi^a, \text{ and} \\ &\rho_\theta^a(\Upsilon, t) := \left\{ \beta^c m_{xx}(\phi) + (1 - \beta^c) m_{yx}(\phi) \right\} 1_{\{\psi^c > 0\}} - \theta^a. \end{aligned} \tag{8}$$

In ODE-based SA literature, it is anticipated that the ODE constructed using conditional expectation (8) provides the limiting behaviour of the SA-scheme. That is, the limits of (6) are given by the limits of the non-autonomous (time-dependent) ODE $\dot{\Upsilon} = \boldsymbol{\varrho}(\Upsilon, t)$. However, as suggested in [15], one can directly derive an approximation result using a simpler autonomous ODE constructed using the time-limit of the mean functions, and hence that of (8). Thus, in view of the assumption (**A.2**), define $\mathbf{h} := (h^c_\psi, h^c_\theta, h^a_\psi, h^a_\theta)$, where:

$$\begin{aligned}
h^c_\psi(\beta^c) &= \beta^c\bigg(m^\infty_{xx}(\beta^c) + m^\infty_{xy}(\beta^c)\bigg) + (1-\beta^c)\bigg(m^\infty_{yy}(\beta^c) + m^\infty_{yx}(\beta^c)\bigg) - 1,\\
h^c_\theta(\beta^c) &= \beta^c\bigg(m^\infty_{xx}(\beta^c) - 1\bigg) + (1-\beta^c)m^\infty_{yx}(\beta^c),\\
h^a_\psi(\beta^c) &= \beta^c\bigg(m^\infty_{xx}(\beta^c) + m^\infty_{xy}(\beta^c)\bigg) + (1-\beta^c)\bigg(m^\infty_{yy}(\beta^c) + m^\infty_{yx}(\beta^c)\bigg), \text{ and}\\
h^a_\theta(\beta^c) &= \beta^c m^\infty_{xx}(\beta^c) + (1-\beta^c)m^\infty_{yx}(\beta^c),
\end{aligned} \tag{9}$$

and then, under assumptions (**A.1**)-(**A.2**), the autonomous, but non-smooth ODE that can approximate $\mathbf{\Upsilon}_n$ in (6) is as follows:

$$\dot{\Upsilon} = \mathbf{g}(\Upsilon) := \mathbf{h}(\beta^c)1_{\{\psi^c>0\}} - \Upsilon, \tag{10}$$

with the initial condition:

$$\Upsilon(0) \in \mathcal{D}_I := \{\Upsilon \in (\mathbb{R}^+)^4 : \theta^c \le \psi^c \le \psi^a \text{ and } \theta^a \le \psi^a\}. \tag{11}$$

Notice that the above subset of the domain is relevant (see (4)) for systems modelling the BPs. Since right hand side of the above ODE is non-smooth, one may not have the classical solution. Nonetheless, we shall derive the results once the ODE has solution in generalised (or extended) sense[4]. Thus, we next assume the existence of the unique generalised (or extended) solution:

A.3 There exists a unique solution $\Upsilon(\cdot)$ for ODE (10) in the generalised sense over any bounded interval.

The first important result of this paper, Theorem 4.1(ii), focuses on the time-asymptotic limits of the ratios (4), derived via the limits of the ODE (10). Thus, next, we recall the definitions of asymptotically stable and saddle points for autonomous ODE (see [25]), that facilitates the desired a.s. convergence of ratios ($\mathbf{\Upsilon}_n$) before stating the required assumptions - some of the definitions are stated differently to suit our purpose and can also be applied for the cases with generalised solutions of ODE.

Definition 1. A set $\mathbf{E} := \{\Upsilon : \mathbf{g}(\Upsilon) = 0\}$ is called as the set of equilibrium points for the ODE (10).

Define open ball, $N_\epsilon(\mathcal{A}) := \{x : d(x, \mathcal{A}) < \epsilon\}$ for some finite set $\mathcal{A}$.

Definition 2. A subset $\mathcal{A}$ of $\mathbf{E}$ is said to be a (locally) stable set for ODE (10) if for any $\epsilon > 0$, there exists a $\delta > 0$ such that every solution of the ODE $\Upsilon(t) \in N_\epsilon(\mathcal{A})$ for every $t > 0$, if initial condition $\Upsilon(0) \in N_\delta(\mathcal{A})$.

Definition 3. A subset $\mathcal{A}$ of the locally stable set is called an attractor or asymptotically stable set and $\mathcal{D}_\mathcal{A} \subset \mathcal{D}_I$ is the domain of attraction for ODE (10) if every solution $\Upsilon(t) \to \mathcal{A}$ as $t \to \infty$ when $\Upsilon(0) \in \mathcal{D}_\mathcal{A}$.

Let $\mathcal{A}^\complement$ be the complement of $\mathcal{A}$.

Definition 4. A set $\mathcal{S} \subset \mathcal{A}^\complement \cap \mathbf{E}$ is said to be saddle set if there exists $\mathcal{D}_S$ such that $d(\Upsilon(t), \mathcal{A}) \overset{t\to\infty}{\longrightarrow} 0$ for some $\Upsilon(0) \in \mathcal{S}^\complement \cap \mathcal{D}_S$ and $d(\Upsilon(t), \mathcal{S}) \overset{t\to\infty}{\longrightarrow} 0$ for some other $\Upsilon(0) \in \mathcal{S}^\complement \cap \mathcal{D}_S$.

Finally, consider the following subset of $\mathcal{D}_I$, which represents the combined domain of attraction towards $\mathcal{A} \cup \mathcal{S}$:

$$\mathcal{D} := (\mathcal{D}_\mathcal{A} \cup \mathcal{D}_\mathcal{S}) \cap \mathcal{D}_I = \{\Upsilon \in \mathcal{D}_I : \Upsilon(t) \to \mathcal{A} \cup \mathcal{S} \text{ as } t \to \infty, \text{ if } \Upsilon(0) = \Upsilon\}. \tag{12}$$

[4]A function $\Upsilon(\cdot)$ is said to be a generalised (or extended) solution of ODE (10) if it is absolutely continuous and satisfies the equation (10) for almost all $t \ge 0$.

Therefore, if the ODE starts in $\mathcal{D}$, it converges asymptotically to $\mathcal{A} \cup \mathcal{S}$.

All of the above provides the required framework for the main results, and we finally proceed towards the same in the next section.

4. Main results

The first result focuses on the finite-time and time-asymptotic behaviours of the BP trajectory. As said before, we shall make use of the ODE based SA-technique for facilitating the analysis. Classically, the SA results imply the convergence possibility to only the attractors of the ODE, however, in our case, we shall also see convergence to the saddle set ($\mathcal{S}$), under an extra assumption given below:

A.4 Let $\mathcal{A} \cap \mathcal{D}_I$ and $\mathcal{S} \cap \mathcal{D}_I$ be the attractor and saddle set as in Definition 3 and Definition 4 respectively. Consider $\mathcal{D}$ as in (12) and let $\mathcal{D}_b := \mathcal{D} \cap \{\psi^a \leq b\}$, for some $b > 0$, be a compact subset of combined domain of attraction. Assume $p_b := P(\mathcal{V}) > 0$, where $\mathcal{V} := \{\omega : \boldsymbol{\Upsilon}_n(\omega) \in \mathcal{D}_b \text{ i.o.}\}$.

Further, we shall see in the coming that the BP trajectory also exhibits a new limiting behaviour, which we describe next for any stochastic process.

Definition 5. The stochastic process $\boldsymbol{\Upsilon}_n$ hovers around a set $\mathcal{S}$ if $\boldsymbol{\Upsilon}_n \in N_\delta(\mathcal{S})$ i.o., for all $\delta > 0$ and $\boldsymbol{\Upsilon}_n \notin N_{\delta_1}(\mathcal{S})$ i.o., for some $\delta_1 > 0$.

Thus, *when a stochastic process hovers around the set $\mathcal{S}$, it means that its trajectory goes arbitrarily close to the set $\mathcal{S}$ i.o., but still comes out of a neighbourhood of it i.o.* Finally, the result is stated below, and its proof is provided in sub-section 4.2 immediately after highlighting the key points about the stated result.

Theorem 4.1. *Under* (**A.1**)-(**A.3**), *we have:*

(i) For every $T > 0$, a.s. there exists a sub-sequence (n_l) such that:

$$\sup_{k: t_k \in [t_{n_l}, t_{n_l}+T]} d(\boldsymbol{\Upsilon}_k, \Upsilon(t_k - t_{n_l})) \to 0 \text{ as } l \to \infty, \text{ where}$$

$\Upsilon(\cdot)$ *is the extended solution of ODE* (10) *which starts at* $\Upsilon(0) = \lim_{n_l \to \infty} \boldsymbol{\Upsilon}_{n_l}$.

(ii) Further, assume (**A.4**). *Then,* $P(\mathcal{C}_1 \cup \mathcal{C}_2) \geq p_b$, *where*

$$\mathcal{C}_1 := \{\boldsymbol{\Upsilon}_n \to (\mathcal{A} \cup \mathcal{S}) \cap \mathcal{D}_I \text{ as } n \to \infty\}, \text{ and } \mathcal{C}_2 := \{\boldsymbol{\Upsilon}_n \text{ hovers around } \mathcal{S}\}. \quad \square$$

Thus, when BP ($\boldsymbol{\Upsilon}_n$) visits some compact subset of $\mathcal{D}$ i.o., then $\boldsymbol{\Upsilon}_n$ either converges to the attractor set ($\mathcal{A}$) or saddle set ($\mathcal{S}$) or hovers around $\mathcal{S}$, with probability at least $p_b > 0$. We will show that (**A.1**)-(**A.4**) are satisfied for BP with attack in Section 6, with $p_b = 1$, i.e, the above results are true a.s.; proving that $p_b = 1$ requires an important result which is the second main result of the paper and is discussed below.

Second main result

The second main result, Theorem 5.1, is stated in the next section as it requires explanation of more machinery. This theorem considers the generic form of ODE that can approximate stochastic processes, like BPs, and provide simpler one-dimensional conditions to verify (**A.3**) and (**A.4**) by identifying the attractors and saddle sets of ODE (10). It will also be proved that the saddle points are indeed q-attractors (these are special saddle points where convergence, when possible, is exponentially fast, see Definition 7 for details), by exploiting the structure of ODE (10).

4.1. Significance of Theorem 4.1

BP trajectories - Theorem 4.1(i) provides *a novel approach for studying the asymptotic trajectory of the BPs using ODE solution.* Consider the solution of ODE (10) initialised with $\lim_{n_l \to \infty} \mathbf{\Upsilon}_{n_l}$. Then, for large enough n_l and for all k with $t_k \in [t_{n_l}, t_{n_l} + T]$, $\mathbf{\Upsilon}_k$ is sufficiently close to ODE solution $\Upsilon(t_k - t_{n_l})$; and, the approximation improves as n_l further increases. *This result only requires* (**A.1**)-(**A.3**), *and hence is true a.s. for all* $T < \infty$ *and is independent of* p_b *in* (**A.4**).

However, the result only approximates the path of embedded chain with corresponding points of ODE solution, while the time information of the original Markov jump process is not captured.

We suggest and numerically illustrate a better finite-time approximation using a non-autonomous ODE in Section 7, inspired by [9]; the results in [9] capture a saturated total population-dependent BP - the one which does not satisfy (**A.1**) and switches from super-critical to sub-critical regime.

Limit proportion - Theorem 4.1(ii) provides an alternate approach to derive limit behaviour via the attractors or saddle points of ODE (10).

In *extinction paths*, where both populations get extinct, $\mathbf{\Upsilon}_n \to \mathbf{0}$ as $n \to \infty$, say with probability $p_e > 0$. Thus, extinction paths are in the set $\mathcal{V}$ of (**A.4**). While in the survival paths, the BP either converges or hovers around $(\mathcal{A} \cup (\mathcal{S} - \{\mathbf{0}\})) \cap \mathcal{D}_I$, with probability at least $p_b - p_e$. As an example of convergence to a saddle point, the vector $\mathbf{0}$ is a saddle point of ODE (10) (proved in Theorem 5.1) and is well-known to be a limit in extinction paths.

It is now clear that the ratios $\mathbf{\Upsilon}_n$ facilitate the analysis of the underlying BP, however, there are few important points one needs to note regarding the individual population, say about the x-type population: (i) if ratio $\Theta_n^c \nrightarrow 0$, then the x-type population does not get extinct, and the growth rate of C_n^x is at least $O(n)$; (ii) if $\Theta_n^c \to 0$, then it necessarily does not mean that $C_n^x \to 0$; instead, it just implies that $C_n^x = o(n)$. However, by dichotomy proved in Proposition 2.1, the ratio $\Psi_n^c \to 0$ if and only if the sum current population, $S_n^c \to 0$. Thus if the process is irreducible, then $\Theta_n^c \to 0$ if and only if $C_n^x \to 0$.

Population independent to population dependent BPs - One can analyse any general BP with limit mean matrix (say) M^∞ using a population-independent BP with mean matrix as M^∞. The knowledge about limits of the latter BP (if any) can be useful in deriving ODE limits and, thus, the limits of the former BP. One still needs to show that the former BP visits the domain of attraction i.o., given that the latter visits the same i.o.

Limitation - By Theorem 4.1, one can not comment on the individual probability of $\mathbf{\Upsilon}_n$ converging to a particular limit in $(\mathcal{A} \cup \mathcal{S}) \cap \mathcal{D}_I$ or the likelihood of hovering around. Further, $\mathrm{B}_n^c \to \{0, 1\}$ does not always imply the extinction of x or y-type population; however, this is true for the BP with attack considered in Section 6, as proved at the end of Appendix (A).

4.2. Proof of Theorem 4.1

Before we provide the proof, we make an important observation which motivates the derivation of SA-based scheme (6) and also to prove a boundedness assumption for ratios $\mathbf{\Upsilon}_n$ (4) required for most SA-based studies.

<u>**Key idea:**</u> Consider a BP with population-independent and positive offspring, i.e., in (**A.1**), assume $\Gamma_{ix}(\phi) + \Gamma_{iy}(\phi) = \overline{\Gamma}$ for all ϕ and all $i \in \{x, y\}$. Let $\overline{\Pi}_n$ represent the sample mean formed by the sequence of offspring plus the initial population size, i.e.,

$$\overline{\Pi}_n = \frac{1}{n} \left(\sum_{k=1}^{n} \overline{\Gamma}_k + s_0^a \right). \tag{13}$$

By strong law of large numbers, $\overline{\Pi}_n \to \overline{m} := E[\overline{\Gamma}_1]$ a.s. The ratio corresponding to the sum of the total population components of BP given in (1), $\Psi_n^a = {(A_n^x + A_n^y)}/{n}$, for this special case exactly matches with $\overline{\Pi}_n$ till the extinction epoch ν_e (i.e., for all $n < \nu_e$), and beyond that (i.e., for all $n \geq \nu_e$) the population does not change anymore. Thus:

$$\Psi_n^a = \overline{\Pi}_n 1_{n<\nu_e} + \nu_e \overline{\Pi}_{\nu_e}/n 1_{n\geq\nu_e};$$

hence Ψ_n^a converges either to 0 (in extinction paths, i.e., $\nu_e < \infty$) or to $\overline{m}$ (in survival paths); and Ψ_n^c respectively converges to 0 or $\overline{m}-1$ a.s. This observation actually completes the proof for this special case with $\mathcal{A} = \{(0,0), (\overline{m}-1, \overline{m})\}$, further when single population (say x-type) is considered. It is well known that the sample mean (13) *can be written as a SA-based scheme* and the iterative scheme for Ψ_n^c in (6) shows that this is true even for the general case. Further, clearly, (13) becomes an upper bound for all components of $\boldsymbol{\Upsilon}_n$, which *helps in bounding* $\boldsymbol{\Upsilon}_n$ *uniformly in n and a.s.* (see (16) given below), again under (**A.1**).

Analogous to $\overline{\Pi}_n$ as in (13), one can construct a lower bounding sequence using $\underline{\Gamma}$ of (**A.1**); this provides a uniform positive lower bound for Ψ_n^c, which will help the proof. We now return to the main proof with general offspring as in (**A.1**).

Proof: The proof of part (i) has two major steps: (a) to construct a sequence of piece-wise constant interpolated trajectories for almost all sample-paths; (b) to prove that the designed trajectories are equicontinuous in extended sense[5]. These steps are majorly as in [15, Theorems 2.1-2.2], but for the changes required for measurable $\mathbf{g}(\cdot)$. We complete part (i) first.

For part (ii), under (**A.4**), the proof is again inspired from [15] and [26, Theorem 2.3.1, pp. 39], even when the solution of ODE (10) is in generalized sense, not the classical one. However, it requires major changes to incorporate saddle points and hovering around aspect, and is considered next.

Part (i): For many steps of the proof, we will work only with θ^c-component of the vector Υ, when the proof for the remaining components goes through in exactly similar manner.

Let $\boldsymbol{\Upsilon}^n(\cdot) := (\Psi^{n,c}(\cdot), \Theta^{n,c}(\cdot), \Psi^{n,a}(\cdot), \Theta^{n,a}(\cdot))$ be the constant piece-wise interpolated trajectory defined as below (see (6), and recall $t_n = \sum_{i=1}^n \epsilon_{i-1}$ for $\epsilon_i := \frac{1}{i+1}$):

$$\Theta^{n,c}(t) := \Theta_n^c + \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i L_i^{\theta,c}, \text{ for all } t \geq 0, \tag{15}$$

$\Psi^{n,c}(t), \Psi^{n,a}(t)$ and $\Theta^{n,a}(t)$ are defined analogously.

Towards proving equicontinuity, we first consider upper-boundedness of $\boldsymbol{\Upsilon}^n(0) = \boldsymbol{\Upsilon}_n$, as the iterates are trivially lower bounded by 0. The claim is immediately true by strong law of large numbers a.s., to be more precise on the set $\{\overline{\Pi}_n \to \overline{m}\}$, because of the following observation (see (13)-(6)):

$$\Psi_n^c \leq \Psi_n^a \text{ and } \Theta_n^c \leq \Theta_n^a \leq \Psi_n^a = \Pi_n \leq \overline{\Pi}_n \text{ for all } n, \tag{16}$$

and then for any sample path $\omega \in \{\overline{\Pi}_n \to \overline{m}\}$ and $\epsilon > 0$, there exists a $n_\epsilon(\omega) < \infty$,

$$\begin{aligned}\sup_n \max\{\Theta^{n,c}(0), \Psi^{n,c}(0), \Theta^{n,a}(0), \Psi^{n,a}(0)\} \\ \leq \max\left\{\max_{n\leq n_\epsilon(\omega)\}} \max\{\Theta^{n,c}(0), \Psi^{n,c}(0), \Theta^{n,a}(0), \Psi^{n,a}(0)\},\ \overline{m}+\epsilon\right\}.\end{aligned} \tag{17}$$

5

Definition 6. Equicontinuous in extended sense ([15, Equation (2.2), pp. 102]): Suppose that for each n, $f_n(\cdot)$ is an $\mathbb{R}^r$-valued measurable function on $(-\infty, \infty)$ and $(f_n(0))$ is bounded. Also suppose that for each T and $\epsilon > 0$, there is a $\delta > 0$ such that

$$\limsup_n \sup_{0\leq t-s\leq\delta, |t|\leq T} |f_n(t) - f_n(s)| \leq \epsilon. \tag{14}$$

Then the sequence $(f_n(\cdot))$ is said to be equicontinuous in the extended sense.

Towards the second part of equicontinuity (see (14) in footnote 5), the interpolated trajectory for $\Theta^{n,c}(\cdot)$ in (15) can be re-written in "almost integral form", for any $t \geq 0$ (see (8)):

$$\begin{aligned} \Theta^{n,c}(t) &:= \Theta_n^c + \int_0^t \rho_\theta^c(\mathbf{\Upsilon}^n(s), s)ds + \mathcal{E}_1^{n,c}(t), \text{ with the difference term,} \\ \mathcal{E}_1^{n,c}(t) &:= \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i L_i^{\theta,c} - \int_0^t \rho_\theta^c(\mathbf{\Upsilon}^n(s), s)ds. \end{aligned} \tag{18}$$

We further re-write the interpolated trajectory using the autonomous ODE (10):

$$\begin{aligned} \Theta^{n,c}(t) &:= \Theta_n^c + \int_0^t g_\theta^c(\mathbf{\Upsilon}^n(s))ds + \mathcal{E}_1^{n,c}(t) + \mathcal{E}_2^{n,c}(t), \text{ where} \\ \mathcal{E}_2^{n,c}(t) &:= \int_0^t \rho_\theta^c(\mathbf{\Upsilon}^n(s), s)ds - \int_0^t g_\theta^c(\mathbf{\Upsilon}^n(s))ds. \end{aligned} \tag{19}$$

In Appendix (B), we show that $\mathcal{E}_1^{n,c}(t) + \mathcal{E}_2^{n,c}(t)$ converges uniformly to 0, as $n \to \infty$, over any finite time window and further show:

Lemma 4.2. *The sequence $(\mathbf{\Upsilon}^n(\cdot))$ is equicontinuous in extended sense a.s.* □

Now, consider the set N of all sample paths for which $(\mathbf{\Upsilon}^n(\cdot))$ is not equicontinuous - by Lemma 4.2, $P(N) = 0$ (see proof of above Lemma for precise definition of N). Then, by extended version of Arzela-Ascoli Theorem [15, section 4, Theorem 2.2, pp. 127], there exists a sub-sequence $(\mathbf{\Upsilon}^{n_m}(\omega, \cdot))$ which converges to some continuous limit, call $\Upsilon(\omega, \cdot)$, uniformly on each bounded interval for $\omega \notin N$ such that:

$$\Upsilon(t) = \lim_{n_m \to \infty} \Upsilon_{n_m}(\omega) + \int_0^t \mathbf{g}(\Upsilon(s))ds. \tag{20}$$

Thus, for every $\epsilon > 0$ and $T > 0$, there exists $n(\omega, \epsilon, T)$ such that:

$$\sup_{l \in L} d(\mathbf{\Upsilon}_l, \Upsilon(t_l - t_{n_m})) \leq \epsilon/2 \text{ for all } n_m \geq n(\omega, \epsilon, T), \tag{21}$$

where $L := \{l : t_{n_m} \leq t_l \leq T + t_{n_m}\}$; observe for any $l \in L$, $\mathbf{\Upsilon}^{n_m}(t) = \mathbf{\Upsilon}_l$ if $t = t_l - t_{n_m}$. Now, we are left to show that $\Upsilon(\cdot)$ in (20), the solution of the fixed point equation (of the integral operator), is the extended solution of ODE (10) starting at $\Upsilon(0) = \lim_{n_l \to \infty} \mathbf{\Upsilon}_{n_l}$, i.e.,

$$\lim_{h \to 0} \frac{\Upsilon(t+h) - \Upsilon(t)}{h} = \mathbf{g}(\Upsilon(t)) = \frac{d\Upsilon(t)}{dt} \text{ for almost all } t.$$

One can easily show that the function $\mathbf{g} \circ \Upsilon$ is locally integrable, and thus, by [27, Theorem 3.21], the claim holds. This completes part (i).

Part (ii): The proof is constructed for sample paths $\omega \notin N$, however, for simplicity, we drop ω. By (**A.4**), $\mathbf{\Upsilon}_n \in \mathcal{D}_b$ i.o. Since $\mathcal{D}_b$ is compact, $(\mathbf{\Upsilon}_n)$ has a limit point $\Upsilon_0 \in \mathcal{D}_b$; then, there exists a sub-sequence (n_k) such that $\mathbf{\Upsilon}_{n_k} \to \Upsilon_0$. Further, by (extended) equicontinuity of $(\mathbf{\Upsilon}^n(\cdot))$, there exists further sub-sequence (denote it again by (n_k), for simpler notations) $(\mathbf{\Upsilon}^{n_k}(\cdot))$ which converges to the extended solution $\Upsilon(\cdot)$ of the ODE (10) uniformly on each bounded interval. Also observe, $\mathbf{\Upsilon}^{n_k}(0) = \mathbf{\Upsilon}_{n_k} \to \Upsilon_0$, and recall $\Upsilon(0) = \Upsilon_0$ is the initial condition for ODE (10). Under characterization of attractor or q-attractor in (**A.4**), the ODE solution $\Upsilon(t)$ converges to some $\Upsilon^* \in (\mathcal{A} \cup \mathcal{S}) \cap \mathcal{D}_I$ as $t \to \infty$.

We will now show that for any $\delta_1 > 0$, $\mathbf{\Upsilon}_n$ visits $N_{\delta_1}(\Upsilon^*)$ i.o. We will also discuss other convergence aspects to complete the proof. Towards this, fix $\delta_1 > 0$.

Step A: To begin with, assume $\Upsilon^* \in \mathcal{A} \cap \mathcal{D}_I$. Then, by (**A.4**) (local stability) it is possible to choose $0 < \delta_2 < \delta_1$ such that any ODE solution, $\widetilde{\Upsilon}(\cdot)$, satisfies the following:

$$\widetilde{\Upsilon}(t) \in N_{\delta_1}(\Upsilon^*) \text{ for all } t \geq 0, \text{ when initial condition } \widetilde{\Upsilon}(0) \in cl(N_{\delta_2}(\Upsilon^*)). \tag{22}$$

Further, by convergence of solution, $\Upsilon(t) \to \Upsilon^*$, thus there exists $T_{\delta_2} < \infty$ such that:

$$d(\Upsilon(t), \Upsilon^*) < \delta_2/2 \text{ for all } t \geq T_{\delta_2}. \tag{23}$$

Now, following similar steps as in part (i) (see (21)), there exists $\overline{n} < \infty$ such that:

$$\sup_{l \in L_k} d(\boldsymbol{\Upsilon}_l, \Upsilon(t_l)) < \delta_2/2 \text{ for all } n_k \geq \overline{n}, \tag{24}$$

for $L_k := \{l : T_{\delta_2} + t_{n_k} \leq t_l \leq 2T_{\delta_2} + t_{n_k}\}$. Using (23) and (24), for all $n_k \geq \overline{n}$:

$$\sup_{l \in L_k} d(\boldsymbol{\Upsilon}_l, \Upsilon^*) \leq \sup_{l \in L_k} d(\boldsymbol{\Upsilon}_l, \Upsilon(t_l)) + \sup_{l \in L_k} d(\Upsilon(t_l), \Upsilon^*) < \delta_2. \tag{25}$$

Thus, $\boldsymbol{\Upsilon}_n$ visits $N_{\delta_2}(\Upsilon^*)$ i.o., and hence $N_{\delta_1}(\Upsilon^*)$ i.o.

Henceforth, the proof is majorly as in proof of [26, Theorem 2.3.1, pp. 39], except for *few changes to consider convergence to q-attractors, not just attractors.* Contrary to the claim, assume that $\boldsymbol{\Upsilon}_n$ exits $N_{\delta_1}(\Upsilon^*)$ i.o. Thus, by (25), $\boldsymbol{\Upsilon}_n$ moves from $N_{\delta_2}(\Upsilon^*)$ to $\mathcal{D}_b - N_{\delta_1}(\Upsilon^*)$ i.o. Let $\overline{\boldsymbol{\Upsilon}}^0(\cdot)$ be the usual linear interpolated trajectory of $\boldsymbol{\Upsilon}_n$, i.e.,

$$\overline{\boldsymbol{\Upsilon}}^0(t_n) = \boldsymbol{\Upsilon}_n, \text{ and } \overline{\boldsymbol{\Upsilon}}^0(t) = \frac{t_{n+1} - t}{\epsilon_n}\boldsymbol{\Upsilon}_n + \frac{t - t_n}{\epsilon_n}\boldsymbol{\Upsilon}_{n+1} \text{ for } t \in (t_n, t_{n+1}).$$

Then, there exists sequence (l_j, r_j) such that (i) $\cdots > r_j > l_j > r_{j-1} > l_{j-1} > \ldots$, (ii) $r_j \to \infty$, (iii) $\overline{\boldsymbol{\Upsilon}}^0(l_j) \in \partial N_{\delta_2}(\Upsilon^*)$, $\overline{\boldsymbol{\Upsilon}}^0(r_j) \in \partial N_{\delta_1}(\Upsilon^*)$, and (iv) $\overline{\boldsymbol{\Upsilon}}^0(t) \in cl(N_{\delta_1}(\Upsilon^*)) - N_{\delta_2}(\Upsilon^*)$, for all $t \in [l_j, r_j]$. Consider the segments (one for each j) of $\overline{\boldsymbol{\Upsilon}}^0(\cdot)$, i.e., consider functions, $\mathbf{q}_j(t) := \overline{\boldsymbol{\Upsilon}}^0(l_j + t)$ for any $t \geq 0$; observe by construction that for each j, we have $\mathbf{q}_j(t) \in \{\Upsilon : \delta_2 < d(\Upsilon, \Upsilon^*) \leq \delta_1\}$ for all $0 < t \leq r_j - l_j$.

Case (a): Suppose there is a $T < \infty$ such that for some sub-sequence (call it j again) $r_j - l_j \to T$. Now, consider a sub-sequence of $(\mathbf{q}_j(\cdot))$ which (again) converges to some solution of ODE, $\widetilde{\Upsilon}(\cdot)$ uniformly over $[0, T]$.[6] Then, $\widetilde{\Upsilon}(0) \in \partial N_{\delta_2}(\Upsilon^*)$ and $\widetilde{\Upsilon}(T) \in \partial N_{\delta_1}(\Upsilon^*)$. This contradicts (22). For $T = 0$, there is an obvious contradiction.

Case (b): If $r_j - l_j \to \infty$, then, $\widetilde{\Upsilon}(0) \in \partial N_{\delta_2}(\Upsilon^*)$ and $\widetilde{\Upsilon}(t) \in cl(N_{\delta_1}(\Upsilon^*)) - N_{\delta_2}(\Upsilon^*)$ for all $t > 0$. Then, it is a contradiction to Υ^* being an attractor.

In all, $\boldsymbol{\Upsilon}_n \to \Upsilon^*$; since $\Upsilon^* \in \mathcal{A} \cap \mathcal{D}_I$ is arbitrary, we have $\boldsymbol{\Upsilon}_n \to \mathcal{A} \cap \mathcal{D}_I$.

Step S: Now consider $\Upsilon^* \in \mathcal{S} \cap \mathcal{D}_I$. If $\nu_e < \infty$, i.e., in extinction sample paths, $\boldsymbol{\Upsilon}_n \to \mathbf{0}$ and we are done. For others, $\liminf_n \Psi_n^c > 0$ by Lemma 2.1. Thus, with $\nu_e = \infty$ and $\Upsilon^* \in \mathcal{S} \cap \mathcal{D}_I$, by Definition 7, the initial condition $\Upsilon_0 \in \mathbb{S}(\Upsilon^*)$ with $\beta^c(\Upsilon_0) = \beta^c(\Upsilon^*)$.

Similar to step A, by asymptotic stability ((**A.4**)), one can show that (22) follows for any ODE solution $\widetilde{\Upsilon}(\cdot)$ when initial condition $\widetilde{\Upsilon}(0) \in N_{\delta_2}(\Upsilon^*) \cap \mathbb{S}(\Upsilon^*)$. Further, clearly (23)-(25) also hold for this case. Thus, $\boldsymbol{\Upsilon}_n$ visits $N_{\delta_1}(\Upsilon^*) \cap \mathbb{S}(\Upsilon^*)$ i.o.

Further, if for every $\delta_1 > 0$, $\boldsymbol{\Upsilon}_n$ does not exit $N_{\delta_1}(\Upsilon^*) \cap \mathbb{S}(\Upsilon^*)$ i.o., then $\boldsymbol{\Upsilon}_n \to \boldsymbol{\Upsilon}^* \in \mathcal{S} \cap \mathcal{D}_I$. Otherwise, for every $\delta_1 > 0$, $\boldsymbol{\Upsilon}_n$ visits and exits $N_{\delta_1}(\Upsilon^*) \cap \mathbb{S}(\Upsilon^*)$ i.o. □

5. Analysis of proportion ODE

Under (**A.2**), ϕ-dependent mean functions converge to just β^c-dependent mean functions, and thus, one may anticipate that the analysis of $\beta^c(\Upsilon(t)) = \beta^c(t)$ plays a crucial role. In fact, we claim and prove that the time limits of β^c, obtained from the following limit ODE for β^c

[6] The equicontinuity in extended sense can easily be extended to linear interpolated trajectories.

(derived using (10)), leads to the required analysis:

$$\begin{aligned} \dot{\beta^c} &= \frac{1}{\psi^c} g_\beta(\beta^c) 1_{\{\psi^c>0\}}, \text{ where} \\ g_\beta(\beta^c) &:= -\beta^c m^\infty_{xy}(\beta^c) + (1-\beta^c) m^\infty_{yx}(\beta^c) \\ &\quad + \beta^c(1-\beta^c)\Big\{ m^\infty_{xx}(\beta^c) + m^\infty_{xy}(\beta^c) - \big(m^\infty_{yx}(\beta^c) + m^\infty_{yy}(\beta^c)\big)\Big\}. \end{aligned} \tag{26}$$

Further from above, g_β depends only on β^c, thus, *one might expect that the asymptotic analysis of β^c is independent of other components of Υ*. We will see that this is indeed true, and in fact, the asymptotic analysis of all components of Υ can be derived using just g_β.

Before moving towards the analysis of the ODE (26), we first define a special type of saddle points which are attracted exponentially to $\mathcal{S}$ along a particular affine sub-space, and to $\mathcal{A}$ in the remaining space. Such saddle points are facilitated by the virtue of ODE structure in (10).

Definition 7. Any non-zero $\Upsilon^* \in \mathcal{S}$ is said to be (quasi) q-attractor if (i) for any $\Upsilon(0) \in \mathbb{S}(\Upsilon^*) := \{\beta^c(\Upsilon) = \beta^c(\Upsilon^*)\}$, $\Upsilon(t) \overset{t\to\infty}{\longrightarrow} \Upsilon^*$ exponentially, and (ii) $\Upsilon(t) \overset{t\to\infty}{\longrightarrow} \mathcal{A}$ for other initial conditions. Further, if $\Upsilon^* = \mathbf{0} \in \mathcal{S}$, it is called q-attractor if the above happens with $\mathbb{S}(\Upsilon^*) := \{\psi^c = 0\}$.

Now, we define:

Definition 8. Any point $\beta^* \in [0,1]$ is (projected) p-stable if $\mathbf{h}(\beta^*)$ is an attractor for ODE (10); a β^* is called p-saddle if $\mathbf{h}(\beta^*)$ is a saddle point, more specifically, q-attractor.

Under certain conditions, we will show that the attractors of the following one-dimensional ODE are p-stable, while the repellers[7] are p-saddle:

$$\dot{\beta^c} = g_\beta(\beta^c). \tag{27}$$

When β^* is a repeller of (27), we have $g_\beta(\beta^*) = 0$. Thus, when ODE (10) is initialised with $\beta^c(\Upsilon(0)) = \beta^*$, the ODE solution may remain in affine sub-space $\{\beta^c(\Upsilon) = \beta^*\}$ and may converge to $\mathbf{h}(\beta^*)$ (see figure 2). While if $\beta^c(\Upsilon(0)) \neq \beta^*$, one might expect the solution of ODE (10) to repel away from $\mathbf{h}(\beta^*)$, by definition of repeller. These observations indicate that β^* should be p-saddle and we precisely prove the same in our *second important result* below. This result is instrumental in deriving $\mathcal{A}$ and $\mathcal{S}$ using the limit set of ODE (27); see Appendix B for the proof.

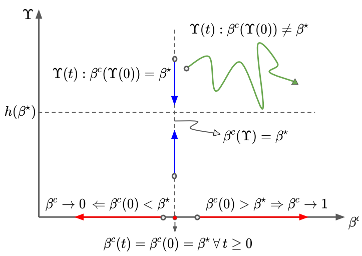

Figure 2. Repeller of (27) leads to saddle point of (10)

Theorem 5.1. *Consider the interval $[0,1]$ such that $g_\beta(0) \geq 0$ and $g_\beta(1) \leq 0$. Let $\mathcal{I} = \{x_i^* : 1 \leq i \leq n\}$ be the set of dis-continuities with $1 \leq n < \infty$ and $\mathcal{J} := \{y_i^* : 1 \leq i \leq m\} \subset \mathcal{I}^{\complement}$ be the set of points with $m < \infty$ ($\mathcal{J}$ is empty when $m = 0$) such that:*

- *(a) $g_\beta(x) = 0$ for each $x \in \mathcal{I} \cup \mathcal{J}$, i.e., $\mathcal{I} \cup \mathcal{J}$ is the set of equilibrium points for (27),*
- *(b) for each $1 \leq i \leq n$, there exists an open/closed/half-open non-empty interval around $x_i^* \in \mathcal{I}$, say $\mathcal{N}_i^*$, such that*
  - *(i) $\cup_{1\leq i\leq n} \mathcal{N}_i^* = [0,1] - \mathcal{J}$ and $\mathcal{N}_i^* \cap \mathcal{N}_j^* = \emptyset$ for $i \neq j$,*
  - *(ii) $g_\beta(\beta) > 0$ for all $\beta \in \mathcal{N}_i^- := \mathcal{N}_i^* \cap [0, x_i^*)$, g_β is Lipschitz continuous on $\mathcal{N}_i^-$,*

[7] Any point $\beta^* \in [0,1]$ is called a repeller of ODE (27) if $g_\beta(\beta^*) = 0$ and $\beta^c(t) \nrightarrow \beta^*$ as $t \to \infty$ when $\beta^c(0) \in \mathcal{N}_\epsilon(\beta^*)$ for some $\epsilon > 0$.

(iii) $g_\beta(\beta) < 0$ *for all* $\beta \in \mathcal{N}_i^+ := \mathcal{N}_i^* \cap (x_i^*, 1]$, g_β *is Lipschitz continuous on* $\mathcal{N}_i^+$.

Then, ODE (10) *satisfies* (**A.3**)*. Further, the set* $\mathcal{I}$ *is an attractor for* (27) *and p-stable for* (10)*; also,* $\mathcal{J}$ *is the set of repellers for* (27) *and p-saddle for* (10)*. Furthermore,* $\mathcal{A} := \{\mathbf{h}(x_i^*) : x_i^* \in \mathcal{I}\}$ *is the attractor set,* $\mathcal{S} := \{\mathbf{h}(y_i^*) : y_i^* \in \mathcal{J}\} \cup \{\mathbf{0}\}$ *is the saddle set in* $\mathcal{D}_I$ *and entire* $\mathcal{D}_I$ *is the combined domain of attraction for* (10). □

The above Theorem can be extended for g_β which is continuous, by standard ODE results, and is considered in [11]. However, we require g_β to be discontinuous for BP with attack (see assumption **K.3** in Section 6), and thus the hypothesis of Theorem 5.1. The last part of the Theorem asserts that the p-stable/p-saddle points are the only attractors/saddle points of ODE (10), other than $\mathbf{0} \in \mathcal{S}$.

6. Branching Process with Attack

Consider a BP with two population types, say x and y. Each individual of any type lives for a random time, $\tau \sim exp(\lambda)$, where $\lambda \in (0, \infty)$. Let τ_n be the death time of the n-th individual which dies first among the alive population. Say an individual of (say) x-type produces $\xi_{xx}(\phi)$ offspring of its type, when the system state is given by ϕ. Further, it attacks/removes $\xi_{xy}(c^y)$ individuals of y-type population; naturally, the attacked population can not exceed the population available to be attacked at the time of death, hence:

K.1 Assume that $\xi_{xy}(c^y) \le c^y$ a.s. and $\xi_{yx}(c^x) \le c^x$ a.s., when the system state is ϕ.

Note that the number of attacks do not depend on the size of the attacking population. The attacked individuals are then deleted from the y-population, and acquired by (i.e., added to) the x-population. No offspring are produced and no individual of opposite population is attacked in between two consecutive death times. Thus, for example, when a x-type individual dies at time τ_n, then the current populations change as follows for any $t \in [\tau_n, \tau_{n+1})$ (in case $\tau_n = \tau_{\nu_e}$, then for all $t \ge \tau_n$):

$$C^x(t) = C_n^x + \xi_{xx}(\Phi_n) + \xi_{xy}(C_n^y) - 1, \text{ and } C^y(t) = C_n^y - \xi_{xy}(C_n^y).$$

The total and y-population also evolve similarly. We call such a BP as *Branching Process with Attack*. The dynamics in (1) capture this BP, when for each i, j:

$$\Gamma_{ii}(\phi) := \xi_{ii}(\phi) + \xi_{ij}(c^j), \text{ and } \Gamma_{ij}(\phi) := -\xi_{ij}(c^j). \tag{28}$$

Next, we assume:

K.2 For each $i \in \{x, y\}$, assume that there exist integrable random variables, $\overline{\xi}$, $\underline{\xi}$, such that $0 \le \underline{\xi} \le \xi_{ii}(\phi) \le \overline{\xi}$ a.s. for each ϕ, $E[\overline{\xi}]^2 < \infty$ and $E[\underline{\xi}] > 1$. Further, let the attack offspring $\xi_{ij}(\phi)$ be integrable for each ϕ and for each $i \ne j \in \{x, y\}$.

The above assumption immediately implies (**A.1**). Define the expectations conditioned on ϕ as $e_{ij}(\phi) := E[\xi_{ij}(\phi)]$ for $i, j \in \{x, y\}$. We further assume (see (28)):

K.3 For $i, j \in \{x, y\}$, let $e_{ij}^\infty \ge 0$ with $e_{xy}^\infty > 0$. Assume $m_{ij}^\infty(\beta^c)$ satisfy the following:

$$m_{xy}^\infty(\beta^c) = -e_{xy}^\infty 1_{\{\beta^c < 1\}}, \; m_{yx}^\infty(\beta^c) = -e_{yx}^\infty 1_{\{\beta^c > 0\}},$$
$$m_{xx}^\infty(\beta^c) = e_{xx}^\infty + e_{xy}^\infty 1_{\{\beta^c < 1\}} \text{ and } m_{yy}^\infty(\beta^c) = e_{yy}^\infty + e_{yx}^\infty 1_{\{\beta^c > 0\}}.$$

Further, assume the conditions of (**A.2**) are satisfied with $\{(m_{ij}, m_{ij}^\infty)\}_{i,j}$ replaced by $\{(e_{ij}, e_{ij}^\infty)\}_{i,j}$.

We are interested in the BP where attack is prominent[8] even at the limit, thus, $e^\infty_{xy} > 0$ without loss of generality in **K.3**. If $e^\infty_{yx} = 0$, then it leads to single-sided attack at limit, but recall anything is possible in transience. Observe the cross-mean function in **K.2** converge to (almost) constant limit, e.g., $e_{xy}(\phi) \stackrel{s^c \to \infty}{\longrightarrow} e^\infty_{xy} 1_{\{\beta^c < 1\}}$. The reason behind the indicator is that there is no attack at limit when $\beta^c = 1$; this is because $C^y_n \to 0$ when $\limsup_{n\to\infty} \beta^c(\mathbf{\Upsilon}_n) = 1$ as proved at the end of Appendix (A).

For BP with attack, the ODE (10) has the following form:

$$
\begin{aligned}
&\dot{\Upsilon} = \mathbf{h}(\beta^c) 1_{\{\psi^c > 0\}} - \Upsilon, \text{ where } \mathbf{h}(\beta^c) := (h^c_\psi, h^c_\theta, h^a_\psi, h^a_\theta), \text{ is such that} \\
&h^c_\psi = \beta^c e^\infty_{xx} + \big(1-\beta^c\big) e^\infty_{yy} - 1, \ h^c_\theta = \beta^c \left(e^\infty_{xx} + e^\infty_{xy} 1_{\{\beta^c<1\}} - 1\right) - \left(1-\beta^c\right) e^\infty_{yx} 1_{\{\beta^c > 0\}}, \quad (29)\\
&h^a_\psi = \beta^c e^\infty_{xx} + \big(1-\beta^c\big) e^\infty_{yy}, \text{ and } h^a_\theta = \beta^c \left(e^\infty_{xx} + e^\infty_{xy} 1_{\{\beta^c<1\}}\right) - \left(1-\beta^c\right) e^\infty_{yx} 1_{\{\beta^c > 0\}}.
\end{aligned}
$$

We begin with the ODE analysis towards providing ODE approximation result for BP with attack using Theorem 4.1.

6.1. Analysis of ODE for BP with attack

Define the parameter vector $\mathbf{e} := \{e^\infty_{ij} : i, j \in \{x, y\}\}$, and consider the following class of limit mean functions (by **K.3**, the vector $\mathbf{e}$ defines M^∞):

$$
\begin{aligned}
&\mathcal{E} := \{\mathbf{e} : e^\infty_{yx} > 0\} \cup \{\mathbf{e} : e^\infty_{yx} = 0 \text{ and } e^\infty_{xx} + e^\infty_{xy} < e^\infty_{yy}\}, \text{ which implies} \quad (30)\\
&\mathcal{E}^\complement = \{\mathbf{e} : e^\infty_{yx} = 0\} \cap \{\mathbf{e} : e^\infty_{yx} > 0 \text{ or } e^\infty_{xx} + e^\infty_{xy} \ge e^\infty_{yy}\} = \{\mathbf{e} : e^\infty_{yx} = 0, e^\infty_{xx} + e^\infty_{xy} \ge e^\infty_{yy}\}.
\end{aligned}
$$

Observe that the first and second sub-classes in $\mathcal{E}$ consider double and single-sided attack, respectively (at the limit); both classes consider acquisition. An important question for a BP with attack is regarding the survival of the individual types and co-survival. Corollary 6.2 of Theorem 4.1 given later provides answers to such questions. Prior to that, the next theorem derives the asymptotic analysis of (27) and also shows that this analysis is sufficient for analysis of (29) (see proof in Appendix B).

Theorem 6.1. *Assume **K.1-K.3**. Then,* (**A.3**) *holds for* (29). *Further, we have:*

(i) For ODE (27)*, no interior $\beta^c \in (0,1)$ is an attractor, $\beta^* = 1$ is always an attractor, but $\beta^* = 0$ is an attractor only if $\mathbf{e} \in \mathcal{E}$.*

Further, again for (27) *in* $[0,1]$*: if $\mathbf{e} \in \mathcal{E}$, then, β^*_r, the unique zero of g_β, is the only repeller; while if $\mathbf{e} \notin \mathcal{E}$, then 0 is the only repeller.*

(ii) The attractors and repellers of ODE (27) *determine the attractor ($\mathcal{A}$) and saddle ($\mathcal{S}$) sets of ODE* (29) *respectively:*

$$
\mathcal{A} = \begin{cases} \{\mathbf{h}(1), \mathbf{h}(0)\}, & \text{if } \mathbf{e} \in \mathcal{E}, \\ \{\mathbf{h}(1)\}, & \text{if } \mathbf{e} \notin \mathcal{E}, \end{cases} \quad \text{and } \mathcal{S} = \begin{cases} \{\mathbf{0}, \mathbf{h}(\beta^*_r)\}, & \text{if } \mathbf{e} \in \mathcal{E}, \\ \{\mathbf{0}, \mathbf{h}(0)\}, & \text{if } \mathbf{e} \notin \mathcal{E}, \end{cases} \quad \text{where}
$$

for example, $\mathbf{h}(1) = (e^\infty_{xx} - 1, e^\infty_{xx} - 1, e^\infty_{xx}, e^\infty_{xx})$ and $\mathbf{h}(0) = (e^\infty_{yy} - 1, 0, e^\infty_{yy}, 0)$.

(iii) The combined domain of attraction of $\mathcal{A} \cup \mathcal{S}$, i.e., $\mathcal{D} = \mathcal{D}_I$ defined in (11). □

6.2. Analysis of random trajectory of BP with attack

By Theorem 4.1, the following holds (proof in Appendix B):

[8] If both $e^\infty_{xy}, e^\infty_{yx} = 0$, then it will lead to two independent (non-attacking) BPs at limit; if required, one can derive the analysis for this case, as done in Theorem 6.1.

Corollary 6.2. *Consider the BP as in* (28)*, and assume* ***K.1****-****K.3****. Then, we have:*

(i) *The assumption* (**A.3**) *holds for ODE* (29)*, and hence Theorem 4.1(i) is applicable.*

(ii) *The following is true w.p.* 1 *for BP with attack:*

- *if* $\mathbf{e} \in \mathcal{E}$*, either* $\boldsymbol{\Upsilon}_n$ *converges to* $\{\mathbf{0}, \mathbf{h}(0), \mathbf{h}(\beta_r^*), \mathbf{h}(1)\}$ *or hovers around* $\{\mathbf{0}, \mathbf{h}(\beta_r^*)\}$*, where* β_r^* *is as in Theorem 5.1 and*
- *if* $\mathbf{e} \notin \mathcal{E}$*, either* $\boldsymbol{\Upsilon}_n$ *converges to* $\{\mathbf{0}, \mathbf{h}(0), \mathbf{h}(1)\}$ *or hovers around* $\{\mathbf{0}, \mathbf{h}(0)\}$*.* □

Recall from Theorem 6.1, ODE for (29) has three types of saddle points: $\mathbf{h}(0)$ when $\mathbf{e} \notin \mathcal{E}$, $\mathbf{h}(\beta_r^*)$ when $\mathbf{e} \in \mathcal{E}$ and vector $\mathbf{0}$ for all cases. The sample paths in which BP hovers around $\mathbf{0}$ or $\mathbf{h}(0)$ or converges to/hovers around $\mathbf{h}(\beta_r^*)$ indicate co-survival. Both populations survive in insignificant numbers in the first case, x-population is comparatively small in the second case and both populations survive in large numbers in the last case. Further, only x or y-population survives when the process converges to $\mathbf{h}(1)$ or $\mathbf{h}(0)$ respectively, see the end of Appendix (A).

We re-iterate that our approach does not provide the probability with which BP converges or hovers around different limit points of the ODE (29).

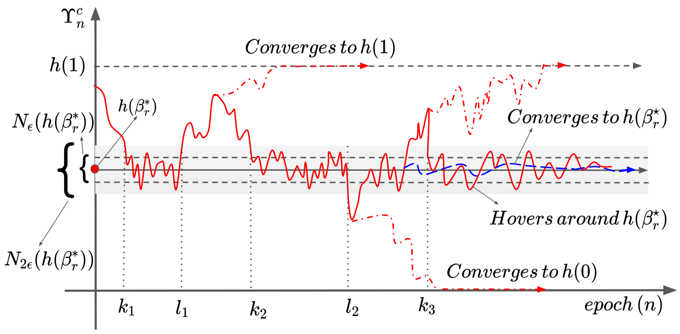

Figure 3. Behavior of BP with attack trajectory when $\mathbf{e} \in \mathcal{E}$

Now, we would like to explain the behaviour of the BP with a pictorial representation in figure 3. Consider $\mathbf{e} \in \mathcal{E}$ and survival paths. Say, the process enters ϵ-neighbourhood of $\mathbf{h}(\beta_r^*)$ at epochs say $k_1, k_2, \ldots$ (for some $\epsilon > 0$), remains in its 2ϵ-neighbourhood for some epochs and then exits at epochs $l_1, l_2, \ldots$ At every exit, it can either get attracted to $\mathbf{h}(0)$ or $\mathbf{h}(1)$ or it can re-enter the neighbourhood. The solid red line in the figure represents the sample path when the trajectory enters and exits the ϵ-neighbourhood i.o., i.e., hovers around $\mathbf{h}(\beta_r^*)$ with $\delta_1 = 2\epsilon$. Some sample paths can converge to $\mathbf{h}(\beta_r^*)$ - see blue dashed line. Similar behaviour is exhibited when $\mathbf{e} \notin \mathcal{E}$. Such hovering around is also observed in [28], where switching between super-to-sub critical regimes occurs due to current-population dependency.

6.3. Analysis of Viral competing markets

In section 2, we briefly discussed the viral markets with content providers (CPs) competing for the propagation of similar posts over OSNs. Recall that the posts from respective CPs (say x and y) can be viewed as x and y-type populations. The unread and unread plus read copies of each post respectively correspond to the current and total population of the respective population-type.

Whenever a new user reads the posts, and shares the preferred post (say x-post), it creates new unread copies of the x-post. Now, some of the copies of the post are shared to the friends who either (i) have not received the other competing post, or (ii) have already received the competing y-post before, or (iii) have received the x-post before. In the first case, copies of the post are referred to as offspring, ξ_{xx}, of the x-type population. While in the second case, the copies of the post are denoted as ξ_{xy}. One can view these copies as the ones where the friends of the user will not prefer the older residing y-post, as it is deep down on their timelines.

The friends of the third kind will not be interested in another copy of the post again, and hence accounts for reduction in the offspring; to be more precise, $e_{xx}(\phi)$ reduces with total population. Thus, the viral competing markets have the attack and acquisition aspect, and can be modelled using the BP with attack (see [13] for modelling details).

In [13], we analysed such markets in a restricted setting, while Corollary 6.2 can handle the generality mentioned here. Both the posts are prominent when the process converges to or hovers around $\mathbf{h}(\beta_r^*)$. While, the convergence to $\mathbf{h}(0)$ or $\mathbf{h}(1)$ represents the dominance of one of the posts.

From Corollary 6.2, one can get more interesting insights. For instance, let y-CP be more influential, and thus y-post is shared more on average in the limit, so $e_{xx}^\infty < e_{yy}^\infty$. If the competition is ignored, the analysis is provided using independent BPs. Such analysis indicates the possibility of co-virality (both posts get viral simultaneously). However, when a typical user receives both posts, it may find x-post more appealing, leading to $e_{xx}^\infty + e_{xy}^\infty > e_{yy}^\infty$ with $e_{yx}^\infty = 0$. Therefore, $\mathbf{e} \notin \mathcal{E}$, thus $\mathbf{h}(1)$ is a limit, which implies that x-post can dominate the post of more influential y-CP. Further, none of the limits indicate co-virality.

On the other hand, when some users prefer the y-post ($e_{yx}^\infty > 0$), while others prefer the x-post, then, co-virality is possible due to interior saddle point $\mathbf{h}(\beta_r^*)$.

7. Finite horizon approximation

In Theorem 4.1(i), we proved the finite-time approximation of $\boldsymbol{\Upsilon}_n$ using the autonomous ODE (10); such an ODE is obtained using the limit proportion-dependent mean functions ($m_{ij}^\infty(\beta^c)$). However, directly using the population-dependent mean functions $m_{ij}(\phi)$, one may anticipate better approximation in transience.

We claim that ODE, $\dot{\Upsilon} = \boldsymbol{\varrho}(\Upsilon, t)$*, constructed using the actual conditional expectation,* $E[\mathbf{L}_n|\mathcal{F}_n] = \boldsymbol{\varrho}(\Upsilon, t)$ *given in* (8) *better approximates the BP*; recall, the difference term $\mathcal{E}_1^n(\cdot)$ of (18) converges to 0 as shown in the proof of Theorem 4.1. The approximation should further improve when the new ODE is initialised with $\boldsymbol{\Upsilon}_{n_m}$, and not with $\lim_{n_m\to\infty} \boldsymbol{\Upsilon}_{n_m}$ as in Theorem 4.1. From (8), the new ODE is non-autonomous and discontinuous. Also by (**A.2**), the right hand side $\boldsymbol{\varrho}(\boldsymbol{\Upsilon}, t)$, converges to that of ODE (10), $\mathbf{g}(\boldsymbol{\Upsilon})$, as $t \to \infty$. Approximation by such non-autonomous ODE is proved for super-to-sub critical total population-dependent BP in [9].

We support our claim using a numerical example; more examples are in [20]. Let $C^x(0) = C^y(0) = 1200$ and let the dynamics be as in BP with attack till S^a is below a certain threshold, and then let the population progress with proportion-dependent mean offspring. Specifically, $M(\phi) = M^t(\phi)1_{\{s^a \le 10^4\}} + M^\infty(\beta^c)1_{\{s^a > 10^4\}}$, where

$$M^t(\phi) = \begin{bmatrix} 4 & -\min(2, c^y) \\ -\min(1, c^x) & 2.2 \end{bmatrix} \text{ and } M^\infty(\beta^c) = \begin{bmatrix} 4\beta^c + 1 & 9\beta^c + 1 \\ 8\beta^c + 1 & 2.2\beta^c + 1 \end{bmatrix}.$$

Observe that (2) holds.

One can view the above dynamics to capture the propagation of new posts over OSNs, which advertise about complimentary products. In the initial phase of sharing (when $s^a \le 10^4$), the users are new to both the products, and therefore, some users prefer product advertised in x-post, while others prefer the y-post (see $M^t(\phi)$). That is, the dynamics are as in branching process with attack. But, thereafter, when both posts achieve good response from the users (in terms of likes, shares), the new users may buy both products (see $M^\infty(\beta^c)$). Thus, *the dynamics shift from being foe-type to friend-type.*

We plot one sample path of BP and the corresponding solutions of autonomous and non-autonomous ODEs[9] (for all $n \ge n_m = 100$ and $T = 12$). The current and total populations are in figure 4, while the proportion $\beta^c(\Upsilon_n)$ is in figure 5. From the plots, one can see that the non-autonomous ODE solution (dashed lines) better approximates the random BP trajectory

[9]The ODE trajectories are estimated using the well known Piccard's iterative method (e.g., [25]).

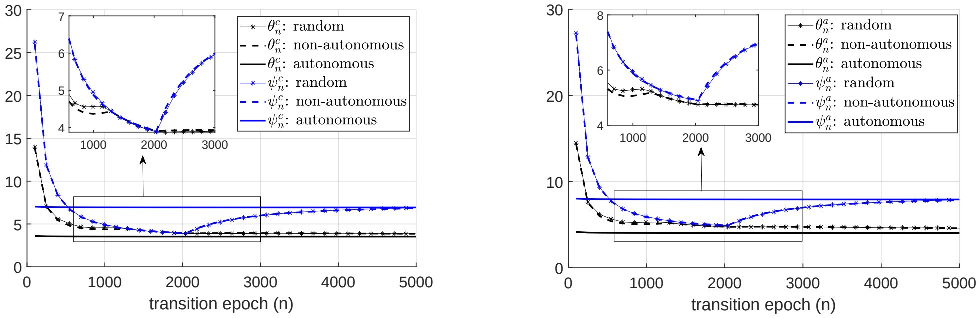

Figure 4. Finite horizon approximation (current on left, and total on right side)

(dotted lines), than the autonomous ODE (solid lines). As seen from the sub-figures, the non-autonomous ODE well captures the transition, unlike ODE (10).

Initially, x-type individuals attack more aggressively than y-type and thus, the y-population depletes faster. In fact, by transition epoch (1300) proportion $\beta_n^c = 1$. Later, $M(\phi) = M^\infty(\beta^c)$ does not have attack component, the y-population is regenerated and β_n^c declines to ≈ 0.51 indicating co-survival. This example also illustrates that the dynamics in transience (here, BP with attack) does not influence the limiting behaviour.

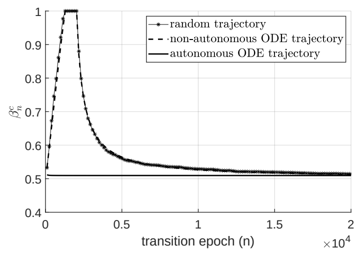

Figure 5. Proportion trajectory, β_n^c

8. Summary and conclusion

We studied time-asymptotic proportion for a class of two-type continuous-time total-current population-dependent Markov BPs. We extended the stochastic approximation result to include the notion of "hovering around the saddle points" of an appropriate ODE and to analyse BPs. The summary to derive the limiting behaviour is:

(s.i) if the BP satisfies the assumption (**A.1**), then the sum current population exhibits dichotomy with probability 1 (see Lemma 2.1);

(s.ii) identify the limit mean functions $m_{ij}^\infty(\beta^c)$ satisfying (**A.2**), if required using the discussion in Appendix (A) for BPs with negative offspring or attack;

(s.iii) identify the attractors and repellers of one-dimensional ODE (27);

(s.iv) identify the attractor and saddle sets of ODE (10) using (s.iii) and Theorem 5.1; these provide the limit proportion;

(s.v) Theorem 5.1 also facilitates the proof of assumption (**A.4**) to conclude about limiting behaviour of BP via Theorem 4.1.

Interestingly, the limit proportion of any BP depends only on the limit mean matrix, irrespective of the dynamics in transience. A finite-time approximation result is also provided. We analysed a recently introduced variant of BP with attack and acquisition under significantly more general conditions; such BP captures essential aspects of competing content propagation over online social networks. We could also analyse the transient behaviour of many interesting variants, for example, the foe-to-friend BP where the populations attack (and acquire) in the initial phase and later compliment each other.

Acknowledgements

We are thankful to Prof. K.S. Mallikarjuna Rao (IEOR, IIT Bombay) for his valuable suggestions which helped in greatly improving the writing style of this paper.

References

[1] K. B. Athreya, P. E. Ney, P. Ney, Branching processes, Courier Corporation, 2004.

[2] H. Kesten, B. P. Stigum, Limit theorems for decomposable multi-dimensional galton-watson processes, Journal of Mathematical Analysis and Applications 17 (2) (1967) 309–338.

[3] R. Dhounchak, V. Kavitha, Decomposable branching processes and viral marketing, arXiv preprint arXiv:1907.00160 (2019).

[4] P. Jagers, The proportions of individuals of different kinds in two-type populations. a branching process problem arising in biology, Journal of Applied Probability 6 (2) (1969) 249–260.

[5] F. C. Klebaner, Geometric growth in near-supercritical population size dependent multi-type galton-watson processes, The Annals of Probability 17 (4) (1989) 1466–1477.

[6] B. Klein, P. Macdonald, The multitype continuous-time markov branching process in a periodic environment, Advances in Applied Probability 12 (1) (1980) 81–93.

[7] F. C. Klebaner, On population-size-dependent branching processes, Advances in Applied Probability 16 (1) (1984) 30–55.

[8] A. Y. Yakovlev, N. M. Yanev, Relative frequencies in multitype branching processes, The annals of applied probability 19 (1) (2009) 1–14.

[9] K. Agarwal, V. Kavitha, Saturated total-population dependent branching process and viral markets, in: 2022 IEEE 61st Conference on Decision and Control (CDC), IEEE, 2022, pp. 382–387.

[10] S. Hautphenne, M. Li, A fluid approach to total-progeny-dependent birth-and-death processes, Stochastic Models (2022) 1–24.

[11] K. Agarwal, V. Kavitha, Robust fake-post detection against real-coloring adversaries, Accepted at IFIP Performance 2023, available at "https://www.ieor.iitb.ac.in/files/Performance_RobustFakePostDetection.pdf" (2023).

[12] S. Kapsikar, I. Saha, K. Agarwal, V. Kavitha, Q. Zhu, Controlling fake news by collective tagging: A branching process analysis, IEEE Control Systems Letters 5 (6) (2020) 2108–2113.

[13] K. Agarwal, V. Kavitha, Co-virality of competing content over osns?, in: 2021 IFIP Networking Conference (IFIP Networking), IEEE, 2021.

[14] J. Coffey, W. J. Bühler, The galton-watson predator-prey process, Journal of applied probability 28 (1) (1991) 9–16.

[15] H. Kushner, G. G. Yin, Stochastic approximation and recursive algorithms and applications, Vol. 35, Springer Science & Business Media, 2003.

[16] P. Jagers, Coupling and population dependence in branching processes, The Annals of Applied Probability 7 (2) (1997) 281–298.

[17] K. B. Athreya, Some results on multitype continuous time markov branching processes, The Annals of Mathematical Statistics 39 (2) (1968) 347–357.

[18] G. Alsmeyer, On the galton-watson predator-prey process, The Annals of Applied Probability (1993).

[19] C. Gutiérrez, C. Minuesa, A predator-prey two-sex branching process, Mathematics 8 (9) (2020) 1408.

[20] K. Agarwal, V. Kavitha, New results in branching processes using stochastic approximation, arXiv preprint arXiv:2111.14527 (2021).

[21] K. B. Athreya, S. Karlin, Embedding of urn schemes into continuous time markov branching processes and related limit theorems, The Annals of Mathematical Statistics 39 (6) (1968) 1801–1817.

[22] I. Higueras, J. Moler, F. Plo, M. San Miguel, Central limit theorems for generalized pólya urn models, Journal of applied probability 43 (4) (2006) 938–951.

[23] S. Janson, Functional limit theorems for multitype branching processes and generalized pólya urns, Stochastic Processes and their Applications 110 (2) (2004) 177–245.

[24] W. B. Arthur, Y. M. Ermoliev, Y. M. Kaniovski, Non-linear urn processes: Asymptotic behavior and applications (1987).

[25] L. Piccini, G. Stampacchia, G. Vidossich, Ordinary differential equations in $\mathbb{R}^n$, Applied Mathematical Sciences 39 (1984).

[26] H. J. Kushner, D. S. Clark, Stochastic approximation methods for constrained and unconstrained systems, Vol. 26, Springer Science & Business Media, 2012.

[27] G. B. Folland, Real analysis: modern techniques and their applications, Vol. 40, John Wiley & Sons, 1999.

[28] P. Jagers, F. C. Klebaner, Population-size-dependent, age-structured branching processes linger around their carrying capacity, Journal of Applied Probability 48 (A) (2011) 249–260.

[29] M. Yamazato, Some results on continuous time branching processes with state-dependent immigration, Journal of the Mathematical Society of Japan 27 (3) (1975) 479–496.

[30] K. B. Athreya, P. Jagers, Classical and modern branching processes, Vol. 84, Springer Science & Business Media, 2012.

Appendix A. Limit mean matrices for BPs with negative offsprings

In this Appendix, we state some important auxiliary results, which are also helpful in further understanding of the subject at hand. The Proposition 2.1 and the discussion thereafter provide insights into the derivation of the limit mean matrices of (**A.2**).

Proof of Proposition 2.1: Let $C^x(0) = c_0^x$ and $C^y(0) = c_0^y$. Consider a fictitious single (z) type population-independent BP. Let $Z(0) = c_0^x + c_0^y$. Each time an individual dies in the new process, assume that random number of offspring, distributed as $\underline{\Gamma}$ in (**A.1**), are produced. Further, assume that exactly 1 individual is immigrated into the new system, if $Z(t) = 0$ for some $t < \infty$. Thus, $Z(\cdot)$ is a classical continuous time branching process with state-dependent immigration as in [29]. Observe $\sum_{j=2}^{\infty} jP(\underline{\Gamma} = j)log(j) < \infty$ due to finite second moment assumption on $\overline{\Gamma}$ in (**A.1**). Thus, by [29, Theorems 6 and 8], $P(Z(t) \to \infty) = 1$, under (**A.1**).

For completing the proof, we couple the embedded chains of the two BPs, for all $n \le \nu_e$, where ν_e is the extinction epoch of the given system. If $\nu_e < \infty$, then $S_n^c = 0$ for all $n \ge \nu_e$. Otherwise, by coupling arguments, $S_n^c \ge Z_n$ for all n, and thus $S_n^c \to \infty$ as $n \to \infty$. Further, in the latter case, by [30, Theorem 1, Chapter 1], the growth rate of S_n^c is at least as large as that of Z_n, i.e., $\lambda(\underline{m} - 1)$. □

Limit mean matrices for BPs with negative offspring:

In BPs with negative offspring, in the the survival sample-paths, by Lemma 2.1, $S_n^c \to \infty$. In such cases, one needs to identify the limit mean matrix of (**A.2**). Say $0 < \liminf_{n\to\infty} \beta^c(\boldsymbol{\Upsilon}_n) \leq \limsup_{n\to\infty} \beta^c(\boldsymbol{\Upsilon}_n) < 1$. Then, for such sample-paths, both populations would have exploded, i.e., $(C_n^x, C_n^y) \to (\infty, \infty)$. Hence, there are sufficient number of individuals to be attacked of both types, which results in the saturation of the number of attacks[10]; thus, it is appropriate to consider $m_{xy}^\infty(\beta^c)$ as some constant for all $\beta^c \in (0,1)$, and so is the case with $m_{yx}^\infty(\beta^c)$.

On the other hand, say $\limsup_{n\to\infty} \beta^c(\boldsymbol{\Upsilon}_n) = 1$, then, $\beta^c(\boldsymbol{\Upsilon}_n) = 1$ i.o. This implies $\beta^c(\boldsymbol{\Upsilon}_n) = 1$ for all n large enough, as $\beta^c(\boldsymbol{\Upsilon}_n) = 1$ is an absorbing state for processes with attack, like BP with attack and prey-predator BP. Thus, clearly $m_{xy}^\infty(\beta^c) = 0$ for $\beta^c = 1$. Similarly, $m_{yx}^\infty(\beta^c) = 0$ for $\beta^c = 0$.

Appendix B.

Throughout the Appendix, we will consider the solution of the integral operator as the generalized solution of ODE (10). The fact that these two solutions are equivalent, is proved towards the end of the proof of Theorem 4.1(i).

Proof of Lemma 4.2. By (17), $(\boldsymbol{\Upsilon}^n(0))_n$ is bounded. We will now prove (14) for $(\Theta^{n,c}(t))$; it can be proved analogously for other components of $\boldsymbol{\Upsilon}^n(\cdot)$. Observe from (18) and (19) that the interpolated trajectory can be re-written as:

$$\begin{aligned}
\Theta^{n,c}(t) &:= \Theta_n^c + \int_0^t g_\theta^c(\boldsymbol{\Upsilon}^n(s))ds + \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i L_i^{\theta,c} - \int_0^t g_\theta^c(\boldsymbol{\Upsilon}^n(s))ds \\
&= \Theta_n^c + \int_0^t g_\theta^c(\boldsymbol{\Upsilon}^n)ds + M^{n,\theta,c}(t) + \rho^{n,\theta,c}(t) + D^{n,\theta,c}(t), \text{ where} \\
M^{n,\theta,c}(t) &:= \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i \left(L_i^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i)\right), \qquad \text{(B1)} \\
\rho^{n,\theta,c}(t) &:= \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i g_\theta^c(\boldsymbol{\Upsilon}_i) - \int_0^t g_\theta^c(\boldsymbol{\Upsilon}^n)ds, \text{ and} \\
D^{n,\theta,c}(t) &:= \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i \left(\rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) - g_\theta^c(\boldsymbol{\Upsilon}_i)\right).
\end{aligned}$$

Now, fix $T > 0$ and define the set $S_T^\delta := \{(s,t) : 0 \leq t - s \leq \delta, 0 \leq t \leq T\}$. Then:

$$\begin{aligned}
\sup_{S_T^\delta} |\Theta^{n,c}(t) - \Theta^{n,c}(s)| \leq \sup_{S_T^\delta} \left| \int_s^t g_\theta^c(\boldsymbol{\Upsilon}^n)dr \right| + \sup_{S_T^\delta} \left| M^{n,\theta,c}(t) - M^{n,\theta,c}(s) \right| \\
+ \sup_{S_T^\delta} \left| \rho^{n,\theta,c}(t) - \rho^{n,\theta,c}(s) \right| + \sup_{S_T^\delta} \left| D^{n,\theta,c}(t) - D^{n,\theta,c}(s) \right|. \qquad \text{(B2)}
\end{aligned}$$

To prove our claim, we begin with the first term of (B2). From (10) and (17), $|g_\theta^c(\boldsymbol{\Upsilon})| \leq \hat{m}$ for an appropriate $\hat{m} > 1$, for any $\boldsymbol{\Upsilon}$, and, thus:

$$\left| \int_s^t g_\theta^c(\boldsymbol{\Upsilon}^n)dr \right| \leq \hat{m}(t-s), \text{ so, } \sup_{S_T^\delta} \int_s^t |g_\theta^c(\boldsymbol{\Upsilon}^n)|\, dr \leq \delta\hat{m}.$$

[10]To be realistic, the number of attacks by a single individual should saturate, i.e., for example, $\lim_{c^y\to\infty} m_{xy}(c^y) = m_{xy}^\infty < \infty$. The case with unsaturated attacks in easier to analyse, and one can easily prove for BP with attack that only one of the two population types survives with probability 1.

For the second term of (B2), define $M_n^{\theta,c} := \sum_{i=0}^{n-1} \epsilon_i \left(L_i^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) \right)$. Then, it is easy to prove that $(M_n^{\theta,c})$ is a Martingale with respect to $(\mathcal{F}_n)$. Thus, using Martingale inequality, for each $\mu > 0$ (where, $E_n(\cdot)$ denotes the expectation conditioned on $(\mathcal{F}_n)$):

$$P\left\{ \sup_{m \le j \le n} |M_j^{\theta,c} - M_m^{\theta,c}| \ge \mu \right\} \le \frac{E_n \left| \sum_{i=m}^{n-1} \epsilon_i \left(L_i^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) \right) \right|^2}{\mu^2}.$$

Observe, $E\left[\left(L_i^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) \right) \left(L_j^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_j, t_j) \right) \right] = 0$ for $i < j$. Using this:

$$P\left\{ \sup_{m \le j \le n} |M_j^{\theta,c} - M_m^{\theta,c}| \ge \mu \right\} \le \frac{\sum_{i=m}^{n-1} \epsilon_i^2 E_n \left| L_i^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) \right|^2}{\mu^2}.$$

Note that under (**A.1**) and (17), for some $K > 0$:

$$\sup_n E_n |L_n^{\theta,c} - \rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i)|^2 \le \sup_n E_n \left(\overline{\Gamma}_n - 1 \right)^2 + \sup_n E_n |\rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i)|^2 < K.$$

Thus, for every $n \ge m$:

$$P\left\{ \sup_{m \le j \le n} |M_j^{\theta,c} - M_m^{\theta,c}| \ge \mu \right\} \le \frac{K}{\mu^2} \sum_{i=m}^{\infty} \epsilon_i^2.$$

By first letting $n \to \infty$ (and using continuity of probability), then, letting $m \to \infty$,

$$\lim_{m \to \infty} P\left\{ \sup_{m \le j} |M_j^{\theta,c} - M_m^{\theta,c}| \ge \mu \right\} = 0 \text{ for each } \mu > 0. \tag{B3}$$

Now, by (B3) and continuity of probability, for each $\mu > 0$:

$$P\left\{ \lim_{m \to \infty} \sup_{m \le j} |M_j^{\theta,c} - M_m^{\theta,c}| \ge \mu \right\} = 0. \tag{B4}$$

Let $A_k := \lim_{m \to \infty} \sup_{m \le j} |M_j^{\theta,c} - M_m^{\theta,c}| < 1/k$, then, $P(A_k) = 1$ for each $k > 0$. We further restrict our attention to sample paths $\omega \notin N := (\cap_k A_k)^c \cup \{\overline{\Pi} \nrightarrow \overline{m}\}$. Now, the second term in (B2) is upper bounded by $2 \sup_{t \ge 0} |M^{n,\theta,c}(t)|$. For any $\omega \notin N$:

$$\sup_{t \ge 0} |M^{n,\theta,c}(t)| = \sup_{t \ge 0} |M_{\eta(t_n+t)}^{\theta,c} - M_n^{\theta,c}| = \sup_{j \ge n} |M_j^{\theta,c} - M_n^{\theta,c}|$$

$$\implies \lim_{n \to \infty} \sup_{S_T^\delta} |M^{n,\theta,c}(t)| \le \lim_{n \to \infty} \sup_{\eta(t_n+t) \ge n} |M_{\eta(t_n+t)}^{\theta,c} - M_n^{\theta,c}| < 1/k,$$

where the last inequality holds because we have considered sample paths which are not in N. Letting $k \to \infty$, we get, $M^{n,\theta,c}(\cdot) \to 0$ uniformly on each bounded interval.

For the third term in (B2), observe that when $t = t_k - t_n$ $(k > n)$, $\rho^{n,\theta,c}(t) = 0$. Thus, for any $|t| \le T$ (following similar steps as in first term, and noting $\epsilon_{\eta(t_n+t)} \le \epsilon_n$):

$$|\rho^{n,\theta,c}(t)| = \left| \int_{t_{\eta(t_n+t)} - t_n}^{t} g_\theta^c(\boldsymbol{\Upsilon}^n) ds \right| < \epsilon_n \hat{m}.$$

Thus, $\rho^{n,\theta,c}(\cdot)$ uniformly converges to 0 as $n \to \infty$ on each bounded interval.

For the last term in (B2), we claim that $D^{n,\theta,c}(t)$ also converges to 0 uniformly on each bounded interval in $(0, \infty)$ as $n \to \infty$, for each $\omega \notin N$. Towards this, first consider $\omega \in$

$N^c \cap \{S_n^c \to 0\}$, i.e, extinction paths. Then, $\rho_\theta^c(\boldsymbol{\Upsilon}_i, t_i) = 0$ and $g_\theta^c(\boldsymbol{\Upsilon}_i) = 0$ for all $i > \nu_e$. Thus, trivially $\lim_{n\to\infty} D^{n,\theta,c}(t) = 0$ for all $t \in (0, \infty)$.

Next, consider $\omega \in N^c \cap \{S_n^c \nrightarrow 0\}$; for such sample paths, we first derive a uniform positive lower bound for Ψ_n^c, required to prove the claim. To this end, analogous to $\overline{\Pi}_n$ defined in (13), one can define $\underline{\Pi}_n$ using $\underline{\Gamma}$ given in (**A.1**). Then, following similar steps as before, i.e., using strong law of large numbers and computing as in (17), we get $\Psi_n^a \geq \Psi_n^c \geq \Delta$ for an appropriate $\Delta > 0$, for all $n \geq 1$. Thus, we have for each $i \geq 1$ (see θ^c component of (10), (8) and assumption (**A.2**)):

$$|D_i^{\theta,c}| = |\mathrm{B}_i^c(m_{xx}(\Phi_i) - m_{xx}^\infty(\mathrm{B}_i^c)) + (1 - \mathrm{B}_i^c)(m_{yx}(\Phi_i) - m_{yx}^\infty(\mathrm{B}_i^c))| \leq \frac{2}{S_i^c} = \frac{2}{\Psi_i^c \eta(t_i)} \leq \frac{2}{\Delta i}.$$

This implies that, (recall $\epsilon_i = 1/(i+1)$)

$$|D^{n,\theta,c}(t)| = \left| \sum_{i=n}^{\eta(t_n+t)-1} \epsilon_i D_i^{\theta,c} \right| \leq \sum_{i=n}^{\eta(t_n+t)-1} \frac{2}{\Delta i(i+1)} \leq \sum_{i=n}^{\infty} \frac{2}{\Delta i(i+1)}, \text{ for any } t.$$

Thus, $D^{n,\theta,c}(t)$ uniformly converges to 0 as $n \to \infty$. In all, by (B2) and above analysis, it is clear that for each $T > 0$ and for any $\epsilon > 0$, there exists n_ϵ such that $\sup_{S_T^\delta} |\Theta^{n,c}(t) - \Theta^{n,c}(s)| < \epsilon$ for all $n \geq n_\epsilon$; hence $(\Theta^{n,c}(\cdot))$ is equicontinuous in extended sense. □

Proof of Theorem 5.1. Recall $\beta^c(\Upsilon) := \theta^c/\psi^c$. Consider the initial condition $\Upsilon(0) \in \mathcal{D}_I$ with $\psi^c(0) = 0$, then ODE (10) simplifies to $\dot{\Upsilon} = -\Upsilon$, which clearly has a unique solution and further $\Upsilon(t) \to \mathbf{0}$ as $t \to \infty$. We claim that $\mathbf{0} \in \mathcal{S}$ as we next show that with $\psi^c(0) > 0$, the solution Υ converges to other equilibrium points.

Let $\psi^c(0) > 0$, and say without loss of generality, $\beta^c(\Upsilon(0)) \in \mathcal{N}_i^-$ for some i. By Lemma B.3, $\psi^c(t) > 0$ for all $t \geq 0$, thus ODE (10) simplifies to $\dot{\Upsilon} = \mathbf{h}(\beta^c(\Upsilon)) - \Upsilon$. Consider the following smooth ODE, with initial condition $\Upsilon(0)$ (by (c), the right hand side given below is Lipschitz continuous):

$$\begin{aligned} \dot{\Upsilon} &= f_l^i(\beta^c) - \Upsilon, \text{ where} \\ f_l^i(x) &:= \mathbf{h}(x) 1_{\{x < x_i^*\} \cap N_i^*} + \mathbf{h}_l^* 1_{\{x \geq x_i^*\}} + \mathbf{h}_l^o 1_{x \leq \Delta_l^i}, \text{ with} \\ \mathbf{h}_l^* &:= \lim_{x_n \uparrow x_i^*} \mathbf{h}(x_n),\ \mathbf{h}_l^o := \lim_{x_n \downarrow \Delta_l^i} \mathbf{h}(x_n), \text{ and } \Delta_l^i := \inf\{\beta^c(\Upsilon) : \beta^c(\Upsilon) \in \mathcal{N}_i^*\}. \end{aligned} \tag{B5}$$

Then, by [25, Theorem 1, sub-section 1.4, pp. 6], the above smooth ODE has a unique solution, say $\Upsilon^1(t)$. Let $\tau := \inf\{t : \beta^c(\Upsilon^1(t)) = x_i^*\}$, then by Lemma B.1, $\tau < \infty$. Observe that the solution of the original ODE (10), with the same initial condition $\Upsilon(0)$, coincides with $\Upsilon^1(\cdot)$ for all $t < \tau$, as $\psi^c(t) > 0$ for all $t > 0$ by Lemma B.3 for such initial condition. Now, let $\Upsilon^\tau := \Upsilon^1(\tau)$ and observe $\beta^c(\Upsilon^\tau) = x_i^*$. Using similar logic, one can prove that x_i^* is an attractor for ODE (27). Further, by uniqueness of the solutions of the smooth[11] ODEs, the solution of ODE (10) for $t > \tau$ is given by:

$$\Upsilon^2(t) = (\psi^c(t), x_i^* \psi^c(t), \psi^a(t), \theta^a(t)), \tag{B6}$$

where the three components of $\Upsilon^2(\cdot)$, defined as $\Omega(\cdot) := (\psi^c(\cdot), \psi^a(\cdot), \theta^a(\cdot))$ is the solution of the following initial value problem (IVP) for all $t \geq \tau$ (see (10)):

$$\dot{\Omega} = \mathbf{h}_i - \Omega, \text{ with } \Omega(\tau) := \Omega(\Upsilon^*), \text{ where constant, } \mathbf{h}_i := (h_\psi^c, h_\psi^a, h_\theta^a)|_{x_i^*}. \tag{B7}$$

Observe that $\beta^c(t) = x_i^*$ for all $t > \tau$ by (a). With this, $\Upsilon(t) := \Upsilon^1(t) 1_{t<\tau} + \Upsilon^2(t) 1_{t>\tau}$ is the unique solution, which satisfies ODE (10) for all $t \neq \tau$, and with initial condition $\Upsilon(0)$. Thus, (10) satisfied (**A.3**). Clearly from (B7),

$$\Upsilon(t) \to \mathbf{h}(x_i^*), \text{ where } \mathbf{h}(x_i^*) = (h_\psi^c, x_i^* h_\psi^c, h_\psi^a, h_\theta^a)|_{x_i^*}.$$

[11]The ODEs (B5) and (B7) are the two smooth ODEs.

Similarly, one can show that $\Upsilon(t) \to \mathbf{h}(x_i^*)$, if $\beta^c(\Upsilon(0)) \in \mathcal{N}_i^+$.

Thus, $\mathbf{h}(x_i^*)$ is an attractor for ODE (10), with domain of attraction as $\mathcal{D}_i := \{\Upsilon \in \mathcal{D}_I : \beta^c(\Upsilon) \in \mathcal{N}_i^*\} \cap \{\psi^c > 0\}$. Since $x_i^* \in \mathcal{I}$ is arbitrary, $\mathcal{A} = \{\mathbf{h}(x_i^*) : x_i^* \in \mathcal{I}\}$, with corresponding domain of attraction as $\mathcal{D}_{\mathcal{A}} = \cup_{1 \le i \le n} \mathcal{D}_i$. Also, $\mathcal{I}$ is an attractor for (27).

By hypothesis (b.i), any initial condition $\Upsilon(0)$ with $\beta^c(\Upsilon(0)) \in [0,1] - \mathcal{J}$ is already considered above. Now consider $\Upsilon(0)$ with $\beta^c(\Upsilon(0)) = y_i^* \in \mathcal{J}$, i.e., $\Upsilon(0) \in \mathbb{S}(h(y_i^*))$. Then, the analysis follows as in (B6)-(B7) to show that $\Upsilon(t) \to \Upsilon(y_i^*)$ as $t \to \infty$; the exponential convergence is clear from ODE (B7). This proves that $\mathbf{h}(y_i^*)$ is a saddle point for ODE (10). Clearly, by (a), (b.ii)-(b.iii), $y_i^* \in \mathcal{J}$ is a saddle point for ODE (27). Hence, the theorem follows, as similar things are true for $\mathbf{0}$. □

Lemma B.1. *The time τ defined in the proof of Theorem 5.1 is finite.*

Proof. By hypothesis (b), $g_\beta(\cdot) > 0$ and continuous, for all $\beta^c \in \mathcal{N}_i^-$. Further, x_i^* is a point of discontinuity for g_β and $g_\beta(x_i^*) = 0$; thus $\beta^c(\mathbf{h}_l^*) = \lim_{x_n \uparrow x_i^*} g_\beta(x_n) > 0$ (see (B5)), which implies, $\inf_{\{\beta^c \in \mathcal{N}_i^-\}} g_\beta(\beta^c) > 0$. Observe τ is determined by β^c-component of $\Upsilon^1(\cdot)$, the solution of ODE (B5). From (B5), the latter is a continuous extension of the original ODE (10), thus, the β^c-component of the ODE (B5) can be uniformly lower bounded by $\inf_{\{\beta^c \in \mathcal{N}_i^-\}} g_\beta(\beta^c) > 0$. Thus, by Lemma B.2(a.ii), $\tau < \infty$. □

Lemma B.2. *Consider an initial value problem $\dot{z} = f(z,t)$, with $z(0) \in (z_0^l, z_0^u)$ where f is a measurable function with finitely many discontinuities.*

(a) Say $f(z,t) > 0$, for all $z \in (z_0^l, z_0^u)$ and all t. Then:

- *(i) $z(\cdot)$ is an increasing function of t till $\tau^u := \inf\{t : z(t) \ge z_0^u\}$.*
- *(ii) Say $f(z,t) > \delta$ for some $\delta > 0$, for all $z \in (z_0^l, z_0^u)$ and all t. Then, $\tau^u < \infty$.*

(b) If $f(z,t) < 0$, for all $z \in (z_0^l, z_0^u)$ and all t, then $t \mapsto z(t)$ is a decreasing function till $\tau^l := \inf\{t : z(t) \le z_0^l\}$, and if in addition $f(z,t) < -\delta$ for some $\delta > 0$, for all $z \in (z_0^l, z_0^u)$ and all t, then $\tau^l < \infty$.

Proof. We will provide the proof for part (a), and it can be done analogously for part (b). Contrary to the claim, let $\tau_1 < \tau_2 < \tau^u$ be two time points such that $z(\tau_1) \ge z(\tau_2)$, with $z(\tau_1), z(\tau_2) \in (z_0^l, z_0^u)$. Then, we have:

$$0 \ge z(\tau_2) - z(\tau_1) = \int_{\tau_1}^{\tau_2} f(z(s), s) ds,$$

which is a contradiction to the hypothesis. Now if possible, let $\tau_u = \infty$, then $z(t) < z_0^u$ for all t and $t \mapsto z(t)$ is an increasing function (as proved before). Further, since $z(t) = z(0) + \int_0^t f(z(s), s) ds > z(0) + t\delta$, there exists $T_\delta > 0$ such that $z(t) \ge z_0^u$ for all $t \ge T_\delta$, which contradicts $\tau^u = \infty$. □

Lemma B.3. *Let* $(\mathbf{A.2})$ *and* $(\mathbf{A.3})$ *hold. Define*

$$\begin{aligned} \underline{\varepsilon} &:= \inf\{m_{ix}^\infty(\beta^c) + m_{iy}^\infty(\beta^c) : \beta^c \in [0,1], i \in \{x,y\}\}, \text{ and} \\ \overline{\varepsilon} &:= \sup\{m_{ix}^\infty(\beta^c) + m_{iy}^\infty(\beta^c) : \beta^c \in [0,1], i \in \{x,y\}\}. \end{aligned} \tag{B8}$$

For any $0 < \epsilon < \underline{\varepsilon} - 1$, define $A_\epsilon := [\underline{\varepsilon} - 1 - \epsilon, \overline{\varepsilon} - 1 + \epsilon]$. In case, $\psi^c(0) \in int(A_\epsilon)$ (interior) for some $\epsilon > 0$, then $\psi^c(t) \in A_\epsilon$ for all $t \ge 0$. Thus, if $\psi^c(0) > 0$, then, $\psi^c(t) > \psi^c(0) - \delta$ for all $t \ge 0$ and for any $\delta > 0$.

Proof. Recall from (10), ODE for ψ^c is $\dot{\psi^c} = h_\psi^c(\beta^c) 1_{\psi^c > 0} - \psi^c$. Now, one can lower bound $h_\psi^c(\beta^c) - \psi^c$ as, for all t (by $(\mathbf{A.1})$ and (B8)):

$$h_\psi^c(\beta^c) - \psi^c \ge \beta^c \underline{\varepsilon} + \big(1 - \beta^c\big)\underline{\varepsilon} - 1 - \psi^c = \underline{\varepsilon} - 1 - \psi^c. \tag{B9}$$

It is easy to observe (by Weierstrass Theorem) that there exists a strict positive uniform lower bound l_I for any closed interval $I \subset (0, \underline{\varepsilon} - 1)$ as below:

$$\dot{\psi}^c \geq l_I > 0 \text{ when } \psi^c \in I \text{ and for all } t. \tag{B10}$$

For the first part, in the above, consider $I = [\psi^c(0), \underline{\varepsilon} - 1 - \frac{\epsilon}{2}]$, where $\psi^c(0) \notin A_\epsilon$. By Lemma B.2(a), we have $\tau^u := \inf\{t : \psi^c(t) \geq \underline{\varepsilon} - 1 - \frac{\epsilon}{2}\} < \infty$, i.e., $\psi^c(\cdot)$ enters A_ϵ from the left.

We will now explicitly show that $\psi^c(\cdot)$ can not exit A_ϵ, once it enters/starts in it (set $\tau^u = 0$ when $\psi^c(0) \in int(K_\epsilon)$). In contrast, say ψ^c leaves A_ϵ and to the left. Observe $\psi^c(\tau^u) > \underline{\varepsilon} - 1 - \epsilon$. For ψ^c to exit A_ϵ, by continuity of ψ^c (and Intermediate Value Theorem, IVT), there exist $\underline{\varepsilon} - 1 - \epsilon < \underline{\nu} < \overline{\nu} < \underline{\varepsilon} - 1$ such that for some $t_2 > t_1 > \tau^u$, $\psi^c(t_2) = \underline{\nu}$ and $\psi^c(t_1) = \overline{\nu}$. Then, by MVT, we have:

$$\dot{\psi}^c(s) = \frac{\psi^c(t_2) - \psi^c(t_1)}{t_2 - t_1} = \frac{\underline{\nu} - \overline{\nu}}{t_2 - t_1} < 0,$$

for some $s \in (t_1, t_2)$. This is a contradiction as $\dot{\psi}^c(t) > 0$ for $\psi^c \in (0, \underline{\varepsilon} - 1)$ and any t. Conclusively, ODE solution $\psi^c(\cdot)$ enters A_ϵ from left when $\psi^c(0) < \underline{\varepsilon} - 1 - \epsilon$, and does not exit A_ϵ from left.

Similarly from (10), $h^c_\psi(\beta^c) - \psi^c$ can be upper bounded as (by (**A.1**) and (B8)):

$$\mathbf{h}^c_\psi(\beta^c) - \psi^c \leq \beta^c \overline{\varepsilon} + \big(1 - \beta^c\big)\overline{\varepsilon} - 1 - \psi^c = \overline{\varepsilon} - 1 - \psi^c, \tag{B11}$$

and $\dot{\psi}^c \leq \overline{\varepsilon} - 1 - \psi^c \leq u_I < 0$ for all t and for any $\psi^c \in I$ where $I \subset (\overline{\varepsilon} - 1, \infty)$ is any closed interval. Then, applying similar arguments as above, one can show that $\psi^c(\cdot)$ enters and does not exit A_ϵ from/to right as well. □

Proof of Theorem 6.1. We first study ODE (27), using which we then analyse ODE (10)/(29). Observe by definition of $m^\infty_{xy}(\cdot)$, $m^\infty_{yx}(\cdot)$ in **K.3** that $0, 1$ are equilibrium points of ODE (27). Further, $g_\beta(\beta^c)$ *is convex or concave in only* $(0, 1)$, *respectively if* $m^\infty \leq 0$ *or* ≥ 0, as can be seen from below (see **K.3** for definitions):

$$\begin{aligned} g_\beta(\beta^c) &= \big(-e^\infty_{yx} + \beta^c \widetilde{m}^\infty - (\beta^c)^2 m^\infty\big)\, 1_{\beta^c \in (0,1)}, \text{ where} \\ \widetilde{m}^\infty &:= e^\infty_{xx} + e^\infty_{xy} - e^\infty_{yy} + e^\infty_{yx}, \text{ and } m^\infty := e^\infty_{xx} - e^\infty_{yy}. \end{aligned} \tag{B12}$$

At first by Lemma B.3, $\mathbf{0}$ is a saddle point for ODE (10) and hence for (29). Now, let $m^\infty \geq 0$, and consider the following two sub-cases.

Sub-case 1: $e^\infty_{xy} > 0$ and $e^\infty_{yx} > 0$. Since $g_\beta(\cdot)$ is continuous in $(0, 1)$:

$$g_\beta(0^+) = \lim_{\delta \to 0} g_\beta(\delta) = -e^\infty_{yx} < 0, \text{ and } g_\beta(1^-) = \lim_{\delta \to 0} g_\beta(1 - \delta) = e^\infty_{xy} > 0. \tag{B13}$$

Therefore, there exists a unique zero of g_β, say $\beta^*_r \in (0, 1)$. Further by concavity, $g_\beta(\beta^c) < 0$ when $\beta^c < \beta^*_r$ and $g_\beta(\beta^c) > 0$ when $\beta^c > \beta^*_r$. Thus, the result follows for this case by Theorem 5.1 with $x^*_1 = 0$, $x^*_2 = 1$ and $y^* = \beta^*_r$. That is, $\{0, 1\}$ is the attractor set, $\{\beta^*_r\}$ is the repeller set for ODE (27). Thus, $\mathcal{A} = \{\mathbf{h}(0), \mathbf{h}(1)\}$ is the attractor set and $\mathcal{D} = \{\mathbf{0}, \mathbf{h}(\beta^*_r)\}$ is the saddle set for ODE (10), with combined domain of attraction, $\mathcal{D}$ as in (v) of the Theorem.

Sub-case 2: $e^\infty_{xy} > 0$ and $e^\infty_{yx} = 0$. Observe $e^\infty_{xx} < e^\infty_{yy}$ is not possible here, as it would contradict $m^\infty \geq 0$. Thus, $e^\infty_{xx} \geq e^\infty_{yy}$. Therefore, for any $\beta \in (0, 1)$, $g_\beta(\beta) = \beta(1 - \beta)(e^\infty_{xx} - e^\infty_{yy}) + \beta e^\infty_{xy} > 0$. Further, $g_\beta(1^-) > 0$, as in case 1. Thus, the result follows for this case as well by Theorem 5.1 with $x^*_1 = 1$ and $y^* = 0$.

This completes parts (i) and (ii) for the case when $m^\infty \geq 0$. Analogously, one can prove (i) and (ii) when $m^\infty \leq 0$. Then, the proof is complete using Theorem 5.1. □

Proof of Corollary 6.2. Given limit mean functions as in **K.3**, the assumption (**A.3**) is guaranteed by Theorem 6.1. We now prove the assumption (**A.4**).

$\mathcal{A}$ and $\mathcal{S}$ are the attractor and saddle sets of ODE (29) respectively, with subset of the combined domain of attraction as $\mathcal{D}_I$, as identified in Theorem 6.1. Towards getting a compact sub-domain of $\mathcal{D}_I$, from (16), (28) and **K.2**, one can bound Ψ_n^a as:

$$\Psi_n^a \leq \overline{\Psi}_n^a := \frac{1}{n}\left(\sum_{k=1}^{\min\{\nu_e, n\}} \left(\overline{\xi}_{xx,k} + \overline{\xi}_{yy,k}\right) 1_{\{\Psi_k^c > 0\}} + s_0^c \right).$$

As before, $\overline{\Psi}_n^a \to E[\overline{\xi}_{xx,1} + \overline{\xi}_{yy,1}]$ a.s. in survival paths and $\overline{\Psi}_n^a \to 0$ in extinction paths, as $n \to \infty$. Thus, $\mathcal{D}_b := \mathcal{D}_I \cap \left\{\Upsilon : \psi^a \in [0, E[\overline{\xi}_{xx,1} + \overline{\xi}_{yy,1}]]\right\}$ is the compact subset of $\mathcal{D}_I$ and $p_b := P(\boldsymbol{\Upsilon}_n \text{ visits } \mathcal{D}_b \text{ i.o.}) = 1$. Hence, by Theorem 6.1 and Theorem 4.1(ii), we have $\boldsymbol{\Upsilon}_n \to \mathcal{A} \cup \mathcal{S}$ with probability 1. □